

SERIES IN AFFECTIVE SCIENCE



# Thinking *about* Feeling

CONTEMPORARY  
PHILOSOPHERS  
ON EMOTIONS

*Edited by*

Robert C. Solomon

Thinking about Feeling

SERIES IN AFFECTIVE SCIENCE

*Series Editors*

Richard J. Davidson

Paul Ekman

Klaus Scherer

*The Nature of Emotion: Fundamental Questions*

edited by Paul Ekman and Richard J. Davidson

*Boo!: Culture, Experience, and the Startle Reflex*

by Ronald Simons

*Emotions in Psychopathology: Theory and Research*

edited by William F. Flack Jr. and James D. Laird

*What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS)*

edited by Paul Ekman and Erika Rosenberg

*Shame: Interpersonal Behavior, Psychopathology, and Culture*

edited by Paul Gilbert and Bernice Andrews

*Affective Neuroscience: The Foundations of Human and Animal Emotions*

by Jaak Panksepp

*Extreme Fear, Shyness, and Social Phobia: Origins, Biological Mechanisms, and Clinical Outcomes*

edited by Louis A. Schmidt and Jay Schulkin

*Cognitive Neuroscience of Emotion*

edited by Richard D. Lane and Lynn Nadel

*The Neuropsychology of Emotion*

edited by Joan C. Borod

*Anxiety, Depression, and Emotion*

edited by Richard J. Davidson

*Persons, Situations, and Emotions: An Ecological Approach*

edited by Hermann Brandstätter and Andrzej Elias

*Emotion, Social Relationships, and Health*

edited by Carol D. Ryff and Burton Singer

*Appraisal Processes in Emotion: Theory, Methods, Research*

edited by Klaus R. Scherer, Angela Schorr, and Tom Johnstone

*Music and Emotion: Theory and Research*

edited by Patrik N. Juslin and John A. Sloboda

*Handbook of Affective Sciences*

edited by Richard J. Davidson, Klaus Scherer, and H. Hill Goldsmith

*Nonverbal Behavior in Clinical Settings*

edited by Pierre Philippot, Erik J. Coats, and Robert S. Feldman

*Thinking about Feeling: Contemporary Philosophers on Emotions*

edited by Robert C. Solomon

# THINKING ABOUT FEELING

Contemporary Philosophers on Emotions

EDITED BY

Robert C. Solomon

**OXFORD**  
UNIVERSITY PRESS

2004

**OXFORD**  
UNIVERSITY PRESS

Oxford New York  
Auckland Bangkok Buenos Aires Cape Town Chennai  
Dar es Salaam Delhi Hong Kong Istanbul Karachi Kolkata  
Kuala Lumpur Madrid Melbourne Mexico City Mumbai Nairobi  
São Paulo Shanghai Taipei Tokyo Toronto

Copyright © 2004 by Oxford University Press, Inc.

Published by Oxford University Press, Inc.  
198 Madison Avenue, New York, New York 10016  
www.oup.com

Oxford is a registered trademark of Oxford University Press

All rights reserved. No part of this publication may be reproduced,  
stored in a retrieval system, or transmitted, in any form or by any means,  
electronic, mechanical, photocopying, recording, or otherwise,  
without the prior permission of Oxford University Press.

C. Calhoun, "Subjectivity and Emotion," is reprinted from the *Philosophical Forum*, 1989, with  
the permission of Blackwell Publishing Ltd. M. Nussbaum, "Emotions as Judgments of Value and  
Importance," is reprinted from P. Bilimoria and J. N. Mohanty, eds., *Relativism, Suffering and Beyond*  
(Delhi: Oxford University Press, 1997), with the permission of Oxford University Press India.

Library of Congress Cataloging-in-Publication Data  
Thinking about feeling : contemporary philosophers on emotions / edited  
by Robert C. Solomon.

p. cm.

Includes bibliographical references and index.

ISBN 0-19-515317-0

1. Emotions (Philosophy) I. Solomon, Robert C.

B815 .T48 2003

128'.37—dc21 2003001427

9 8 7 6 5 4 3 2 1

Printed in the United States of America  
on acid-free paper

*for Kathryn Ann Higgins (1925–2003)*

*This page intentionally left blank*

# CONTENTS

Contributors ix

Introduction 3

*Robert C. Solomon*

## Part I: Emotions, Physiology, and Intentionality

1. Primitive Emotions 9  
*John Deigh*
2. Emotion: Biological Fact or Social Construction? 28  
*Jenefer Robinson*
3. Embodied Emotions 44  
*Jesse Prinz*

## Part II: Emotion, Appraisal, and Cognition

4. Emotions: What I Know, What I'd Like to Think I Know,  
and What I'd Like to Think 61  
*Ronald de Sousa*
5. Emotions, Thoughts, and Feelings: Emotions as Engagements  
with the World 76  
*Robert C. Solomon*

## Part III: Emotions and Feelings

6. Emotion, Feeling, and Knowledge of the World 91  
*Peter Goldie*

7. Subjectivity and Emotion 107  
*Cheshire Calhoun*

Part IV: Emotions and Rationality

8. Emotions, Rationality, and Mind/Body 125  
*Patricia Greenspan*
9. Some Considerations about Intellectual Desire and Emotions 135  
*Michael Stocker*

Part V: Emotions, Action, and Freedom

10. Emotion and Action 151  
*Jon Elster*
11. Emotions and Freedom 163  
*Jerome Neu*

Part VI: Emotion and Value

12. Emotions as Judgments of Value and Importance 183  
*Martha Nussbaum*
13. Feelings That Matter 200  
*Annette Baier*
14. Perturbations of Desire: Emotions Disarming Morality in the  
“Great Song” of *The Mahābhārata* 214  
*Purushottama Bilimoria*

Part VII: On Theories of Emotion

15. Is Emotion a Natural Kind? 233  
*Paul E. Griffiths*
16. Emotion as a Subtle Mental Mode 250  
*Aaron Ben-Ze'ev*
17. Enough Already with “Theories of the Emotions” 269  
*Amélie Oksenberg Rorty*

*Bibliography* 279

*Index* 293

## CONTRIBUTORS

ANNETTE BAIER, Department of Philosophy, University of Pittsburgh,  
now a resident of New Zealand

AARON BEN-ZE'EV, Department of Philosophy, University of Haifa, Israel

PURUSHOTTAMA BILIMORIA, Department of Philosophy, Melbourne  
University

CHESHIRE CALHOUN, Department of Philosophy, Colby College

JOHN DEIGH, Department of Philosophy and the Law School, University  
of Texas, Austin

RONALD DE SOUSA, Department of Philosophy, University of Toronto

JON ELSTER, Department of Philosophy/Social Studies, Columbia  
University

PETER GOLDIE, Department of Philosophy, King's College, London

PATRICIA GREENSPAN, Department of Philosophy, University of  
Maryland, College Park

PAUL E. GRIFFITHS, Department of History and Philosophy of Science,  
University of Pittsburgh

JEROME NEU, Division of Humanities, University of California, Santa  
Cruz

MARTHA NUSSBAUM, Department of Philosophy and the Law School,  
University of Chicago

JESSE PRINZ, Department of Philosophy, University of North Carolina,  
Chapel Hill

JENEFER ROBINSON, Department of Philosophy, University of Cincinnati

AMÉLIE OKSENBERG RORTY, Department of Philosophy, Brandeis  
University

ROBERT C. SOLOMON, Department of Philosophy, University of Texas,  
Austin

MICHAEL STOCKER, Department of Philosophy, University of Syracuse

Thinking about Feeling

*This page intentionally left blank*

# Introduction

ROBERT C. SOLOMON

In this volume, I have tried to bring together some of the best Anglo-American philosophers now writing on the philosophy of emotion.

That field, the philosophy of emotion, is by one measure quite recent. In the Anglo-American tradition, the subject of emotion was for a considerable period disreputable, typically dismissed as “mere subjectivity” or, worse, as nothing but physiology plus dumb sensation. An ethical theory known as emotivism took center stage during and just after the Second World War, in which all of ethics was dismissed as nothing but expressions of emotion with no more cognitive content than “Boo!” or “Hooray” (Ayer 1952). It was only with occasional pieces by Princeton philosopher George Pitcher and Edinburgh philosopher Errol Bedford and then a book by Anthony Kenny that the subject started to become noticed at all, although it was several years more before it began to attract an audience and deserve recognition as a “field” (Pitcher 1965, Bedford 1953, Kenny 1963). Today, by contrast, it is evident to most philosophers that emotions are ripe for philosophical analysis, a view supported by a considerable number of excellent publications. Emotions have now become mainstream.

This is not to say, of course, that the philosophy of emotion is something new. Philosophers since Aristotle have explored it with considerable interest, usually motivated by an interest in ethics. The Stoics and Epicureans carried on a lively debate over several centuries on the nature of emotion and the passions’ place in ethics and the quest for the good life. Medieval philosophy is filled with concern about the emotions, both as “higher passions” (e.g., love and faith) and as “lower” passions, a.k.a. “sins.” And in this century, “Continental” European philosophy remains keenly aware of the importance of the emotions in human life, thanks in

part to the two giants Martin Heidegger and Jean-Paul Sartre. It was no surprise that Anthony Kenny in his *Action, Emotion, and Will* (1963) drew heavily from Aristotle and the medieval Scholastics, and it is no surprise now that many contemporary Anglo-American philosophers feel compelled to take on Sartre (less often Heidegger) in their efforts to understand the nature of emotion. I see the new emphasis on emotion in Anglo-American philosophy not so much as something new as the rediscovery of a discipline that is very old and has always been essential to the “love of wisdom.”

The essays in this collection represent a variety of positions on a number of topics: the nature of emotion, the category of emotion, the rationality of emotions, the relationship between an emotion and its expression, the relationship between emotion, motivation, and action, the biological nature versus social construction of emotion, the role of the body in emotion, the extent of freedom and our “control” of emotions, the relationship between emotion and value, and the very nature and warrant of “theories” of emotion. Since I have my own say in the book (and elsewhere), I will not preempt my contributors by commenting on these various topics here. I will simply say that these are all lively and very current issues of considerable interest and importance to a wide variety of theorists in the various theoretical, experimental, and clinical branches of psychology, philosophy, philosophical psychology, and moral psychology, as well as cognitive science, the social sciences, and literary theory, and I hope that the controversies that become evident in this volume will contribute to the cross-fertilization of these disciplines.

As in any “collection,” there will be questions about choices—who was invited to contribute, who was not. To answer these potentially embarrassing questions as quickly and straightforwardly as possible, I have tried to solicit chapters from those theorists who have already established solid reputations in the field of emotion research and are presently (still) working on the emotions. There are a substantial number of younger scholars and researchers who are not represented here, I am sorry to say, and there are several well-known figures in the field who are no longer working on emotions but have moved on to other interests. I have also restricted my attention (for the purposes of this volume) to philosophers, although I have profound difficulties with the often artificial distinctions between that “queen of the sciences” and its kin: psychology and the other social sciences, history, and literature. There are several psychologists, to be sure, who have philosophically rich and suggestive things to say about emotion, and there are historians who have written and are writing valuable histories of the emotions and the histories of ideas about emotions. And, of course, it is often and rightly said that the most lucid insights about the nature and “logic” of emotions are to be found not in the terse

prose of the philosophers but in the pathos-ridden and often more lyrical writing of poets and novelists. But that all leads, I am afraid, to a collection that would constitute a library, not a single volume. A modest selection of excellent pieces from some of the most prominent current philosophical researchers on the scene is all that I can promise here.

Special thanks to Catherine Carlin and John Rauschenberg, and to Farrah Ghazi Zughni for her help with the index.

*This page intentionally left blank*

I

Emotions, Physiology, and Intentionality

*This page intentionally left blank*

# I

## Primitive Emotions

JOHN DEIGH

Two facts about emotions stand out among the many that a theory of the subject ought to cover. The first is that emotions are intentional states in the sense that they are directed at something. Hope, for example, is a state of mind directed at a future condition or event. One hopes for sunny weather on the day of the picnic or calm seas on the day of the regatta. In this respect, hope is unlike giddiness or drowsiness, states of mind that can occur undirected at anything. The difference is nicely illustrated in the opening scene of Shakespeare's *Romeo and Juliet*, when two servants of the house of Capulet and two servants of the house of Montague cross paths. "Do you bite your thumb at us, sir?" Abram, one of the Capulet servants, asks Sampson, the offending Montague retainer. "No sir," replies Sampson. "I do not bite my thumb at you, sir. But I bite my thumb, sir."<sup>1</sup> Sampson's reply, as we know from an aside to his confederate, is insincere. His thumb biting is aimed at the Capulet servants, and we understand its being aimed at them by recognizing the state of mind it expresses. It expresses contempt, an intentional state, and its target, so to speak, that at which it is directed, is the Capulet pair. If Sampson's reply had been sincere, then his thumb biting would not have expressed this intentional state, though it might still have signified a state of mind, like chronic and undirected jitters. But in this case, the state it would have signified would not have been an emotion.<sup>2</sup>

The second fact about emotions that a theory of the subject ought to cover is that emotions are common to both humans and beasts. This is not to say that humans and beasts are liable to the same set of emotions. On the contrary, the set to which humans are liable is much greater than the set to which beasts are liable. Shame over a moral failing, for instance, is

an emotion to which humans are liable and beasts are not. It is to say, though, that some emotions are common to both sets. These are, in many cases, what I will call primitive emotions. They are the emotions liability to which is instinctive. That is, a human's or beast's liability to them is an inherited trait whose development, to the extent that it depends on the existence of environmental conditions, depends only on those necessary for meeting basic biological needs. Fear, anger, and delight all have primitive forms. The terror of horses fleeing a burning stable, the rage of a bull after provocation by a tormentor, and the delight of a hound in finding and retrieving his quarry are all examples.

A successful theory of emotions must account for both of these facts. It cannot skirt them. Yet accounting for both has proven to be surprisingly difficult. Some theories, particularly the cognitivist theories that have been so influential in philosophy and psychology over the last thirty years, use the first fact as their point of departure and leading idea, but they then have trouble accommodating the second.<sup>3</sup> Other theories, particularly those that have developed under the influence of Darwin's seminal work *The Expression of the Emotions in Man and Animals* (1998), take the second fact as their springboard, but they then have trouble accommodating the first. The reason, in either case, is the gap between the way intentional states of mind are typically understood and the way primitive emotions are typically understood. The problem of closing this gap seems to outstrip the resources of these theories. The point is not generally recognized, however. It tends to lie beyond the theories' horizons. The object of this essay is to bring it forward and to vindicate it.

The gap appears most clearly when one considers the trouble that cognitivist theories have in accounting for primitive emotions. On standard cognitivist theories, an intentional state of mind is either a thought or a compound state that includes a thought as a component, and in either case the content of the thought is represented as a proposition. Consider again the hope of a picnic planner for sunny weather. The emotion contains a thought about the advantages of sunny weather for picnicking, and the content of that thought is naturally represented by a proposition in which being sunny is predicated of the day of the picnic. Indeed, sometimes we make the propositional character of such thoughts explicit, as when we describe a person who is planning a picnic as hoping that the day of the picnic will be sunny. But propositional thought presupposes linguistic capacities, which are unique to human beings and, in fact, human beings who have grown past infancy. Consequently, if one represents the thought content of every intentional state as a proposition, one cannot account for primitive emotions. One's theory of emotions in that case will be like the theory of the ancient Greek and Roman Stoics. They held that emotions were judgments, which is to say, affirmations and denials of propositions, and therefore that beasts and babies were incapable of emotions. Such a

view is no longer tenable, however. Like Descartes's cognate view that human beings alone among the animals of the world have minds, it has passed into history. So cognitivist theories of emotion must give up taking the thoughts emotions contain as in every case a proposition. They must find a way to explain some of those thoughts as nonpropositional so as to avoid making the possession of linguistic capacities a condition of being liable to emotions.<sup>4</sup>

This demand may not seem all that difficult to meet. After all, you might think that a defender of a cognitivist theory could just assume that the thoughts primitive emotions contained were like the thoughts contained in the distinctively human emotions that her theory takes as the paradigms of its subject, except that they lacked propositional form. Indeed, you might think that the thoughts contained in the former were just unencoded versions of the thoughts contained in the latter. To think this, however, would be a mistake. The concept of an encoded thought is that of a thought expressed in the words of some language or its equivalent.<sup>5</sup> When the thought is a complete one, then it is expressed by a complete, declarative sentence of that language. Consequently, if there is a version of this thought that is unencoded, it must be a complete thought in abstraction from every complete, declarative sentence that expresses it, and this is just what logicians mean by a proposition. A proposition, on their understanding of it, just is the meaning of a complete, declarative sentence of some language. It is what one grasps when one understands the sentence and what one preserves when one accurately translates it into a sentence of another language. If the translation is accurate, then the two sentences have the same meaning. They express the same proposition. Hence, defenders of cognitivist theories cannot use the idea of an unencoded version of an encoded thought to explain the thoughts that primitive emotions contain, for this idea just is the logicians' idea of a proposition.

Nonetheless, you might still think that the difficulty is not that great. For you might think that, even if the thoughts primitive emotions contain are not unencoded versions of the thoughts contained in the distinctively human emotions that cognitivist theorists take as the paradigm of their subject, we can still understand them as like the thoughts contained in the latter except that they lack propositional form. But to think that we could so understand them is to suppose that there is some way in which they and the thoughts that these distinctively human emotions contain are alike, and it is unclear what the form of this likeness could be. Of course, both are alike in being identical with or a component of an intentional state of mind, but to say that they are alike in this way is merely to reaffirm what is true of both types of emotion in virtue of their being intentional states. It is merely to reaffirm that intentional states are or include thoughts. A more specific account of what makes them alike is necessary if their being alike is to explain the character of the thought that primitive emotions

must contain in virtue of their being intentional states, and no such account seems available.

The difficulty is an old one. It goes back to the problems on which Locke's doctrine of abstract ideas came to grief. Locke, you may recall, advanced this doctrine in opposition to the Cartesian belief that some thoughts, the intellectual ideas, did not originate in sensory experience while others, in particular, those that are the product of our imagination, did. On Descartes's theory of human cognition, the intellectual ideas were radically unlike the ideas of imagination. They were clear and distinct. The latter were confused and obscure. And what explained the difference was the dependence of the latter on the operations of the body's sensory apparatus. These operations produced sensory images and internal feelings, and the ideas of imagination were composed of memories and replicas of these images and feelings. The intellectual ideas, by contrast, did not depend on the operations of the body's sensory apparatus and were comprehensible apart from the sensory images and internal feelings they produced. This distinction corresponds to a distinction between thoughts common to both humans and beasts and thoughts that are distinctively human, though of course no Cartesian would have embraced this latter distinction, since it presupposes what they denied, namely that beasts had minds. Locke, however, did not deny that beasts had minds. On the contrary, he took sensory experience to be common to both humans and beasts and, in consequence, held that both were capable of retaining the resultant ideas in memory and of discriminating among them. Distinctively human thought, Locke maintained, consisted in applying the power of abstraction to these ideas, for humans alone possessed this power.<sup>6</sup> Humans alone, that is, had the power of attending exclusively to some feature of an idea while neglecting all the others. An abstract idea, then, was an idea that one held in memory or formed in imagination and that one understood to represent generally a property of things that corresponded to the feature of the idea one had abstracted. Accordingly, Locke identified the ideas that came immediately from sensory experience as thoughts common to both humans and beasts and abstract ideas as thoughts that were distinctively human. At the same time, abstract ideas, on Locke's view, were not radically unlike other ideas. To the contrary, they were ideas of memory and imagination to which the power of abstraction was applied. Thus Locke defined a kind of idea that was distinctively human in ways analogous to Descartes's intellectual ideas and that was nonetheless like the ideas that were common to humans and beasts.

Yet for this definition to cover the same cognitions that Descartes explained as intellectual ideas, it had to capture the thoughts that words and sentences express when they are used with their customary meaning. This requirement is evident from Descartes's point, at the start of Meditation VI, that our knowledge of the difference between a chiliagon and a myriagon

cannot come from comparing the ideas we form of these figures in our imagination, since any idea we form in imagination of either figure will be indistinguishable from the idea we form of the other.<sup>7</sup> Our knowledge must come instead, Descartes observed, from our comparing the intellectual ideas we have of these figures or, as we would now say, our concepts of them. The knowledge, then, to which Descartes appealed in this passage consists in our conceptual understanding of these figures, and this is the same as our understanding of what it means to say that a figure is a myriagon and not a chiliagon. It is the same, that is, as our understanding of what thought the sentence “A myriagon is not a chiliagon” expresses when the sentence is used with its customary meaning. Nor was Descartes’s point peculiar to mathematical objects. One could make the same point about our knowledge that a coyote is not a dog. So the success of Locke’s opposition to Descartes’s theory depended on his capturing with his definition of abstract ideas the thoughts that words and sentences express when they are used with their customary meaning.

Locke, of course, though he may not have recognized the force of Descartes’s point, meant his definition of abstract ideas to capture such thoughts.<sup>8</sup> In this regard, he initiated a long tradition in modern empiricist philosophy of programs for reducing what words and sentences mean to a set of sensory images and internal feelings common to all speakers. He supposed that the thoughts we express in language precede and are independent of our knowledge of language, and he further supposed that we came to have such thoughts by first making comparisons among the great many sensory images and internal feelings that fill our minds and then exercising the power of abstraction to isolate in thought those features and facts that interest us and that we use words and sentences to denote. In short, he conceived of the thoughts we express in language as wholly independent of our linguistic capacities, for he conceived of the powers of comparison and abstraction as operating independently of such capacities. Hence, Locke’s doctrine of abstract ideas, if it were sound, would close the gap between the way we typically understand emotions as intentional states and the way we typically understand primitive emotions. On his doctrine, the thoughts we attribute to emotions in virtue of their being intentional states do not presuppose linguistic capacities and are therefore attributable to primitive emotions as well as to the distinctively human ones that cognitivist theories take as the paradigms of their subject.

The difficulty with the doctrine, however, is that it fails to account for the thoughts we express in language. Specifically, the power of abstraction, when understood as a power that operates independently of linguistic capacities, cannot yield such thoughts. It cannot, for instance, yield the thought we express when we say that a coyote is not a dog. For no amount of abstraction from the sensory images of dogs will isolate in one’s thought features that show, in view of one’s abstract idea of a coyote, that a coyote

is not a dog. The reason, moreover, is not or not merely that whatever features one abstracts will be at too great a level of phenotypic generality to be features that coyotes lack, though this is no doubt true. The reason, rather, is that we distinguish coyotes from dogs because of their genotype and regardless of any phenotypic difference between them. A coyote is not a dog because none of the ancestors it has in common with dogs was a dog, and no dog has a coyote as an ancestor. Descartes, then, was right to treat intellectual ideas as radically unlike the ideas that come from sensory experience and are held in memory or produced in imagination. What we now call concepts and the propositions they help to constitute are not explicable on Locke's doctrine of abstract ideas.

The failure of Locke's program and of programs like it to reduce the thoughts we express in language to sensory images and internal feelings means that Descartes's theory of human cognition survives the attack that its traditional empiricist opponents made on it. What is more, the failure of their attack leaves unopposed the view that the thoughts we express in language are radically unlike the thoughts common to both humans and beasts. And in the absence of a viable alternative to this view, cognitivist theories of emotions must therefore abandon giving a uniform account of the thoughts in virtue of which emotions are intentional states. They must, in other words, take the thoughts in virtue of which primitive emotions are intentional states to be radically unlike the thoughts in virtue of which the emotions they take as paradigms of their subject are intentional states. Yet how they can do this consistently with their signature thesis that the thoughts in virtue of which emotions are intentional states are the principal determinants of the nature of emotions is a significant challenge. It is hard, after all, to maintain that one has satisfactorily explained the nature of something if one also allows that its nature could be determined by either of two radically dissimilar things.

Let us leave it to the defenders of these theories to stew over this problem and turn next to theories of emotions that take the second fact, that emotions are common to humans and beasts, as their guide. As I said at the outset, these theories draw their inspiration from Darwin's work on the expression of emotions in humans and other animals, and accordingly I will refer to their defenders as Darwinians.<sup>9</sup> Darwin himself was unconcerned with the question of the nature of emotions. He reflexively accepted the empiricist conception of them that was the orthodoxy of his time.<sup>10</sup> This conception identifies emotions with feelings as distinct from thoughts. They are, in Locke's words, "internal sensations," a phrase that nicely reveals the assimilation of emotions to sensations characteristic of traditional British empiricist psychology.<sup>11</sup> Consequently, on this traditional psychology, emotions, being pure feelings, are mental states that are not essentially directed at anything. Hence, the standard British empiricist conception immediately runs into trouble when applied to the first fact, that emotions

are intentional states. For this reason, among others, it now has few defenders. In particular, the Darwinians do not defend it. Though they are inspired by Darwin's work, they assume a different conception from his. Nevertheless, it too has trouble accommodating the first fact. To understand why will require some explication of their program, and to do this it is best first to explain how it emerges from Darwin's.

Darwin was chiefly concerned with involuntary expressions of emotion. He was particularly interested in the involuntary facial expressions common among human beings. The study of these has a long history, going back at least to Descartes's explanations of how emotions are manifested in laughter, tears, blushing, paling, the wrinkling of the brow, the quivering of the lips, and so forth.<sup>12</sup> Darwin did not follow Descartes's lead, however. Indeed, as far as I can tell, he was unaware of Descartes's work, though he does mention the work of the seventeenth century painter Charles LeBrun, who had based his teachings of how to paint the face on Descartes's theory.<sup>13</sup> In any case, Darwin's interest in his predecessors in this field was more local. The main writer whose views interested him was Charles Bell, a prominent physiologist whose book *Anatomy and Philosophy of Expression* Darwin praised for having "laid the foundation of the subject as a branch of science."<sup>14</sup> What especially interested Darwin in Bell's work was Bell's view that the musculature of the human face was unique to human beings. Nothing like it, Bell maintained, occurred in other species. Moreover, Bell regarded this fact, or rather what he mistakenly thought was a fact, as evidence of God's design in creating human beings. God, Bell held, gave human beings these special facial muscles for the purpose of expressing the emotions distinctive of humankind. Needless to say, Bell's thesis offered Darwin a ripe opportunity for showing, in a new area of natural history, the superiority of evolutionary theory to explanations that appealed to God's design.<sup>15</sup> *The Expression of the Emotions in Man and Animals* was the result of Darwin's having seen this opportunity and seized it.

Darwin focused his study on facial expressions that seemed purposeless. Of course, if Bell were right, their purpose would be to express emotions just as nodding and shaking one's head express affirmation and negation. The point, in either case, is to communicate an attitude or thought. But on this view the connection between a movement of the face and the emotion it expresses would be entirely arbitrary: as long as God's design of our facial muscles had as its sole purpose to facilitate our expressing emotions, raised eyebrows might just as well have expressed dejection as surprise, a curled lip might just as well as have expressed admiration as scorn. Darwin thought, to the contrary, that one could see a purpose in many of these movements that would be well suited to the prehistoric environments in which the distant ancestors of modern human beings lived, even though it was ill suited to the environments of modern human life. Accordingly, he proposed that human beings inherited the disposition to make these

movements from their ancestors, who had themselves acquired it because of the usefulness of the movements in their prehistoric environments. On this explanation, then, the connection between surprise and raised eyebrows, for example, was not arbitrary because the movement was advantageous in the circumstances that typically provoked surprise in prehistoric environments. Thus raising one's eyebrows, Darwin observed, is necessary to opening one's eyes widely, and wide open eyes enable one to scan one's surroundings quickly, something it would be very useful to do in circumstances in which unexpected sights and sounds were often omens of danger.<sup>16</sup> And though these circumstances may not be so common in modern life as to explain why raised eyebrows accompany surprise, they were common enough in the life of our prehistoric ancestors to explain why they acquired the disposition to raise their eyebrows when surprised.

Darwin, in offering these explanations, usually characterized the ancestors who first acquired the disposition as themselves hominids. Sometimes, though, to reinforce his explanations he cited similar facial and bodily movements of other primates, thus implying that the ancestors who first acquired the disposition to make these movements were not hominids. And sometimes he appealed to the bodily movements of animals even lower down on the phylum that expressed the same emotion. Thus, to explain why our hair stands on end when we are frightened, Darwin noted that the same thing occurs in many other animals and that, though it serves no purpose in our life, it does serve the purpose in theirs of discouraging predators and cowering rivals. This is because their hair typically constitutes a coat and consequently, when erect, makes them appear larger and fiercer.<sup>17</sup> In these cases, then, the ancestors from whom humans inherited the disposition are the progenitors of many other animal species as well.

Clearly, Darwin could not argue for the superiority of these explanations to theological ones like Bell's if the dispositions to make the movements that were their explananda were not inherited. He could not argue for their superiority, for instance, if human beings acquired these dispositions by mimicking their parents' behavior or following their parents' instructions. Evolutionary theory, in that event, would be inapplicable. Hence, Darwin had to establish that the movements on which he focused were the products of inherited dispositions. Obviously, in a case like bristling hair that, as an expression of emotion, is common to many species, the evidence that the underlying physiological mechanism is inherited is irrefutable.<sup>18</sup> But in many other cases, such as raising one's eyebrows as an expression of surprise or curling one's lip as an expression of scorn, the burden of establishing that the disposition to make these movements is inherited is more difficult to meet. Darwin was aware of this burden and devoted a good deal of his research to meeting it. Specifically, he recognized that he needed to isolate the involuntary expressions that were the proper object of his study from expressions that were merely conventional. The latter, he

observed, were most likely to be learned in childhood and to vary across cultures. For example, as we saw in *Romeo and Juliet*, biting one's thumb was an expression of contempt in Shakespeare's Verona. It does not, by contrast, express contempt, at least not conventionally, in contemporary America. Rather, the conventional, vulgar expression of contempt in America is displaying a fist with a raised middle finger. True expressions of emotion, therefore, Darwin declared, differ from conventional ones in being instinctive or innate. As such, they are likely to be invariant across human cultures. Or, at any rate, showing that the same movements were recognized in many different and disconnected cultures as expressing the same emotion would be powerful evidence that the disposition to make those movements, when experiencing that emotion, was inherited.<sup>19</sup> A substantial part of Darwin's work, then, entailed gathering and presenting such evidence for a broad range of movements, particularly facial movements, that express emotions.

It is this part of Darwin's work that inspires the Darwinians. Their program for studying the emotions follows his account of the true expressions of human emotion and uses the same principle he used to organize the chapters that presented his account. Thus, the Darwinians focus on human emotions and divide them into a small number of basic, general categories, each of which is defined by facial movements that, according to their research, qualify as the true facial expressions of all emotions that belong exclusively to that category. There are some disagreements among the Darwinians about how many basic categories there are and what emotions belong to them, but these disagreements are minor. A typical Darwinian division includes, as its basic categories, joy, anger, disgust, surprise, fear, distress, and sadness.<sup>20</sup> Each of these is then understood to cover a range of cognate emotions. The category of joy covers happiness, delight, gladness, satisfaction, and so forth; that of anger covers annoyance, indignation, rage, resentment, and so forth. And what makes the emotions in a given category cognate, that is, what explains why they belong to the same basic category, is that their true facial expressions consist of the same facial movements. To be sure, some emotions also have conventional facial expressions, but these are irrelevant to the determination of the basic category to which they belong. Holding one's nose, for instance, is a conventional expression of distaste, which is a mild form of disgust. But it is not by virtue of this expression that distaste, on the Darwinians' program, belongs in the category of disgust. Rather its membership in this category is due to its being expressed by wrinkling one's nose, raising one's nostrils, and lowering the inner corners of one's eyebrows.<sup>21</sup>

What separates the Darwinians from Darwin is their belief that the study of an emotion's true expressions illuminates the very nature of the emotion. Darwin, as I noted, conceived of emotions as analogous to sensations in accordance with traditional British empiricism, and on that con-

ception an emotion's true expressions no more illuminate its nature than swollen gums illuminate the nature of a toothache. The Darwinians, however, conceive of emotions differently. An emotion, on their conception of it, is a neurophysiological event whose manifestations typically include the facial and overt bodily movements that are the emotion's true expressions.<sup>22</sup> The event that is the emotion occurs when certain neurophysiological mechanisms are activated, and activating the mechanisms produces these movements along with covert physiological changes such as changes in heartbeat and electrogenic activity.<sup>23</sup> In human beings, the most perspicuous true expressions of emotion are facial movements, and the Darwinians take these to be the determinants of the neurophysiological mechanisms whose activation, on their conception, produces the emotion. These movements, that is, in virtue of being the emotion's true facial expressions, fix as the referent of the basic category to which the emotion belongs those neurophysiological mechanisms whose activation produces the emotion. In this way the true facial expressions of an emotion are evidence of its very nature. Indeed, one philosophical enthusiast for the Darwinians' program has declared that the success of their program establishes the basic categories into which they divide the emotions as natural kinds by virtue of this reference-fixing character of the facial movements that define the basic categories and the applicability of evolutionary theory to the neurophysiological mechanisms reference to which they fix.<sup>24</sup>

It should be clear, then, that the Darwinians view their division of the emotions into these basic categories as corresponding to major real differences among the emotions. Because they hold that the different facial movements that define the categories determine, for each category, the neurophysiological mechanisms whose operations the emotions in that category consist of, they understand the real distinction between joy, say, and all of its cognate emotions, on the one hand, and sadness and all of its cognate emotions, on the other, to be a distinction between the neurophysiological mechanisms that produce the facial movements defining those categories. And the same is true of other major differences they find among the emotions. The Darwinians recognize minor differences among the emotions too. Thus they explain differences among the emotions that belong to the same category—distaste and revulsion, for instance, which are forms of disgust—as reflecting differences in the intensity, duration, and course of the neurophysiological events that constitute these emotions. And they explain differences among emotions that belong to two or more categories—horror, for instance, which is a mixture of fear and disgust—as reflecting differences in the combination of neurophysiological mechanisms whose operations those emotions consist of. These explanations follow more or less directly from how the Darwinians explain the major differences among emotions and, consequently, do not introduce any substantially new premises into their program. Thus, in considering how well their program ac-

counts for the two facts I have highlighted, we can concentrate exclusively on their explanations of these major differences.

That the Darwinians can readily account for the second fact, that emotions are common to humans and beasts, is evident. The mechanisms to which they attribute the major differences among the emotions have the same place in their theory that the inherited dispositions to make facial movements have in Darwin's explanations of the true facial expressions of emotions. Indeed, on the Darwinians' most ambitious hypothesis, they suppose that each of the basic categories into which they divide the emotions corresponds to a complex program that is genetically hard-wired in the brain, as it were, and that coordinates activation of the different neurophysiological mechanisms whose operations an emotion consists of. These affect programs, to use the term those who advanced the hypothesis favor, are inherited dispositions, and the ancestors from whom human beings inherited them are the progenitors of many different species of animal.<sup>25</sup> Hence, the Darwinians, given their conception of an emotion as a neurophysiological event that occurs when an affect program is activated, can account for there being emotions common to humans and beasts by identifying many of the emotions that belong in one or another of their basic categories as emotions to which some beasts are liable in virtue of their having affect programs that are homologous to the human affect programs to which those categories correspond. The question, then, is how, given this conception of emotion the Darwinians can account for the first fact, that emotions are intentional states.

To do so, they must explain intentional states differently from the way standard cognitivist theories of emotion explain them. That is, they must explain differently how emotions are directed at objects. Standard cognitivist theories, as I observed earlier, explain this feature by attributing propositional thought to emotions, for they take the emotions distinctive of human beings as the paradigm of their subject and the thought content of these emotions is propositional. The Darwinians, by contrast, start with a conception of emotions as common to humans and beasts and indeed as having first occurred in beasts millions of years before the first language using animals appeared on the earth. They will therefore have no interest in this or any explanation that is based on the emotions distinctive of human beings. Instead, they must give an explanation of how emotions can be directed at things that is immediately consistent with the emotions of beasts. The explanation they give, that is, must immediately fit what goes on, say, when a dog, angered by a stranger's invasion of his territory, growls or barks at the stranger. The dog's anger, in this case, is directed at the stranger, and what the Darwinians must explain is how to understand this feature of the dog's emotion, given that it cannot be explained by a belief the dog has or a judgment he makes that the stranger has invaded his territory.

What is going on in a dog's mind when he growls at someone? Suppose, for example, you need to enter your neighbors' 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 sensitivity, 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 this 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. The question, then, is whether these two features of the perceptions that excite and sustain the dog's emotion, your being the object of his attention and his tracking you by virtue of his sensitivity to some property you have, are sufficient to constitute you as the object of that emotion. If they are, then the explanation is one the Darwinians should find congenial. For they could still conceive of the emotion, on this explanation, as a neurophysiological event. They would locate its object in the perceptions that excited and sustained that event, which is to say, in the sensory images that activated the neurophysiological mechanisms whose operations, on the Darwinians' view, it consisted of. Since the Darwinians do not expressly offer an explanation of the fact that emotions are intentional states, let us assume that they would endorse this one.

The commonest objection to explanations like this one is that they confuse the cause of an emotion with its object. In the above example, you are both the cause and the object of the dog's emotion.<sup>26</sup> One can give other examples of emotions, however, in which the two are different, and the objection is that explanations like this one fail to capture that difference. I believe, though, that a little tinkering with the explanation can save it from this objection. Consider an example of Norman Malcolm's. Imagining his dog chasing a neighbor's cat, Malcolm writes, "[The cat] runs full tilt toward an oak tree, but suddenly swerves at the last moment and disappears up a nearby maple. The dog doesn't see this maneuver, and on arriving at the oak tree, he rears up on his hind legs, paws the trunk, and barks excitedly into the branches above."<sup>27</sup> The dog, we might say, is barking at a cat he thinks is up the tree.<sup>28</sup> And if we further suppose that rustling leaves due to movements of a small bird that the dog doesn't see, or perhaps just the wind, cause the dog to continue to bark, then the object

of the dog's excitement, the cat, is no longer its cause. In this case, although the dog can no longer be said to be tracking the cat, he can still be said to be responding to a sensory image of the cat that is sustained in his mind by the sound of rustling leaves. Accordingly, the cat is still the object of the dog's thought, and the dog is still responsive to the cat's image, and to changes therein, by virtue of his sensitivity to some apparent property of the cat presented in the image. Hence, by replacing the notion of tracking with that of responsiveness to an image, in the explanation of how certain features of the perceptions that excite and sustain an emotion constitute the emotion's object, we can preserve in this explanation the distinction between an emotion's cause and its object.

The real problem with the explanation lies elsewhere. When the neighbor's dog growls at you as you approach the gate, he senses something invasive about your behavior. This is something about you he doesn't like. If there were nothing about you he didn't like, then we could make no sense of his growling *at you* or the anger it expresses. In other words, the dog could not be angry *at you* unless there were something about you or, more exactly, unless there were something about the way you appeared to him that made him angry, and it must be something, like invasiveness, to which anger is an intelligible response. Indeed, by identifying the emotion as anger and the dog's growling as its expression we make intelligible behavior that would otherwise be no more intelligible than a fit of hiccups. And we do so because, in identifying the emotion as anger, we identify how its object appears to the dog and not simply that its object is the object of the dog's attention and is being tracked by him. Hence, for something to be the object of an emotion, whether the emotion is anger, disgust, pity, embarrassment, shame, or what have you, it must appear to the subject in a way that makes his feeling an emotion of that type intelligible or it must be thought by him to have a property whose possession by the object makes his feeling an emotion of that type intelligible. To be the object of anger, for instance, something must appear or be thought to be invasive, injurious, offensive, or the like. To be the object of disgust, something must appear or be thought to be foul or rotten or putrid. To be the object of pity, something must appear or be thought to be in some sorry or wretched condition. The real problem, then, with the explanation that we are assuming the Darwinians would endorse is that it misses this intelligibility condition on something's being the object of an emotion. As far as the explanation goes, you could be the object of someone's attention, that person could be tracking you, and these features could be features of perceptions that activated neurophysiological mechanisms whose operations produced the facial movements that were the true facial expressions of a certain type of emotion, and yet you might still not be the object of an emotion of that type. For it might still be the case that you do not appear to its subject in any way that makes his feeling an emotion of that type intelligible.

This problem, unlike the last, defeats the attempt to come up with an explanation of the intentionality of emotions that the Darwinians would find congenial. One could, to be sure, revise the explanation again to resolve the problem. But one could not do so and still maintain the Darwinians' conception of emotions. And this gets to the heart of the trouble the Darwinians have in trying to accommodate the fact that emotions are intentional states. Because they conceive of an emotion as a neurophysiological event whose type is determined by certain facial movements, namely those that are produced by the operations of the neurophysiological mechanisms that the emotion consists of, they have to allow, as a conceptual possibility, that those operations and the facial movements they produce can occur, and occur in response to perceptions of a particular object, even though the object does not appear to the subject in any way that makes his feeling the emotion intelligible. For they cannot deny this possibility without also denying the possibility that the facial expressions defining one of the basic categories into which they divide the emotions occur on some occasion as the true expressions of an emotion that belongs to some other basic category. And this surely can happen. That it can happen, moreover, confounds the Darwinians' theory. For when it happens, the object of the emotion appears to the subject in a way that makes his feeling that emotion intelligible but that does not make intelligible his feeling any of the emotions that, according to the Darwinians' scheme, belong in the category defined by his facial expressions. When it happens, in other words, the Darwinians must insist on his feeling some emotion that he is in fact not feeling.

Consider, as an example, the phenomenon known as Beatlemania. It has been wonderfully captured in a video of the Beatles 1965 concert at Shea Stadium.<sup>29</sup> Here is a brief description. The Beatles suddenly appear from a tunnel and run across the field directly to the stage on which they will play. At their appearance, thousands of teenage girls in the stands begin to scream and then to shriek. The noise is deafening. The girls continue to scream and shriek as the Beatles start to play, and their screaming never stops. Many of the girls, at some point, break down into tears. When they do, the tears flow freely. The girls weep. They sob. Their bodies slump. Their faces lose all composure and become blubbery and slack. If one were presented with pictures of these girls' faces and did not know the context in which the pictures were taken, one would say they were the faces of great sorrow, anguish, or grief. Yet the girls are experiencing none of these emotions. The object of their emotions is the Beatles, and nothing about the Beatles on this occasion would make sorrow, anguish, or grief an intelligible response. For any of these emotions to be an intelligible response, one or more of the Beatles would have to have suffered some grave misfortune or at least to have appeared to have suffered such misfortune, and none of them obviously has. To the contrary, all of them appear to be

having a great time, though they are rather perplexed by all the screaming. The girls' faces, then, are products of a different emotion. It is ecstasy or rapturous joy at seeing and being near the objects of their most ardent devotion. On the Darwinians' theory, however, there is no basis for attributing ecstasy or joy to these girls. On their theory, whatever emotion the girls are experiencing, it must belong either to the category of sadness or to the category of distress.<sup>30</sup> Their theory, because it excludes considerations of intelligibility from its definitions of the basic categories of emotions, cannot correctly identify the girls' emotions in this case.

Needless to say, it is crucial to this criticism that the expressions on the girls' faces are true expressions of their emotions in Darwin's sense. They must be true expressions and not conventional, for otherwise nothing in the Darwinians' theory would require identifying the girls' emotions as belonging to either the category of sadness or that of distress. If the girls' behavior were merely histrionic, for instance, then it would not confound the Darwinians' theory. But to contend that it was merely histrionic would be implausible. It is evident that there is no artifice in the girls' expressions of emotion. Indeed, it is hard to see how a defender of the Darwinians' theory could deny that the girls' expressions were true. A true expression of emotion, recall, is one that is instinctive in the sense that it manifests an inherited disposition to make the movements that the expression consists of, and the girls' weeping is as much a manifestation of such a disposition as weeping that expresses anguish or grief when it too is brought on by a screaming fit. In neither case of weeping that results from screaming is there any reason to think that the display of emotion is less true than in the other.

Darwin's account of such weeping makes this point clear.<sup>31</sup> According to Darwin, strong contraction of the muscles around the eyes produces tears as a result of its stimulating the lachrymal glands, and such contraction occurs when one screams. Indeed, as Darwin explained, any violent expiration of air, such as violent coughing, sneezing, or laughter, will cause these muscles to contract and bring tears as a result. Of course, people do not cough or sneeze or laugh when they suffer great pain or become aware of a grave personal loss. But they do scream. Screaming, after all, is the common and presumably universal response to pain among infants even before their lachrymal glands develop to the point where they can shed tears. And while learning to control the impulse to scream is part of learning how to deal with pain, the impulse remains even after one has acquired some control over it, and it still produces screaming when the pain or the loss is great enough. Thus weeping and sobbing, and the facial movements that occur when weeping results from prolonged and intense screaming, are widely recognized as expressions of anguish and grief. The phenomenon of Beatlemania, by contrast, and similar crowd phenomena in which prolonged and intense screaming occurs not as the result of pain or grave

personal loss, but as the result of the sudden appearance of the objects of ardent devotion, are too peculiar for the facial expressions of the emotions characteristic of them to be widely recognized outside of the contexts in which they occur as expressions of those emotions. Nevertheless, these expressions are the very same ones as the expressions that, presented out of context, people readily identify as expressions of anguish or grief. They result from the operations of the same neurophysiological mechanisms and are therefore no less true facial expressions of the emotions characteristic of Beatlemania, when they express those emotions, than they are true facial expressions of anguish and grief, when they express them.

This conclusion points up a significant confusion in the Darwinians' theory. The source of the confusion is the assumption behind the Darwinians' division of emotions into basic, general categories according to the true facial expressions of the emotions in those categories. For the assumption behind this division is that the facial expressions defining a basic category are not only true expressions of the emotions in that category but are also never true expressions of an emotion that belongs to a category defined by different facial expressions. And the example of Beatlemania shows that this assumption is false. In the same way, the example shows that the Darwinians' identification of each basic category with the neurophysiological mechanisms whose operations produce the facial expressions defining that category represents a misconception of the nature of the emotions belonging in that category. For it shows that the difference between such emotions as joy, delight, gladness, and elation and such emotions as sadness, sorrow, grief, and dejection cannot consist, not even in part, in their being different types of neurophysiological event whose differentiae are determined by the expressive behavior that manifests them. This misconception is most sharply realized in the Darwinians' ambitious hypothesis that each of the basic categories corresponds to an affect program whose activation produces the type of neurophysiological event that every emotion in that category consists in. If prolonged, intense screaming, whether initiated by the excitement of suddenly seeing the object of one's ardent devotion or the shock of suddenly getting news of a tragic and personal loss, activates the same affect program, then the Darwinians who advance this hypothesis have to hold that the emotion in either case belongs to the same basic category, and therefore two emotions as seemingly opposed to each other as rapturous joy and devastating grief are really generically the same. This result may, to be sure, leave our philosophical enthusiast for this hypothesis undeterred. He may just shrug it off with the remark, "Well, modern biology teaches us that birds and lizards are generically the same, so why think this result is any weirder?"<sup>32</sup> But most of us would not be so sanguine. For most of us the incoherence of the result is sufficient to warrant adding the hypothesis to the large class of failed theories in scientific psychology.

Elsewhere I have written that the great changes in our understanding of emotions that took place in the twentieth century are rooted in the ideas of William James and Sigmund Freud.<sup>33</sup> James's ideas are the source of the view that one can fruitfully study emotions by studying the neurophysiological processes that occur with experiences of them. Of course, James did not identify emotions with these neurophysiological processes. He identified them with feelings. His famous definition is that emotions are the feelings of the bodily changes that "follow directly the perception of an exciting object."<sup>34</sup> But on this definition, emotions become epiphenomena, and the proper object of study becomes the physiological processes the feelings of which are identified with emotions. Freud's ideas are the source of the view that emotions transmit meaning or purpose to the feelings and behavior that manifest them. Though Freud often described emotions as flows of nervous energy, his view of them as transmitters of meaning and purpose was nonetheless implicit in his notion of an unconscious emotion and in the way he used this notion to make sense of feelings, behavior, and physical maladies that seemed otherwise inexplicable. Widespread acceptance of his explanations has thus led to studying emotions for the ways they render feelings, behavior, and bodily conditions meaningful products of the mind. Theorists of emotion who develop their theories from an understanding of emotions as phenomena common to humans and beasts are readily drawn to the view of how to study them that comes from James, for the neurophysiological mechanisms in human beings on which such studies focus are homologous to neurophysiological mechanisms in other animals, and by appeal to these homologies they can then explain how humans and beasts are liable to many of the same emotions. Theorists of emotion who take as their leading idea that emotions are intentional states and develop cognitivist theories based on this idea accept the view that comes from Freud, for to make thought essential to emotions is to introduce an element in emotions that can explain how emotions give meaning to the feelings, behavior, and bodily conditions they produce. These two different programs have, relative to the chief fact about emotions each takes as central to understanding the phenomena, yielded powerful and illuminating theories. The main problem for the study of emotions now is how to develop a theory that reconciles these two facts.

#### NOTES

I am grateful to Dan Brudney, Russell Dancy, Joshua Gert, Martha Nussbaum, and the audiences at Johns Hopkins University and Florida State University, where I presented an earlier draft of this essay.

1. Shakespeare, *Romeo and Juliet*, I.i.42–48.

2. I am ignoring the possibility of objectless emotions here. Whether there are such emotions and how, if there are, their existence can be squared with this

first fact are issues on which theorists of emotion disagree. Their disagreement, however, does not affect the argument of this essay. For discussion of these issues see Deigh 1994, 824–54.

3. See *ibid.* for a survey and critical discussion of these theories.

4. The best and most sustained effort at providing such an explanation is Nussbaum 2001, 89–138. I have discussed Nussbaum's explanation and raised some objections to it in Deigh 2000, 293–307. For Nussbaum's replies, see Nussbaum 2000, esp. 358–62.

5. By an equivalent, I mean a code like Morse code whose meaningful strings of symbols one must translate into a language to recover their meaning.

6. Locke 1975, 157–58 (book 2, chap. 9, secs. 5–7), 159–60 (book 2, chap. 9, secs. 10–11).

7. Descartes, 185–86.

8. Locke 1975, 159 (book 2, chap. 11, sec. 9).

9. Among the Darwinians, I include Silvan S. Tomkins, Robert Plutchik, Carroll E. Izard, and Paul Ekman. For representative writings, see Tomkins 1962 and 1963, vols. 1 and 2; Plutchik 1980; Izard 1977; Ekman 1980, 73–102; and Ekman 1984, 319–43.

10. Darwin, 33n.

11. Locke 1975, 229–30 (book 2, chap. 20, sec. 3).

12. Descartes 1967, 380–90.

13. See Ross 1984.

14. Darwin, 7.

15. *Ibid.*, 17–19.

16. *Ibid.*, 280–81.

17. *Ibid.*, 99–100.

18. "With mankind some expressions, such as the bristling of the hair under the influence of extreme terror or the uncovering of the teeth under that of furious rage, can hardly be understood except on the belief that man once existed in a much lower an animal-like condition." *Ibid.*, 19.

19. "Whenever the same movements of the features or body express the same emotions in several distinct races of man, we may infer, with much probability, that such expressions are true ones—that is, are innate or instinctive. Conventional expressions or gestures, acquired by the individual during early life, would probably have differed in different races, in the same manner as do their languages." *Ibid.*, 22.

20. There appears to be some uncertainty among the Darwinians about whether and how to distinguish the last two categories.

21. See Ekman's comment about these facial movements in Darwin, 256–57.

22. This conception is implicit in the Darwinians' belief that they are studying the very nature of emotions by studying the true expressions of emotion. In their statements of what an emotion essentially consists of, however, some of the Darwinians add other components besides neurophysiological processes. Izard (1977, 4), for instance, includes subjective feelings or some inner experience. But he seems to include this component as a concession to common sense. It is not well integrated into his study of the facial expressions of emotion, which is the major part of his research. Furthermore, he does not entertain the possibility of

the subjective experience of an emotion occurring in the absence of the activation of the neurophysiological mechanisms that produce the expressions of the emotion. His ignoring it strongly suggests that the subjective feeling is merely a parasitic element in his conception, for it is hard to see how someone could acknowledge the possibility that the subjective feeling or inner experience of anger, say, could occur in the absence of the activation of the neurophysiological mechanisms that produce the expressions of anger without having to conclude that the subjective feeling, the inner experience, is the emotion and the operations of the neurophysiological mechanisms that produce its expressions are just its normal material base. Ekman (1980, 79–82) evades defining what an emotion is. Emotion, he says, is complex, and there are many things that are relevant to its study such as subjective experiences and coping behavior. But his research, which concentrates on facial expressions and correlative processes in the autonomic nervous system, leaves no doubt that he is working with a conception of an emotion as a neurophysiological event.

23. See Ekman 1984, 324–28.

24. Griffiths 1997, 11–14.

25. See Ekman 1980, 80–84. Ekman cites Tomkins as the source of the idea of an affect program and the inventor of the term.

26. Some philosophers hold that the cause and the object of an emotion are always and necessarily distinct, and the reason they hold this position is that they understand the thought in virtue of which an emotion is an intentional state as having indirect rather than direct (in Frege's sense) reference to the world. My view is different. On my view, some emotions are intentional states, even though the thought in virtue of which they are intentional states makes direct reference to the world. The difference between the two views raises issues outside the scope of this essay and does not affect the argument I am making here. For a statement and defense of the position that the cause and the object of an emotion are always and necessarily different, see Solomon 1977, 172–85.

27. Malcolm 1977, 49.

28. This remark, of course, requires some gloss to avoid taking it as implying that the dog is capable of propositional thought. What we must mean, then, when we say the dog thinks the cat is up the tree—and we say such things all the time—is something like the dog has been impressed with images of the cat's going up the tree and these images are guiding his behavior. Malcolm's response to the issue is to distinguish between saying the dog thinks such-and-such and saying the dog has the thought that such-and-such. According to Malcolm, the latter implies propositional thought but the former does not. See Malcolm, 50–51.

29. See *The Beatles Anthology* (a video history of the Beatles; EMI, 1996).

30. As noted above (n. 20), there is some uncertainty among the Darwinians about whether and how to distinguish these two categories.

31. Darwin, 146–75.

32. See Griffiths 1997, 77–79.

33. Deigh 2001, 1247–56.

34. James 1950, 2:449.

## Emotion

### *Biological Fact or Social Construction?*

JENEFER ROBINSON

#### EMOTIONS AS JUDGMENTS

Currently the most widely favored theory of emotion among philosophers is the “judgment” theory of emotion. According to this theory, at the heart of emotion is a cognitive state: an emotion either is or essentially includes a judgment or belief. If I am in love with Joe, this means not just that I get warm and fuzzy feelings inside when Joe approaches, but that I have certain beliefs about Joe—that he is a worthy, lovable kind of guy. Similarly, if I am afraid of a bear, I don’t just experience a twinge or a pang; I believe or judge the bear to be dangerous or threatening to me. At the heart of love, it would seem, is the judgment that the beloved is a wonderful person; at the heart of fear is the judgment that I am being threatened.

This theory is plausible as a bit of folk psychology. Being afraid of the bear *does* seem to entail that I believe it is threatening me. Likewise it seems contradictory to say that I love Joe, but there’s nothing about him that I believe to be appealing. Furthermore, a change in the relevant evaluative judgment may ipso facto produce a change in one’s emotional state. I cannot be angry that you have insulted me if I learn that you did not in fact insult me. If I thought you said “You cow!” and then I discover that you really said “Oh wow!” my anger is likely to change to relief and amusement. A change in the belief or judgment seems to entail a change in the emotion and/or the abandonment of the emotion (Solomon 1976, 1980).

Moreover, we argue with people about their emotions; we say that I *should not* be angry with you, that your fear is *unjustified*, that you *ought* to be ashamed of yourself, and so on. This suggests that we are arguing

about evaluative judgments: you are trying to convince me that I am right or wrong to make a particular evaluative judgment. If emotions were nothing but feelings, argument would be beside the point: normally you would not try to argue somebody out of a pang or a pain.

One thing that judgment theorists tend to agree upon is that the judgments involved in emotion are *evaluative* judgments about a situation in terms of one's own *wants, wishes, values, interests, and goals*. These evaluations are evaluations of the personal significance of something going on in the external or internal environment—either the external environment of other people, things, and events, or the internal environment of one's own thoughts, memories and imaginings. As one of the judgment theorists puts it, an emotion involves “an evaluation of some object, event or situation in the world about me in relation to me, or according to my norms” (Lyons 1980, 59). Another goes so far as to say that our emotions are “the very core of our existence, the system of meanings and values within which our lives either develop and grow or starve and stagnate” (Solomon 1976, xvii).

Another thing judgment theorists tend to agree about is that the way to *distinguish* one emotion from another is by the evaluative judgments they embody: anger involves a judgment that one has been offended, sadness that one has suffered a loss, fear that one is in danger, and so on.

#### PROBLEMS WITH THE JUDGMENT THEORY

The idea that being in an emotional state either is or entails making an evaluative judgment has some serious problems, however. On the one hand, it seems as if you can make an evaluative judgment of the appropriate kind, yet not be in the corresponding emotional state, and on the other hand, it turns out that you can be in an emotional state without making any judgment of the sort the judgment theory has in mind. I will address these issues in turn.

To see that one can make the appropriate evaluative judgment yet fail to be in the corresponding emotional state, it is enough to note that I can judge that you cut me off in traffic and that this was offensive and insulting, without getting angry: I may be resigned or saddened or even cynically amused. Or I may simply judge dispassionately that I have been offended without getting emotional about it at all. Likewise I can judge that I have mistreated my children and that this is bad, without being ashamed. I may be making an emotionless judgment, or I may experience another emotion: resignation at my bad character or heartless rejoicing in it.

Judgment theorists have recognized this objection and tried to rebut it. Robert Solomon has suggested that an emotion is a special *kind* of judgment: emotions are “self-involved and relatively *intense* evaluative judgments. . . . The judgments and objects that constitute our emotions are

those which are especially important to us, meaningful to us, concerning matters in which we have invested our Selves” (Solomon 1976, 188). Elsewhere he characterizes emotions as “urgent” judgments: “Emotional responses are emergency behavior” (Solomon 1980, 264). An emotion is “a necessarily hasty judgment in response to a difficult situation” (265).

However, to call a judgment “intense” seems odd unless you are already convinced that some judgments are emotions: emotions may be intense, but judgments are not normally thought of as the sorts of things that admit degrees of intensity. And at least some emotions do not seem to require hasty judgments: I can start to be *afraid*—say, about the state of my stock portfolio—only after months of painstaking statistical analysis.

William Lyons (1980) suggests that while an evaluative judgment all by itself is not sufficient for an emotion to occur, if the evaluation causes an “abnormal physiological response,” this *is* sufficient for emotion to occur. This is a valuable suggestion that will in a certain sense turn out to be accurate. The trouble with it is that it does not explain why sometimes an evaluative judgment leads to physiological change and hence emotion, while at other times what appears to be the very same evaluative judgment *fails* to lead to physiological change and emotion. Thus, the judgment that I have been insulted and offended sometimes produces an angry emotional response, but sometimes it produces resignation, sometimes sadness, and sometimes merely a philosophical shrug of the shoulders. The very same judgment with the same propositional content sometimes produces physiological change and sometimes not. This is something that needs to be explained. But Lyons’s suggestion is useful because it suggests that even if emotion requires an evaluative judgment, it also requires physiological changes.

#### EMOTION AND PHYSIOLOGICAL CHANGE

William James, the father of modern psychology, held that without a consciousness of physiological changes, an emotion would reduce to a “cold and neutral state of intellectual perception” (James [1890] 1981, 1067). In other words, a judgment or cognitive state all by itself can never produce an emotional state; it is physiological change that puts the emotionality into emotion. I may disapprove of oil drilling in nature preserves, but unless I am also undergoing characteristic physiological changes, we don’t say that I am emotionally upset about it.

It is still an open question whether specific emotions can be identified by means of specific autonomic changes. Some theorists have thought that only generalized arousal is necessary for emotion. Walter B. Cannon ([1929] 1963) identified the characteristic physiological profile of the “emergency” or “fight or flight” reaction (increase in muscular blood flow, increase in

activity of sweat glands, increase in blood glucose level, etc.), but although there is good evidence that this reaction occurs in situations of stress, it does not occur in all emotional states. In particular, it is not associated with more passive emotional states such as sadness.

There is some evidence of systematic autonomic differences among some of the emotions. Thus anger, fear, and sadness all produce a larger increase in heart rate than does disgust, and anger produces a larger increase in finger temperature than fear (Levenson 1994, 255). But such results are hardly enough to justify the claim that each emotion has a uniquely identifying physiological profile.

The only really good evidence that there are distinct physiological changes associated with distinct emotions comes from studies of facial expressions across cultures. Paul Ekman has conclusively demonstrated that for several emotions or emotion families there are universal facial expressions. He has shown that there is remarkable agreement among different researchers using different methods about the facial expressions of anger, fear, enjoyment, sadness, and disgust.

This evidence is based not just on high agreement across literate and preliterate cultures in the labeling of what these expressions signal, but also from studies of the actual expression of emotions, both deliberate and spontaneous, and the association of expression with social interactive contexts (Ekman 1992, 176).

Ekman has carried out several different sorts of experiment. Typical is one study focused on people in New Guinea who had had no prior contact with Westerners. In one set of experiments, subjects were told a story and asked to pick out from photographs the facial expression most appropriate to the story. In another set of experiments, subjects were asked to pose the expression itself; their facial expressions were then videoed and shown to students in the United States, who were asked to identify the expressions.

Interestingly, Ekman has also found evidence that physiological changes characteristic of a particular emotion can be induced "directly" by manipulations of the facial musculature. When subjects moved their facial muscles without knowing what expression they were being asked to configure, there was activity in both skin temperature and heart rate distinctive of the particular emotion (Ekman 1984). In other words, facial expression all by itself is sufficient to produce autonomic nervous system changes characteristic of particular emotions. His general conclusion is that there are at least five *basic emotion systems*, which are found in all human beings (and perhaps in other species), consisting of identifiable physiological changes and facial expressions that are found universally in particular stimulus situations. (Ekman usually also includes surprise on his list, and sometimes contempt.)

The second problem with the judgment theory is that there seem to be instances of emotion without cognition, or at least where a cognitive evaluation is not a necessary precondition for emotion.

First, there is good evidence that our earliest emotions or affects do not require complex cognition. The founder of behaviorism, John Watson (1929), performed some notorious experiments on newborn babies. He found that restraining the newborn's ability to move its head was a universal stimulus for rage, that sudden loss of support (dropping the baby!) was a universal stimulus for fear, and that gentle caresses universally elicited pleasure (although Watson dubbed this response "love"). Yet newborn babies do not seem to have the cognitive resources necessary to make cognitive evaluations such as "That was an offense!"

Nico Frijda has studied what he calls "unlearned stimuli" for emotions among several different species. He points out that many species—including humans—respond "instinctively" to the strange and weird. For example, dogs and apes are afraid of people dressed or behaving in unfamiliar ways, horses are afraid of flapping plastic bags, and Rhesus monkeys are terrified of mechanical moving monsters. Moreover, other fear responses, while not inbuilt or present at birth, are easily acquired with very little learning, such as human fear of spiders and snakes (Frijda 1986).

However, the best evidence that emotions can occur prior to any cognitive evaluation comes from data accumulated by the psychologist Robert Zajonc (Zajonc 1980, 1984, 1994; Kunst-Wilson and Zajonc, 1980; Murphy and Zajonc, 1993). Zajonc claims that at least in some cases, "affect" occurs without any prior "cognition" or cognitive appraisal. Psychologists tend to talk about "affect" and "cognition," rather than "emotion" and "judgment," but both philosophers and psychologists are talking about more or less the same thing. In claiming to show that affect precedes cognition, Zajonc is mounting a direct threat to the judgment theory. I will describe three of the most compelling sets of experiments that he describes.

1. In the "mere exposure effect" experiments, Zajonc and others have convincingly established that people prefer stimuli to which they have been exposed more often, even when the stimuli are presented so fast that the subjects cannot consciously recognize what they are seeing. For example, in one experiment, Zajonc and a colleague flashed slides of irregular polygons to a group of subjects, which—it was independently established—appeared too fast for recognition. The subjects were asked to discriminate which of two polygons they liked better and which they had seen before. In this study the rate of recognition was virtually equivalent to chance, but the liking responses reliably distinguished between those polygons that were "old" or (in some sense) "familiar" and those that were "new" or "unfamiliar" (Kunst-Wilson and Zajonc 1980).

2. More recently, Zajonc has undertaken a series of experiments designed to show that new affective reactions can be induced apparently without the intervention of cognition. In these experiments he and his colleagues have studied what he calls "nonconscious affective priming," that is, inducing an affective reaction to a "neutral" stimulus such as a Chinese ideograph (presented to someone who does not know Chinese) by "priming" the stimulus with an "affective" picture, such as a picture of a happy or an angry human face. The "prime" is exhibited to the experimental subjects so briefly that they are unconscious they have seen anything, and yet the Chinese ideographs that have been primed by a smiling face are much preferred to those primed by an angry face. However, when the ideographs are presented more slowly so that they can be consciously recognized, this priming effect vanishes (Murphy and Zajonc 1993).

3. Among the most compelling evidence that Zajonc cites are some classic experiments carried out in the early 1950s by Lazarus and McCleary, called the "subception" experiments (Lazarus 1991). In these experiments people were presented with ten five-letter nonsense words. The subjects' galvanic skin responses were then conditioned to five out of the ten, using electric shock as the unconditioned stimulus. In other words, five out of the ten "words" were associated with a shock, and when the subjects saw those "words" they responded with a heightened galvanic skin response. Such a response is part of the "emergency" response studied by Cannon. All ten stimuli were presented an equal number of times, so that the subjects would be equally familiar with all ten. After the subjects had been conditioned in this way, a test was administered in which the "words" were presented at speeds too fast for recognition and the subjects were asked to say which "word" they had "seen." In the test itself no shocks were administered, but the subjects were not told this. The remarkable result of this experiment was that the galvanic skin response was much higher for "words" associated with shock than for those associated with nonshock, *even when the subjects were unable to identify* the words that had been presented to them. Lazarus called this effect "subception," indicating a kind of perception that occurs below awareness: the subjects did not know which "word" had appeared to them, but their skin did. We can infer that the subjects *feared* or *disliked* certain "words" even though they did not know what they were.

#### AFFECTIVE EVALUATIONS

What Zajonc's data suggest is that there is an *affective* or *noncognitive* appraisal that concerns those things that "matter" to the organism; that occurs very fast, automatically, and below the threshold of awareness; and that produces physiological changes such as an increased galvanic skin response. Zajonc himself interprets his results as evidence that affect pre-

cedes cognition and that cognitive evaluation is therefore not necessary to emotion. Yet the data he cites can just as plausibly be construed as showing that at least some emotions involve primitive evaluations that occur below consciousness and independently of “higher” cognitive processing. Examples include the preferences evinced by subjects in the mere exposure experiments, in which subjects “evaluated” the tones or polygons they had been exposed to more often as pleasanter than the others. Similarly, in the subception experiments, subjects “evaluated” some syllables as threats. Of course, whether these evaluations are classified as “noncognitive” and “affective” depends on what you mean by affect and cognition. However, if an automatic appraisal of pleasure, aversion, or novelty counts as affect, virtually everyone now agrees that affect can occur prior to any higher cognitive processing.

From an evolutionary perspective, it seems plausible to speculate that fast, automatic appraisals are adaptive just because they take place more quickly than “higher” cognitive processes and instantly get the organism into a state of action readiness. It may be very important to us to know immediately whether something in the environment is friendly or hostile, is a threat or a poison, and to get ready to respond accordingly. Remember that emotional appraisals concern matters that are of significance to me and mine—to our survival and/or our well-being—and that emotions focus our attention on those aspects of the environment that we appraise as crucial to our survival and well-being. When we make emotional appraisals of the environment, we are not asking questions of it in a spirit of philosophical inquiry or intellectual curiosity. We are appraising some aspect of the world in terms of its implications for our survival and well-being and that of our “group.”

Emotional appraisals result in autonomic nervous system changes and motor activity, including facial expressions. In many cases it may be important to signal one’s emotional state and to recognize someone else’s very fast. The angry gorilla’s facial expression alerts other gorillas that they had better back off if they want to avert being attacked. Similarly, if a human being looks angrily at me, I know immediately I had better back off and look humble if I want to avoid trouble. My appraisal of your expression may be so rapid that it’s unreliable. But it is better to have an organism that responds fast and mistakes your smile for a snarl than one that pauses to reflect and gets attacked.

These results suggest that one important biological function of emotion is to pick out from the multitude of competing stimuli those that are important to the wants, goals, and interests of the organism and which need to be “dealt with” as a matter of immediate urgent attention. The subsequent motor and autonomic changes signal what state the organism is in and prepare the organism for subsequent action. Central to emotion is an *affective appraisal* that very rapidly assesses the affective significance of the

stimulus—that is, what is at stake for the wants and interests of the organism in this particular encounter with the environment.

THE JUDGMENT THEORY: A PERSPECTIVE  
FROM NEUROPHYSIOLOGY

In a series of articles culminating in his 1996 book *The Emotional Brain*, the neurophysiologist Joseph LeDoux has argued that there is no unitary phenomenon called “emotion,” but only a variety of emotion systems. Emotions are not peculiar to humans or “higher” animals; the *basic* emotion systems are to be found in many “lower” species, even insects and fish. These basic emotion systems are designed so that the organism can cope with fundamental life encounters. It is important to think of a basic emotion system as adapted through evolution to particular sorts of important interaction between the organism and its environment. LeDoux thinks that a list of basic emotions would correspond to a list of “special adaptive behaviors that are crucial to survival” (1996, 126). His “working hypothesis” is that “different classes of emotional behavior represent different kinds of functions that take care of different kinds of problems for the animal and have different brain systems devoted to them” (127).

The emotion system that LeDoux has studied most extensively is the fear system. The object of his study has been conditioned fear in rats, but his research has wide and important implications for naturally occurring fear—including fear in humans—as well as for the study of emotion in general. LeDoux thinks that whatever your theory of emotion, the “core” of an emotion system is “a mechanism for computing the affective significance of stimuli” (1989, 271). In other words, the organism can somehow “appraise” or evaluate the emotional significance of a stimulus. As I have explained, Zajonc’s results (among others) show that such an “appraisal” can take place very fast and prior to any conscious cognition or complex information processing. LeDoux’s great contribution to emotion research has been to show how this can happen. He and his colleagues have discovered a fear circuit in the brain that operates very fast and without awareness; it can compute the affective significance of a stimulus without the organism’s being able to recognize what the stimulus is.

According to LeDoux, the key to the fear system in the brain is the *amygdala*, a small almond-shaped formation where the emotional significance of threat is registered. LeDoux found that when rats that have been conditioned to fear the sound of a buzzer hear that sound, the *auditory thalamus* is activated. He notes that the thalamus has been called a “way station” in the brain, because it receives incoming stimuli and sends them on to different parts of the cortex that are responsible for “higher” processing in the various different sense modalities (vision, hearing, etc.). The auditory thalamus receives auditory signals and sends them on to the *au-*

*ditory cortex*, where the sound is “cognitively processed”—for example, the sound is identified. The auditory cortex then sends signals to the amygdala, where the emotional impact of the buzzer sound is assessed. However—and this is the important point—the auditory thalamus also sends signals *directly* to the amygdala, *bypassing the cortex* altogether. The amygdala computes the affective significance of the stimulus *before* the more precise information about what the stimulus is gets received from the auditory cortex. Significantly, for a rat it takes about 12 milliseconds (12 one-thousandths of a second) for an acoustic stimulus to reach the amygdala through the direct pathway from the thalamus, whereas it takes almost twice as long for the cortical route. The thalamo-amygdala and cortico-amygdala routes converge in the lateral nucleus of the amygdala. Once in the lateral nucleus, the signal is dispatched to the central nucleus of the amygdala, which controls the characteristic *fear responses*: the freezing response, the release of stress hormones into the bloodstream, and the various cardiovascular and other visceral responses, including increased blood pressure, sweating, and increased galvanic skin response. If information from the thalamo-amygdala route is coordinated with information from the cortico-amygdala route in the lateral nucleus of the amygdala, then presumably it would be possible for the information derived from the auditory cortex to confirm or disconfirm the information received directly from the thalamo-amygdala route; it could confirm whether the “affective appraisal” is appropriate or not, and whether the responses generated should be maintained or aborted.

Significantly, the auditory thalamus cannot make very fine discriminations in a stimulus. It is in the auditory cortex that fine discriminations are made. LeDoux concludes that he has discovered two different pathways for processing the same sound. On the one hand, there is a “quick and dirty processing system,” which responds very fast, warns the organism that something dangerous may be around without identifying it very carefully, and gets the organism to respond appropriately to whatever it is. And on the other hand, there is a slower, more discriminating processing system, which operates through the cortex and figures out whether the thalamo-amygdala’s “affective appraisal” is appropriate or not (LeDoux 1989, 1996).

LeDoux proposes that emotion and cognition should be thought of as “separate but interacting mental functions mediated by separate but interacting brain functions” (1996, 69). The initial response to fear is generated by an “emotional appraisal” in the amygdala that happens very fast and prior to cognitive intervention. The subsequent slower “cognitive appraisal” can identify the stimulus more carefully, assess the appropriateness of the prior automatic response, and presumably attempt to modify and control both the initial appraisal and the organism’s subsequent responses.

As we saw earlier, Solomon claims that emotions are “self-involved and relatively *intense* evaluative judgments” that are “urgent” and “especially important to us, meaningful to us, concerning matters in which we have invested our Selves” (1976, 188); emotional responses are “emergency behavior” (1980, 264). The notion of a noncognitive affective appraisal explains very well the features of emotional “judgment” that Solomon indicates. An affective appraisal operates very fast or “urgently” in situations where the person or other organism has a vital stake in what’s going on; it concerns matters “which are especially important to us.” The appraisal immediately gets the person physiologically prepared for possible “emergency” action. The whole process is marked by “intensity”: what is happening is very important to the person and demands his or her exclusive attention. But, strictly speaking, what is intense is not so much the appraisal as the physiological activity that it induces. This activity in turn helps keep the attention focused on what is perceived as of urgent importance and prepares for emergency action, if necessary. Lyons believes that an emotional response is a physiological response caused by an evaluative judgment. This too has some truth to it. The physiological responses characteristic of particular emotions are indeed caused by affective appraisals.

At the same time we seem to be faced with a dilemma. On the one hand, we have evidence that emotional responses are caused by noncognitive affective appraisals, but these affective appraisals are simple appraisals perhaps conceptualizable as “I don’t like this!” or “This is a threat!” or even “Yuk!” These appraisals cannot make fine discriminations. And on the other hand many of the emotions for which we have names in our language seem to be definable in terms of complex cognitions. Given LeDoux’s description of the quick and dirty pathways they use, it would seem impossible for rough and ready affective appraisals to distinguish the subtle differences between shame and guilt, jealousy and envy. Yet complex cognitions are not sufficient all by themselves to cause an emotional response.

LeDoux’s model of what goes on in a simple episode of emotion suggests an answer to our dilemma. In the simple cases he has studied, there is an affective appraisal that causes autonomic and motor changes and is succeeded by cognitive monitoring. In other words, even a simple episode of emotion is a *process*, involving a number of different events, and, in particular, involving *both* affective *and* cognitive evaluations. On this model it is the affective appraisal that generates autonomic and motor changes and puts the emotionality into emotion, as James might have said. And presumably these affective appraisals distinguish in a rough and ready way between different emotion systems such as fear, anger, sadness, and enjoyment. But it seems likely that more subtle distinctions are made cognitively,

that since cognitive appraisal and reappraisal are also part of the emotion process, there is plenty of opportunity for complex cognition to play a causal role in the development of the emotion process, so that what we call different emotions in ordinary language will be the result of different emotion processes involving different cognitive evaluations.

It is important to remember that in an emotion process there is continuous feedback of various sorts from one event in the process to another. For example, physiological changes may help to fix our attention on whatever it is that is important to us in the situation. Autonomic arousal may prompt us to action. Subjective feelings may also reinforce attention and in addition can serve as a source of information about what I am reacting to and how important it is to me. Cognitive appraisals and reappraisals can modify action tendencies, physiological changes, and subjective feelings alike. After responding instinctively by freezing or tensing, I may see that the situation is no longer threatening, so I relax, my heart gradually stops pounding, and my feelings gradually become calmer. And action tendencies induced by an affective appraisal, such as freezing in fear or tensing in anger, often alter the situation itself—perhaps the enemy retreats—and so indirectly modify my appraisal of the situation and thus in turn affect my subsequent behavior.

The psychologist Phoebe Ellsworth (1994) has pointed out that the various events that occur in an emotion process are themselves processes that to some extent unfold independently and run their own temporal course. Thus autonomic arousal typically continues even after cognitive monitoring has changed my appraisal of a situation. One consequence of this more complicated picture of emotion is that Solomon is wrong to say that once my evaluation that I have been wronged (say) changes and I realize I have not been wronged after all, my emotion of anger will vanish. In fact the physiological changes associated with anger will not—indeed, cannot—vanish as soon as a change in judgment occurs.

Ellsworth also stresses that our emotional life occurs in “streams” that change all the time in response to ever changing appraisals, ever evolving actions and action tendencies, ever changing bodily states and feelings. It is only after the process is over that we can catalog the process “in recollection” (as she puts it) as belonging to “one of the familiar emotions catalogued by ancient and modern taxonomists” (228, 1994).

Suppose Joe, my husband of thirty years, suddenly runs off with a twenty-five-year-old model. I will probably be overwhelmed by a stream of different emotions, by different affective appraisals, unpleasant feelings, and a welter of physiological changes: at one moment I am raging, at the next crying, at one time I am lethargic, at another manic. I may be confused about what I believe about the situation, and about which wants of mine are most at stake. At one time I think I have been betrayed, at another that I am worthless and deserve to be abandoned; at one time I think that I

have suffered an irretrievable loss, at another that such a scoundrel is not worth keeping. Or I may be barely conscious of thinking anything. In any event I can have only the dimmest of ideas about the sequence of affective appraisals and cognitive reappraisals that actually occurred in temporal sequence. Later, however, when I have had a chance to reflect upon my various emotional reactions, I can catalog them in recollection. I can describe my state in terms of the emotion concepts familiar to me in English: I was resentful, grieving, depressed, jealous, frightened, and so on. How I catalog this stream of events in reflection is largely a function of the emotion concepts at my disposal in the language and culture that I inhabit. And of course I might be wrong in my assessments. I may decide that my main emotion is resentment, whereas my behavior, physiological states, and facial expressions show that my predominant reaction is best described as grief.

How we categorize our emotions will be largely a matter of the concepts available to us in our culture. In Japanese culture one can experience the emotion of *amae*, a propensity to “depend or presume upon another’s love,” which according to some “social construction” theorists is not an emotion that is experienced—or experienced in the same way—in the West (Morsbach and Tyler 1986, 290). Similarly, on the Pacific island of Ifaluk, people experience *fago*, translated by Catherine Lutz as “compassion/love/sadness” (Lutz 1988, 119). When we “catalog” an emotion, we typically recognize a particular kind of situation as cognitively appraised in a particular way in the light of a particular set of wants and interests. Given different situations of life, different values and interests, and so on, it is not surprising that different cultures will catalog in different ways. In general, however, explaining behavior by reference to ordinary language concepts for emotions—jealousy, envy, *amae*, *fago*—is making an after-the-fact cognitive evaluation in the terms of folk psychology, summarizing a particular emotion process, a particular sequence of events.

#### BEING ANGRY AND BEING EMOTIONAL

One question I have so far avoided broaching is whether a state of “anger,” say, is always a genuinely emotional state. In my view the answer—paradoxically—is no. I think we need to make a distinction between anger as a certain kind of cognitive or attitudinal state and anger as part of an emotional process. As I remarked at the very beginning of this chapter, I can judge that I have been offended and I can dislike this yet fail to be in an emotional state—fail to make the requisite affective appraisal and to undergo its concomitant physiological changes. But in some sense it seems correct to describe me as “angry.” If that is so, then “anger” refers, perhaps derivatively, not just to an emotional state but also to a related cognitive or attitudinal state. At the same time, recognizing an offense and disliking the

offense do typically result in the affective appraisal characteristic of the emotion anger.

One of the reasons in favor of the judgment theory was that we argue about emotions, saying that you ought not be angry or that your jealousy is unjustified. We can now see that such arguments concern the evaluative or cognitive judgments that distinguish a particular emotional state rather than the emotional state itself. My belief that I have been offended may be unjustified. My affective appraisal “I don’t like this!” or “Yuk!” is immune to any kind of justification. It is fast, automatic, and not directly within our control.

Earlier I criticized Lyons for saying that an emotional response is caused by a cognitive evaluation of an event without explaining why apparently identical cognitive evaluations (“He wronged me!” “I am in danger!”) may sometimes issue in an emotion of anger or fear and sometimes not. An answer now suggests itself. A cognitive evaluation all by itself is not enough to generate an emotional response; an affective appraisal is required as well. Affective appraisals are fast and automatic, and they instantly produce the physiological responses characteristic of emotion. It seems not implausible to suppose that the initiator of a genuinely emotional process will always be an affective or noncognitive appraisal, which appraises the situation in terms of one’s wants and interests and generates action tendencies as well as autonomic and motor responses, including facial and vocal expressions that alert others to the state of the organism. This affective appraisal will then itself be monitored by ensuing *cognitive* appraisals.

We might still wonder, however, why I sometimes respond emotionally to an event or situation—and make an affective appraisal of it—and sometimes not. Why do I sometimes get angry with my small son and sometimes not, even when he is doing exactly the same thing on both occasions? There are a number of reasons, I think. One is that an affective appraisal is probably more likely when the wants and interests that are at stake in an encounter are intensely important to us. Another is the vividness of the perception or mental image that prompts the affective appraisal: we are more likely to get emotional about refugees when we see them on TV than if we read about them in the newspaper. A third relevant fact is our bodily state, whether we are fatigued or hyped up on caffeine or under the influence of alcohol. Our current mood will also be a factor. The tired and hungry mother beset with anxiety about other things is much more likely to get upset with her small son. Affective appraisals occur more readily in some bodily and mood states than others. And then there are individual differences among people: some people are more emotionally labile than others and readily change mood and emotional state.

In the remainder of this chapter I would like to show that the picture of emotion I have painted is consistent with folk psychological accounts of emotion in terms of our everyday emotion concepts, so that we do not need two entirely different accounts for basic emotions and for more cognitively complex emotions (as Paul Griffiths [1997] has recently argued). My suggestion is that there is a set of inbuilt affective appraisal mechanisms, which in more primitive species and in neonates are automatically attuned to particular stimuli, but which, as human beings learn and develop, can also take as input more complex stimuli, including complex “judgments” or thoughts. Furthermore, although these affective appraisals are at the heart of the emotion process, they always give way to cognitive appraisals and reappraisals that may put an end to the emotion episode or modify it in various ways. The fear system responds not just to “large black bear” but also to much subtler threats requiring cognitive processing, such as the veiled insult from my boss or a potential loss on the stock market. And although its results may always include some of the symptoms of the emergency reaction (an increased heart rate and so on), subsequent cognitive appraisal and reappraisal may result in a wide variety of other behaviors.

In short, I am suggesting that it is always an affective appraisal that *initiates* an emotional response, but that this affective appraisal may itself be an appraisal of some cognitively complex information (“My boss insulted me”). Furthermore, although it causes physiological changes willy-nilly, it will give way to further cognitive activity, which will modify my responses. (Although my heart is racing, I realize it’s best if I smile cheerfully.) In retrospect an emotion process of this sort may be catalogued in recollection as one of the emotions we recognize in our folk psychology (in this case, fear). Let us now briefly examine why this hypothesis is reasonable as a general hypothesis about emotion.

The responses studied by LeDoux are very simple responses, in which there is an affective appraisal of a simple stimulus followed by characteristic fear responses. However, it is perfectly possible that the same neural mechanisms could operate with much more complex input. Consider the case of phobias. In a phobic state, a person has acquired an emotional memory of some traumatic event that may be associated with a specific stimulus. Pat Greenspan (1988) has an example of a somewhat phobic fear of skidding that she has developed as the result of a bad skid in snow. Now even a slight skid in safe conditions on a dry road sets her gasping out of fear. What seems to happen is that a particular stimulus—in this case the skid—has the power to evoke an emotional response in an automatic way. Greenspan has learnt to associate a skid with an affective appraisal of fear, re-

ardless of whether or not she is really in danger. In this case the input is a slight skid on a dry road, but the response is a bona fide fear response. Similarly, Amélie Rorty (1980) has an example of a man who is resentful of women in authority even though he sincerely believes he isn't. Rorty explains the situation as due to his ambiguous feelings about his mother, of whose love he has never been sure. As Rorty tells the story, the man automatically reacts with resentment to a new female boss before he knows anything about her. Regardless of the details of the case, what is interesting for my purposes is that a highly complex cognition—"Here is a woman in authority!"—results immediately and automatically in an unfavorable affective appraisal.

The moral I want to draw from these cases is that an affective appraisal can be evoked by a complex learned stimulus just as easily as by a stimulus that is preprogrammed to produce such a response. In all of these cases a complex stimulus—such as a *thought* that here is a woman in authority—has been fixed in *emotional memory*, so that it automatically elicits emotion. LeDoux has explained the distinction between a declarative and an emotional memory of a frightening event such as a car accident. The declarative memory is mediated by the temporal lobe system and has no emotional consequences: I remember such things as whom I was with and what kind of car I was driving. However, I can simultaneously have an "aversive emotional memory" and a current bodily response that is mediated by "an emotional memory system" such as the implicit fear memory system involving the amygdala. The two memories may be fused in consciousness so that I remember (declaratively) having been emotionally aroused by the traumatic situation; or I may forget the whole thing as far as declarative memory is concerned, and yet a cue, such as the sight of a car exactly like the one I was driving, will trigger an emotional reaction. There is ample evidence from LeDoux and others that there are distinct memory systems for declarative memory and "emotional memory." This shows that the neural pathways exist whereby learned stimuli could instantly produce an affective appraisal and emotional responses.

If this idea that emotion requires affective appraisal is right, it nicely explains why emotions have always been thought of as passive states, and the extent to which this is true. I have suggested that an affective appraisal is an essential part of an emotional state, and an affective appraisal is by definition an automatic appraisal. In the simplest cases, an affective response is preprogrammed in an organism, as when we—and lots of other organisms—respond to a sudden loud sound with the startle response. More complex cases involve wants and interests that have been acquired or learned but that still evoke an automatic appraisal of "I like it!" or "I don't like it!" or perhaps "This is an offense!" or "This is a threat!" So the notion of a noncognitive appraisal produced automatically and automatically resulting in physiological changes explains why traditionally emotions have

been treated as “passions,” as phenomena that act upon us and are not directly under our control, rather than as “judgments” that we consciously and deliberately make. Noncognitive affective appraisals are not controllable except indirectly. I can try to program myself or brainwash myself. I can go to a therapist. But I can’t just decide not to react when I return to the scene of the accident, just as Greenspan cannot just decide not to react when she feels herself skidding.

#### CONCLUSION

The judgment theorists claim to be *explaining* human emotions in folk psychological terms, usually in terms of the wants and beliefs characteristic of a particular emotion. But their explanations do not give a scientific account of the sequence of events in an emotion process. At best they are giving a kind of after-the-fact summary of that process. The actual sequence of cognitions, of affective and cognitive appraisals, of physiological responses, action tendencies, and so on is unlikely to be reliably revealed by introspection or armchair psychology. Yet in a sense the judgment theorist is right after all to insist that we can describe the emotions recognized by our culture in terms of particular beliefs and wants. What the judgment theorist is doing in “explaining” the nature of “jealousy” or “shame” is summarizing in terms that the culture understands what kind of situation has occurred, what beliefs about it the person had, and which of his or her wants, goals, interests and values were at stake. He helps us to organize the emotional terrain in our particular culture.

As I have tried to indicate, there is no reason—yet, at any rate—to reject wholesale the insights of folk psychology into the emotions. Emotions are indeed distinguished by the particular wants and interests that are at stake in some situation and by the way that situation is cognitively evaluated in light of those wants and interests. And these wants and interests and evaluations may be specific to some particular culture. But this is not inconsistent with the idea that it is always an affective appraisal that sets off an emotion process. More precisely, it is the affective appraisal that “evaluates” a situation in terms of a few simple categories (“That’s an enemy!” “That’s a friend!” “I like this!” “I don’t like this!”) and that sets off the physiological activity, action tendencies, and changes in facial expression that constitute the emotional response.

## Embodied Emotions

JESSE PRINZ

In one of the most frequently quoted passages in the history of emotion research, William James (1884, 189ff.) announces that emotions occur when the perception of an exciting fact causes a collection of bodily changes, and “*our feeling of the same changes as they occur IS the emotion.*” The same idea occurred to Carl Lange (1885) around the same time. These authors were not the first to draw a link between the emotions and the body. Indeed, this had been a central theme of Descartes’s exquisite opus *The Passions of the Soul*. But James and Lange wanted to push things farther than most, suggesting that emotions are exhausted by bodily changes or perceptions thereof. Other kinds of mental episodes might co-occur when we have an emotion state. For James, an emotion follows an exciting perception. But the exciting perception is not a part of the emotion it excites. (Ellsworth reads James differently [1994], but see Reisenzein, Meyer, and Schützwohl’s convincing response [1995].) The majority of contemporary emotion researchers, especially those in philosophy, find this suggestion completely untenable. Surely, emotions involve something more. At their core, emotions are more like judgments or thoughts than perceptions. They evaluate, assess, or appraise. Emotions are amendable to rational assessment; they report, correctly or incorrectly, on how we are faring in the world. Within this general consensus, there is a further debate about whether the body should figure into a theory of emotions at all. Perhaps James and Lange offer a theory that is not merely incomplete but entirely off base. While they view judgments as contingent and nonconstitutive concomitants of emotions, it is actually bodily perceptions that deserve this demotion. Perhaps emotions can be, and often are, disembodied in some fundamental sense.

I propose to defend James and Lange, though not completely. They should be criticized for their failure to reckon with what can broadly be regarded as the rationality of emotions. That failure, however, has a remedy that does not depart from the spirit of the James-Lange approach. Emotions are somatic, but they are also fundamentally semantic: meaningful commodities in our mental economies.

I will not be especially concerned with presenting the somatic theory exactly as it appears in James and Lange. I will not assume that emotions are always consciously felt, as James sometimes implies, or that the relevant bodily changes must have the origins in the vasomotor system, as suggested by Lange. The core idea I will defend is that emotions are perceptions (conscious or unconscious) of patterned changes in the body (construed inclusively). I begin by briefly presenting some of the positive evidence for this core idea. Then I discuss six stubborn objections. I argue that the objections can be answered without abandoning the core idea put forward by James and Lange, but they do demand an important amendment.

#### ARGUMENTS SUPPORTING THE JAMES-LANGE THEORY

James and Lange offer several considerations in favor of the hypothesis that emotions are perceptions of bodily changes. Further support comes from more recent work in emotion research. There is no knockdown argument for the hypotheses, but the collective force of several considerations adds considerable plausibility.

The first set of considerations I will consider serves merely to establish a link, or correlation, between emotion and bodily perturbations. The link between emotions and the body is quite obvious. Every culture seems to have bodily expressions for talking about emotional states, from broken hearts in our own culture to bad intestines in Tahiti (see Heelas 1986). Many of these expressions are obviously metaphorical or byproducts of bad folk theories, but there is also empirical evidence for correlations between emotional states and changes in the body. James (1890, 447ff.) quotes long passages from authors, such as Darwin, who had carefully observed such correlations. Darwin's list of fear symptoms includes widely opened eyes and mouth, raised eyebrows, dilated nostrils, stiff posture, motionlessness, a racing heart, increased blood supply to the body, pallor of the skin, cold perspiration, piloerection, shivering and trembling, hurried breathing, dry mouth, faltering voice, fists that are alternately clenched and opened, and so on. One hundred years later, Levensen, Ekman, and Friesen (1991) systematically studied the autonomic changes associated with Ekman and Friesen's (1971) six basic emotions (anger, disgust, fear, joy, sadness, and surprise), and found that each of these corresponds to a unique bodily pattern. It has also been observed that the principle brain structures underlying our

emotional states have all been independently associated with bodily response (Damasio et al. 2000). Some structures, like the amygdala, are involved in orchestrating bodily changes, not in perceiving them. But the amygdala seems to play a role in emotion elicitation and is not essential for emotions themselves. The structures that seem more essential for emotions themselves, such as the insular cortex, second somatosensory cortex, and portions of the cingulate cortex, are associated with bodily perception.

In addition to the correlational evidence, there is evidence that bodily changes can induce emotions. This speaks to the sufficiency of those changes. In particular, it suggests that emotions can arise without the mediation of anything like an appraisal judgment. In making this case, Lange was impressed by the fact that affect could be altered by consuming alcohol. More recent authors have emphasized evidence from facial feedback (Zajonc, Murphy, and Inglehart 1989). Mere change in facial musculature seems sufficient for an emotional response, even when we do not realize we are making emotional expressions. There is also anatomical evidence that emotions can be elicited via pathways from early visual structures, such as the pulvinar and superior colliculus, to the amygdala, which instructs other structures to perturb the body (Ledoux 1996; Morris, Öhman, and Dolan 1999). These pathways trigger an emotional bodily response without the mediation of any kind of judgment. The relevant perceptual centers don't support even categorical object recognition, much less sophisticated appraisal, and the amygdala pairs inputs with somatic responses by association, not by assessment. One could try to argue that the bodily states induced by this pathway would not qualify as emotions unless we supplemented them with appraisal judgments, but that would be desperate. The said bodily changes feel like emotions, and they do not require appraisals to occur.

Showing that bodily changes are sufficient does not establish that the somatic theory is true. For that, one would also need to show that bodily changes are necessary for emotions. James and Lange defend the necessity claim by appeals to introspective intuition. James asks, "What kind of an emotion of fear would be left, if the feelings neither of quickened heartbeats nor of shallow breathing, neither of trembling lips nor of weakened limbs, neither of goose-flesh nor of visceral stirrings, were present, it is quite impossible to think" (1884, 193ff.). And Lange, without any awareness of James's work, echoes, "If from one terrified the accompanying bodily symptoms are removed, the pulse permitted to beat quietly, the glance to become firm, the color natural, the movements rapid and secure, the speech strong, the thoughts clear,—what is there left of his terror?" (1885, 675). The authors want us to mentally subtract all the bodily symptoms from an imaged emotional state and see what remains. James says we will discover that "we have nothing left behind, no 'mind-stuff' out of which the emotion can be constituted."

James finds further evidence for the necessity of bodily changes in contemporary observations of individuals with no awareness of their bodily states. He considers anecdotal reports from doctors who examine patients with disease-induced bodily anesthesia, and even a French study of individuals who had bodily anesthesia induced temporarily through hypnotic suggestion. In these latter cases, and in some of the disease cases, individuals report a profound reduction in emotion. Some other patients were reported to show preservation of emotional behavior, but there was no reported evidence of preserved emotional experience. In a more systematic study, Hohmann (1966) investigated the emotional states in a population of twenty-five people with spinal cord injuries. He found reductions throughout the group, and those reductions became more acute with injuries higher in the cord. With less bodily feedback, Hohmann concludes, there is less emotion. He did find that these patients tended to experience an increase in "sentimentality," characterized by crying and feeling choked up. This is unsurprising on the James-Lange view, however, because those bodily states involve changes that are above the injury, and hence perceivable to the patient.

In sum, the hypothesis that emotions are perceptions of bodily changes is consistent with a range of observations. Some of these suggest bodily changes are sufficient for emotions, some suggest that bodily changes are necessary, and some merely establish a correlation. Collectively, these observations tend to favor taking a James-Lange approach, all else being equal. But all else may not be equal. I turn now to six objections that attempt to show that bodily perceptions are either unnecessary or insufficient for emotional response.

#### ARGUMENTS AGAINST THE JAMES-LANGE THEORY

##### 1. Diminished Bodily Perception Does Not Always Diminish Emotions

As we have seen, some studies of individuals with spinal cord injuries support the James-Lange theory. When Hohmann (1966) asked twenty-five spinal patients to compare their present emotions to their past emotions, they reported a significant decrease. The decrease was greatest for those whose injuries were highest in the cord. Hohmann does report, however, that some emotions remain. Subjects in his study say that they became more "sentimental" after their injuries, where being sentimental is characterized in terms of crying and feeling choked up. Other investigators tried to replicate Hohmann's study more carefully (e.g., by adding control groups), and they found that some spinal patients continue to experience a full range of emo-

tions (Chwalisz, Diener, and Gallagher 1988). These results seem to contradict the James-Lange theory.

Damasio (1999, 289ff.) offers a multipronged response to this objection. First, he notes that most spinal cord injuries are incomplete. Second, much information about the body can travel through the blood stream, the vagus nerve, and cranial nerves that remain intact after the spinal cord is sectioned. Third, spinal patients can feel changes above the spinal cord, including changes in the throat, face, and central nervous system. Is it any wonder that Hohmann's patients could feel choked up and cry? Fourth, in all studies there is some attenuation of affect, and the attenuation increases with the height of the injury.

Damasio also suggests that emotional responses may be able to bypass the body by means of an "as-if loop." If an emotion is a perception of a bodily change, then the very brain state that underlies that perception must be able to arise in the absence of a bodily change, acting as if the body had changed. James anticipated this suggestion in a footnote about pathological morbid dread. He says that morbid dread can occur without bodily change, but he does not see this as a threat to his theory: "It is of course possible that the cortical centres normally percipient of dread as a complex of cardiac and other organic sensations due to real bodily change, should become *primarily* excited in brain-disease, and give rise to an hallucination of the changes being there,—an hallucination of dread, consequently, co-existent with a comparatively calm pulse, &c." (1885, n. 4). Damasio places much more emphasis on this shortcut than James, and he speculates that emotions may bypass the body even in nonpathological cases.

The as-if loop can also help answer an objection posed by Walter Cannon (1927) in an influential critique of James and Lange. He complained that emotions cannot be perceptions of visceral changes because the viscera react too slowly. One can respond to Cannon by pointing out that some bodily changes, including those in striate muscles, are quite fast. But this reply concedes that visceral changes are significantly less important for emotional response than James had presumed. The as-if loop helps us avoid this concession. If emotions often work by anticipatory perception—by simulating a perceived bodily state before that state has occurred—the speed of visceral response would not constrain the speed of emotional response.

Critics may find the whole idea of an as-if loop a bit desperate. Isn't it ad hoc to assume that spinal patients simply hallucinate bodily changes when they are in an emotional state? Doesn't this make the theory impossible to refute? I don't think so. If emotions are evolved from reflexive bodily response, as I suggest above, the brain may have adapted a way of anticipating bodily movements before they happened, in order to help us make a faster response. In addition, there is evidence that some neurons involved in body control also serve in body perception. Most notably, mirror neurons (Rizzolatti, Fadiga, Gallese, and Fogassi 1996) in the motor cortex respond

when we move our hands and when we see hands move. Perhaps some of the cells that underlie perceptions of bodily changes also orchestrate such changes. If so, we would be able to experience changes in our bodies through the very act of instructing our bodies to change. When this occurs, the experience of change would occur prior to and independent of actual change in the body. The hypothesis could be tested by looking to see whether the neurons involved in body change work like mirror neurons. We could also use neuroimaging to test the more general hypothesis that bodily changes are being hallucinated by spinal cord patients. We need only ask whether the brain centers that are ordinarily involved in body perception are active when they report emotions.

## 2. Some Emotions Do Not Involve Bodily Change

A common criticism of the James-Lange theory is that it doesn't generalize. Critics like to point out that some emotions are not associated with bodily changes at all (Harré 1986). Is there a bodily correlate of guilt? What about loneliness? In addition, there are some emotions that have bodily manifestations under certain circumstances and not others. Consider long-standing emotions (Solomon 1976). One can be in love for a long time, even though one's body isn't in an enduring state of perturbation. The fact that one's body returns to a baseline level of arousal during the hours, days, or even years that one is in love does not entail that one's love periodically wanes during that period. Even James seemed to admit that some emotions fall outside the explanatory purview of his theory. He says, "I should say first of all that the only emotions I propose expressly to consider here are those that have a distinct bodily expression. That there are feelings of pleasure and displeasure, of interest and excitement, bound up with mental operations, but having no obvious bodily expression for their consequence, would, I suppose, be held true by most readers" (1884, 189). As examples, James cites certain moral, aesthetic, and intellectual feelings. He refers to these as cerebral emotions, and contrasts them with the "standard" or "coarser" cases, which are always embodied (1890).

In response, one could concede that emotions form a mongrel category: some fit the somatic model and others do not. But this concession would leave us with a puzzle. How do we recognize disembodied emotions as belonging to the same category as our most visceral sorrows and fears? Why does a single word, *emotion*, lord over such a motley? I suspect there is far greater unity in the emotion category than often appreciated. I think the somatic approach can subsume anything that deserves to be called an emotion.

Consider the alleged counterexamples just adduced. Can we be certain that there is no bodily correlate of guilt? Perhaps it overlaps with the down-

trodden bodily state of grief. We talk of guilt pangs and agonies. Perhaps guilt has a face, with downcast eyes and lowered chin. There may even be a blush of guilt, borrowed from the more primitive emotion of embarrassment. Loneliness may not even be an emotion, but, in any case, it is surely embodied. Like grief, again, loneliness seems to be marked by consuming enervation. James's cerebral emotions may also have bodily concomitants. Moral passions are widely believed to prod us into action, and aesthetic response can send tingles down our spines. Intellectual emotions can overlap with surprise or delight and almost certainly have a somatic mark. Consider the self-satisfied grin of a mathematician who has just discovered an elegant proof.

This is not to say that we never ascribe emotions in the absence of bodily perturbations. James mentions the case of art critics, who have mastered the skill of aesthetic judgment so well that they can praise and demolish artworks without visceral response. Retracting his initial concession, James (1885, 201ff.) implies that cerebral emotions either have subtle bodily concomitants, or they do not deserve to be called emotions at all. Mocking the well-trained critic, he writes, "A sentimental layman would feel, and ought to feel, horrified, on being admitted into such a critic's mind, to see how cold, how thin, how void of human significance, are the motives for favour or disfavour that there prevail" (1894, 202). In other words, some alleged cases of disembodied emotions can be dismissed as vague imitations. If a critic claimed to find delight in an artwork but showed absolutely no somatic response, we might justifiably question her sincerity.

Something similar can be said about long-standing emotions. Imagine someone who claimed to be in love but *never* showed signs of somatic response. I think we would regard this person as disingenuous or confused. Long-standing emotions deserve to be called emotions only because they *dispose* us to enter into patterned bodily responses. We do not say that these emotions disappear when they are unfelt, because the disposition is there all the time. Compare the claim that Sally is sickened by seafood. That does not imply that Sally is in a perpetual state of sickness—only that seafood is disposed to make her sick. The fact that her standing state of being sickened by seafood does not involve a constant perturbation of the body certainly does not imply that being sickened is not a somatic state. Likewise, the quiet phases of our standing passions do not cancel out their carnal nature. I would defy the critic of James and Lange to identify a single emotion that lacks a bodily mark, at least dispositionally.

### 3. Bodily Changes Require Interpretation

Suppose we admit that all emotions are embodied. We can still ask whether perceiving a bodily change is sufficient for being in a particular emotional state. One challenge to the sufficiency claim comes from a famous study by

Schachter and Singer (1962). They argue that bodily changes qualify as emotions only when coupled with judgments that attribute those changes to emotionally relevant objects or events. To show this, they injected subjects with adrenalin, which causes autonomic arousal. All subjects were told that they had been given a drug that was designed to improve vision. While waiting for a vision test, some subjects were seated in a room with a stooge who engaged in silly behavior, such as playing with hula hoops and making paper airplanes. Other subjects were given an offensive questionnaire to fill out and seated with a stooge who feigned being irate about the questions contained therein. All subjects were secretly observed as they interacted with the stooges, and all were given a questionnaire about their physical and psychological states after waiting in the room. Schachter and Singer observed that subjects with the silly stooge behaved as if they were happy, and subjects with the irate stooge behaved as if they were angry. There were also control subjects who had been given a placebo and subjects who were forewarned about the effects of the drug. Both showed less response to the stooges. The experimenters conclude that bodily change is indeed necessary for emotion, but cognitive interpretation is needed to determine what emotion a bodily change amounts to.

These results may look embarrassing for James and Lange, but closer inspection suggests that they are actually harmless. First, strictly speaking, James and Lange do not need to insist that every emotion has distinctive physiology. They can say that the identity of an emotion depends in part on context. Gordon (1987) draws the useful analogy between sunburns and windburns; these are physiologically indistinguishable, but they are different ailments, in virtue of having different causes.

Second, the experiment does not actually establish that the subjects in the two conditions have different emotional states. While their behavior is different, subjects in both groups report being relatively happy when they filled out the questionnaire about their current emotional state in the final part of the experiment. Schachter and Singer dismiss this, saying the subjects may have been trying not to offend the experimenters, but the same logic could be used to explain their behavior while interacting with the stooges. Perhaps they were playing along with the stooges just to be sociable. On the face of it, this would not explain why the control subjects were less responsive to the stooges, but there is an explanation for this as well. If the adrenalin made the subjects happy, they may have become more sociable, and thus more likely to mimic the stooge. Subjects without the drug were simply less sociable. Subjects who were informed about the effects of the drug may have recognized that their expected states of arousal felt pretty good. They would have concluded that their happiness was caused by the drug, and knowing that it wasn't caused by being in the presence of another person, they may have been reluctant to act in the sociable way that happiness otherwise promotes.

Third, Schachter and Singer did not actually measure the physiological states of subjects at the end of the experiment. Earlier, I mentioned evidence that some emotions (including anger and happiness) do have distinctive body states, but these overlap. Perhaps the generic state of arousal caused by the drug transformed into emotion specific states over the course of the experiment as a function of context. Subjects without the drug didn't develop strong emotions, because they weren't given a head start. Subjects who were informed about the effects of the drug may or may not have had an altered emotional state, but either way they would have blamed the drug for a good portion of their feelings and resisted acting out in characteristically emotional ways.

The experimental results are inconclusive. There have been other alleged replications of Schachter and Singer's results along with attempted replications that failed. In a major review, Reisenzein, Meyer, and Schützwohl (1983) conclude that Schachter and Singer's conclusions are not supported by the data.

#### 4. Some Perceived Bodily Changes Are Not Emotions

A related objection to James and Lange is that some bodily changes are not experienced as emotions at all. The arousal caused by exercise, the shivers caused by cold, and the sluggishness of fatigue all come to mind. Why are these states not regarded as emotions if emotions are just perceptions of the body?

This simple question has a simple answer. James and Lange are not committed to the view that every bodily change corresponds to an emotional state. Only some are. James (1894) says that emotional states tend to involve a number of bodily changes and that these almost always include changes in visceral organs. Shivers caused by the cold are simply too local to qualify as emotional states.

This reply is OK for shivers, but it is less satisfying for exercise arousal, fatigue, starvation, and other more global bodily states. Mere locality cannot distinguish emotional changes in the body from nonemotional changes. James and Lange could just list all the patterned changes that qualify as emotions and leave it at that, but this would miss the force of the objection. If the essence of being an emotion is being a perception of a (relatively global) bodily change, then fatigue and starvation should qualify. This suggests that emotions must have some other essence. The James-Lange theory leaves the most fundamental question unanswered: What is it to be an emotion?

A response to this challenge requires the contemporary defender of James and Lange to formulate the account a bit more precisely. We can distinguish two questions. On the one hand, we can ask what kind of inner

mental states are the vehicles of our emotions. Here, a list a particular body-pattern perceptions may suffice. On the other hand, we can ask what makes those vehicles qualify as emotions at all. Here, it is necessary to say something about functional roles. The bodily states whose perceptions are experienced as emotions characteristically arise under certain kinds of circumstances. They arise when an organism faces what Lazarus (1991) has called core relational themes: organism/environment relations that bear on well-being. Starvation and fatigue certainly bear on well-being, but they are not relationships between an organism and the environment; they are states of the organism. Core relational themes include dangers, losses, threats, achievements, status demotions, and transgressions. In each case, there is an object, situation, or event that bears some relation to the organism. As a first stab, we can invite James and Lange to say that emotions are perceptions of those bodily states that are characteristically caused when an organism enters a relation that falls under a core relational theme.

#### 5. There Are Not Enough Bodily Changes to Go Around

Earlier, I reported that there are emotion-specific bodily states. Levensen, Ekman, and Friesen (1991) identified patterned autonomic changes associated with the six basic emotions investigated by Ekman and Friesen (1971) in their pioneering work on cross-cultural facial expressions (anger, disgust, fear, joy, sadness, and surprise). But what about the many other emotions that we experience? Are there distinctive bodily profiles for guilt, shame, jealousy, love, indignation, amusement, resentment, nostalgia, *schadenfreude*, and existential dread? Research on this question has been less extensive than one would like, but it is difficult to imagine positive results in every instance. Will indignation really have a different bodily expression than anger? Will *schadenfreude* stand out from joy? Will dread have bodily changes not found in fear or grief? If emotions were nothing but perceptions of the body, and there are only a few bodily patterns associated with emotions, then there would be only a few emotions. That is an unfortunate result because there seem to be many emotions. Anger and indignation are assuredly distinct.

The objection has two interconnected replies. First, while anger and indignation are distinct, they are also closely related. So closely, in fact, that it would be appropriate to call the latter a species of the former. Indignation is something like anger at an injustice. Similarly, *schadenfreude* is joy about someone's misery, and jealousy is a combination of anger, sadness, fear, and disgust brought on by a perceived infidelity. If the bodily states of emotions are shared, it is because many emotions belong to common families (cf. Ekman 1999).

This reply is incomplete. It explains the overlap, but it doesn't explain how we distinguish different members of the same emotion family. That's where the second reply comes in. Emotions can be distinguished by their eliciting conditions. This goes back to the point made in the previous section. We can think about perceptions of bodily states as the vehicles of emotions, and we can individuate emotions by their vehicles. We can also individuate emotions by their eliciting conditions. Anger and indignation feel alike, from a somatic perspective, but indignation is caused by injustice, and many instances of anger are not. James and Lange are not sufficiently clear on this point, but it is consistent with their theoretical approach.

## 6. The James-Lange Theory Cannot Accommodate the Intentionality of Emotions

Within philosophy, the most persistent objection the James-Lange theory is that it cannot explain the fact that emotions have intentional content. Indeed, emotions are intentional in two senses. They have formal and particular objects. All fears concern dangers (the formal object), and each particular episode of fear concerns a particular danger, such as an assailant, a great height, a loud noise, a dental visit, an upcoming exam, and so on (particular objects). Intentionality renders emotions amenable to rational assessment. They can be right or wrong, appropriate or inappropriate, warranted or unwarranted, rational or irrational (see, e.g., Pitcher 1965; Solomon 1976). If emotions were merely perceptions of the body, they would qualify as intentional in only an uninteresting sense; they would represent the body as being in such-and-such a state. This, according to the objection, is not enough.

On a standard cognitive appraisal theory, emotions contain evaluative judgments that explicitly characterize their formal objects. Fear might be the judgment that there is a danger (see Lazarus 1991). These formal objects are what Lazarus calls core relations themes—dangers for fear, losses for sadness, insults for anger, and so on. Particular objects elicit fear because they can be appraised as exhibiting the core theme expressed in the fear appraisal. A great height is appraised as dangerous. The total mental state, say fear of heights, contains a representation of the particular object (a high place) as well as a representation of the property that makes it fearful (the danger of falling). Emotions place representations of particular objects under concepts that represent core themes. Since heights are in fact dangerous, fear is warranted. If fear persists after reliable precautions have been taken, the fear is no longer warranted.

On the face of it, the James-Lange theory is totally incapable of accommodating the intentionality or rationality of emotions. Even Damasio gives in on this point. He says the James-Lange theory places inadequate em-

phasis on the role that evaluation plays in the induction of emotions (1999, 130). To make up for this shortcoming, he recommends that we identify emotions with sensations of bodily changes *coupled with* a mental evaluative process (139). This process can involve innate perceptual triggers in the case of the primitive (or “primary”) emotions exhibited by animals and children, but it will involve more complex cognitive processes mediated by the frontal cortex in the more advanced (or “secondary”) emotions of human adults.

I think Damasio concedes too much. In responding to the last objection, I suggested that James and Lange can make some headway by emphasizing the conditions under which emotions are elicited. Emotions, I said, can be individuated by their reliable elicitors. This suggestion can be taken a step further. According to prevailing theories of mental representation, a mental state gets its intentional content in virtue of being reliably caused (or having the function of being reliably caused) by something (Dretske 1981, 1988; Fodor 1990). Let’s assume that a theory of this kind, whatever the details, is correct. There is some causal relation that confers content. If emotions are perceptions of bodily states, they are caused by changes in the body. But if those changes in the body are reliably caused by the instantiation of core relational themes, then our perceptions of the body may also represent those themes. In other words, leading theories of how mental representations *entail* that emotions represent core themes, if they are reliably caused by those themes in the right sort of way.

Consider fear. It seems quite likely that we are wired to undergo a perceived (or imagined) bodily change under a variety of threatening conditions. A similar bodily pattern is triggered when the auditory system detects a loud sudden noise, or when the visual system detects a looming object, or when we proprioceptively detect a sudden loss of support. The perception of that bodily state represents danger, because it is under the reliable causal control of dangerousness. Danger is the property in virtue of which these highly desperate eliciting conditions have come to perturb our bodies. If loud noises and looming objects were not dangerous, they would not have their characteristic effects. We can think of all of these body-change elicitors as belonging to a mental file—an elicitation file. That file may start out with a handful of triggers and expand over the life span. As we learn of new dangers, we may add new entries to the elicitation file. Elicitation files can even come to include evaluative judgments of the kind emphasized by defenders of appraisal theories. Each addition to an elicitation file will be sufficient for triggering the relevant bodily response, though getting admitted to the elicitation file in the first place will depend on similarity to or association with triggers that have already been attained. Consequently, all the representations that trigger the bodily response will do so in virtue of being recognized as dangerous, either explicitly or implicitly by similarity to previously established elicitors.

This proposal answers half of the intentionality objection. If bodily perceptions represent core themes by reliable causation (or something along these lines), then they have formal objects. They also have conditions of correctness. If fear happens to be caused by something benign, it would count as inappropriate because fear bears the meaning-conferring causal relation to dangerous things. New emotions can be generated by modifying elicitation files or establishing new elicitation files that are offshoots of those that have already been established. The new files may expand or contract the conditions under which an emotion is elicited. For example, an anger elicitation file can spawn an “indignation” offshoot that contains representations of different kinds of injustice. Culture can exert considerable influence on how elicitation files are modified and created.

This explains how emotions qua bodily perceptions can have formal objects, but what about particular objects? How does a thought about heights latch onto a perception of the body? A full answer to this question would have to include details about how mental states are bound in thought. The idea would be that a representation of heights gets coactivated with a somatic perception and linked to it in such a way that the former causes the latter to occur, and the latter wanes when the former becomes inactive. There is no reason to provide the details here, because any theory of the emotions will have to explain how dependencies arise between representations of particular objects and somatic states. Everyone agrees that bodily responses often occur *in virtue of* some particular perception or thought. The defender of the James-Lange theory can say that this dependency, however it gets fleshed out, determines the particular content of an emotional state.

Although it is only a sketch, I hope this response makes it clear that one could in principle capture the intentionality of emotions without abandoning the idea that emotions are perceptions of bodily states. Strictly, the intentionality objection does not require one to abandon the core idea behind the somatic theory. This is a very significant finding, because the majority view is that the somatic theory is utterly unworkable. But it would be a hollow victory if there were reasons to favor a theory that identified emotions with compounds of bodily perceptions *and* evaluative states. Are there such reasons?

#### EMBODIED APPRAISALS

I have been suggesting that perceptions of the body could represent core relational themes by being reliably caused by such themes. This proposal offers a way out for James and Lange, but it faces a pressing objection. Above, I introduced the idea of elicitation files, and I said that these determine what the content-conferring causes of our emotions. I also said that elicitation files may include evaluative judgments. But why not identify the

contents of an elicitation file with an emotion? Shouldn't we follow Damasio in saying that emotions are bodily perceptions *plus* evaluations?

I'll mention three considerations that can help us resist this move. First, the fact that elicitation files help establish the intentional content of emotions does not entail that they should be regarded as constituent parts. The nociceptive fibers that link bodily injuries to pains mediate the relation between pains and their contents, but there is little temptation to say that pains are such fibers. Second, the items in an elicitation file do not activate with the right time course to qualify as components of an emotion. An emotion can be triggered by a passing thought or fleeting perception and then linger. Consider a bout of fear caused when one mistakes a shadow for a bug moving across the floor. Third, the contents of an elicitation file will be very heterogeneous. In contrast to appraisal theories, I submit that there is no privileged representation mediating the link between core relational themes and bodily changes. The items in an elicitation file range from very abstract cognitive appraisals to very concrete perceptual representations of specific objects. If emotions are elicited by different representations on different occasions, there is no reason to think that any one of those representations qualify as constituent parts of the emotion. If elicitation-file contents were constituent parts, emotions would change from occasion to occasion. That consequence can be avoided.

These are not knockdown arguments, but they show that we are under no obligation to regard the items in an elicitation file as constituent parts of the emotions they elicit. Far from it. There might be motivation to do so if these items were the only things that could serve as bearers of the right intentional contents. I have been arguing that this is not the case: emotions can represent core relational themes even if they are perceptions of bodily states. In fact, they must represent such themes if leading semantic theories are correct. I call this amendment to the James-Lange theory the "embodied appraisal theory." Emotions are embodied, just as James and Lange proposed. They are perceptions of changes in our somatic condition. But, ironically, they are also appraisals. Let us define an appraisal not as an evaluative judgment, but as any representation of an organism-environment relation that bears on well-being. Evaluative judgments can serve as appraisals, but they are not alone. If a nonjudgmental state represents an organism-environment relation that bears on well-being, it too will count as an appraisal on this definition. My suggestion is that certain bodily perceptions have exactly this property. They represent roughly the same thing that explicit evaluative judgments represent, but they do it by figuring into the right causal relations, not by deploying concepts or providing descriptions. Our perceptions of the body tell us about our organs and limbs, but they also carry information about how we are faring.

Is this the view that James and Lange had in mind? Probably not. Their contributions predate the relevant developments in philosophical thinking

about representation. But the embodied appraisal theory preserves their core insight and endows emotions with semantic properties that can be used to deflect the arrows of dissent. In developing a theory of emotion, we should not feel compelled to supplement embodied states with meaningful thoughts; we should instead put meaning into our bodies and let perceptions of the heart reveal our situation in the world.

## II

### Emotion, Appraisal, and Cognition

*This page intentionally left blank*

## Emotions

*What I Know, What I'd Like to Think I Know, and  
What I'd Like to Think*

RONALD DE SOUSA

The editor's invitation to contribute to this volume appeared to license telling more than I know. Accordingly this essay will move quickly from an all too brief survey of what I know to raise some of the increasingly speculative questions that currently preoccupy me.

### 1. WHAT I KNOW

On second thought, there's nothing I'm that sure of.

### 2. WHAT I'D LIKE TO THINK I KNOW

#### Reconstructing Cognitivism

There has been much made in recent decades of the idea that emotions are "cognitive." The term is used in a confusing diversity of senses. Sometimes by "cognition" one means merely to insist that emotions are not "merely subjective" phenomena. But that is hardly helpful, since there are by my count at least a dozen different things one can mean by "subjective" (de Sousa 2002a). A more contentful thesis is that emotions are genuine representations not just of the inside world of the body but also, through that, of the external world of value. As representations, they have a mind-to-world direction of fit. In this they are like beliefs rather than desires. At the minimalist end, Martha Nussbaum (whose own view, to be sure, is more robustly "cognitivist") has remarked that a theory is "already in [a] sense a cognitive theory [if] the transmission of information within the animal is central to it" (Nussbaum 2001, 114). But if that is true, it is difficult to see

what ground is left for noncognitivism to occupy. By contrast, Solomon's view that emotions are *judgments* (Solomon 1973)—call this *straightforward cognitivism*—remains refreshingly extreme. The burden of my own proposal will be to suggest that insofar as a “cognitive” view of the emotions can be sustained, it is better construed on the model of perception than on the model of knowledge or judgment.

Experimental psychology and commonsense observation have accumulated an impressive range of facts that refute straightforward cognitivism (DeLancey 2002; Stocker and Hegerman 1996). Fear of flying is a classic example of a groundless emotion. It is experienced by some who are well aware that no method of transport is safer than flying. So while fear has something to do with danger, believing that X is dangerous is neither necessary nor sufficient for experiencing fear of X.

More controversial, because less conclusively shown to be irrational, is the fear of death. According to a notorious argument of Epicurus, no one should experience any fear of her own death *as such* (as opposed to its many unpleasant concomitants—the process of dying, the loss to one's loved ones entailed by one's own death, the unknown fate which those who don't really believe death is *death* contemplate in its stead, and so forth). For death is experienced neither while we are alive—since we're not dead yet—nor when we're dead—since we're no longer there to experience anything. The argument is neat, but many share Philip Larkin's impatience with

specious stuff that says *No rational being*  
Can fear a thing it will not feel, not seeing  
That this is what we fear.

Epicurus's argument assumes that to fear X in itself must be to fear *the experience of X*. And that is question begging, though perhaps not wholly gratuitous.

Yet many who are well aware of these counterexamples continue to call themselves cognitivists. And indeed the dispute over Epicurus also establishes that it makes sense to criticize an emotion for *irrationality*, if only by impugning its appropriateness to its object. The pertinence of such a charge may suffice to justify the label of “cognitivism.” More important, not all cognition is belief: *perceptions* are also a form of cognition, and so may provide a better model on the basis of which to think of emotions as cognitive. Compelling visual illusions that persist in the face of the knowledge that they are illusory constitute an immediate analogy to groundless emotions. (See Hoffman 1998 for numerous examples.) I am sure enough that something of this sort is right; how the details of the analogy are to be worked out, on the other hand, belongs under my third, most tentative heading.

## What Is Universal Overflows the Species

Recently my five-year-old's teacher described the dynamics in the life of the day care's denizens. The two dominant children never fought or interfered with one another: each had his or her own sphere of influence. Among the others, most didn't have the strength actively to resist their dominant patron's demands, but some knew better than others how to evade it. Some were demanding yet quick to sympathy; some, on the contrary, seemed rather to relish others' distress. Some jealously demanded exclusive attention; others continually sought protection from the more powerful. In short, all the commonplaces of human social interaction were already to be seen there. The picture she was sketching was a perfect match for the world scene of politics as well as any lesser realm of human life (see Moldoveanu and Nohria 2002). Whether it is displayed in the daily newspaper or dissected by Hobbes or Machiavelli, the dynamics of social life seem to turn on emotional dispositions equally powerful at all stages of life. These are merely more transparent in small children unable to articulate the confabulations with which adults dress them up. And as Frans de Waal (1996) has richly illustrated, while social life among other primates may not be quite as varied in its details as that of humans, the range of patterns present in other primate species are very much the same: affiliation and exclusion, rank, dominance and submission, sympathy and antagonism, reciprocity and revenge, compassion, jealousy, envy, anger, guilt, and shame: none will be news to the primatologist any more than to my child's kindergarten teacher.

This simple observation raises two questions. The first might well arise from the resistance likely to meet any reminder of our kinship with our mammalian cousins. (For it is part of our primitive animal vanity to pretend we are not animals.) *If we are so much alike, how come we are obviously so different?* And does this leave room for any differences that can reasonably be ascribed *uniquely to our species?*

My answer, in a nutshell, is that human emotions, thanks to a gift of language that brings with it a particularly strong form of intentionality, are differentiated from the emotions of other animals as well as from one another by the acquisition of a *narrative form*. This adds to our experience of reality a whole new dimension—or rather a whole new set of dimensions, which both transform and attempt to usurp the function of our primitive emotions.

That answer, however, must in all fairness be relegated to the category of things I don't quite know but would like to think.

### Varieties of Intentionality

Emotions face both in and out: they reflect facts about the subject but refer also to something outside, to which they typically are responses. In this

respect they offer both an analogy and a contrast with sensory perception. Perceptions, as recently pointed out by Alva Noë, are “transparent” in the sense that when you attempt to depict your visual field you just end up drawing a picture of the room you are in (1999, 124–26). By comparison, emotions are relatively *opaque*: often the effect of passion is precisely the reverse: when the angry man, or the joyful bride, or the jealous husband attempt to describe the world, they succeed only in describing their own state of mind, or perhaps even just the hormonal (or as it used to be said the *humoral*) balance in their body. Descartes may have implied something like this in his odd phrase about “passions of the soul” being affections “excited by the animal spirits” yet “referred to the soul” (1989, part 1, arts. 25, 27). Although his phrase is not altogether clear, it seems to reflect something of the complex ways in which emotion relates to many “objects”: often to some specific target in the world (the “object of my affections”), but also to the state of my body, to my phenomenological consciousness, and to aspects and features of the target most naturally expressed in terms of adjectives or adverbs. Thus if I am *moved* by your performance, this has something to do with the tears welling in my own eyes; if I am *angry at* you, this may be *because* of some things about my own past, no less than because of something you have done, or *how* you have done it, or just what you are like. All this may be signaled, as Descartes ([1649] 1989), James (1884), and Damasio (1994) have pinned down with increasing accuracy, by some “perception” of the state of my own body.

## Emotions and Consciousness

“Perception” needs to be put in scare quotes because, as all three writers realized, it need never reach awareness. This point is worth stressing. Freud (1915, 78) claimed that strictly speaking the notion of unconscious affect is a contradiction in terms. But much recent work on consciousness has established that any simple dichotomy between what is conscious and what is not (even when supplemented with slots for the “subconscious” or “pre-conscious”) is hopelessly simplistic. The “stream of consciousness” is something of a “grand illusion” (Noë 2002; see Churchland 1986). Insofar as emotions are construed on the model of perception, therefore, we should not expect them to be simply conscious or unconscious. A fortiori, since the body’s role in emotions is more complex, Freud’s remark seems even less likely to be true. In many ways, some of which I return to in a moment, an emotion can genuinely affect not just behavior but even our whole orientation to the world and the events of our lives, without the subject having any particular insight into either the identity of the emotion or the nature of its influence.

## Emotions and the Body

It is often stressed that emotions are *bodily* phenomena. A contrast is presumably intended with “purely cognitive” phenomena, but in some sense perceptions and beliefs are bodily too, so the import of the claim is not clear. The insistence on the bodily character of emotion is presumably intended to underline two differences between emotion and other “states of mind.” One is that in many cases specific emotions are associated with certain relatively gross physiological changes: blushing, accelerated heart-beat, skin conductance changes—in short, the kind of measures that impressed James and that polygraphs are intended to record. But some of these involve highly specific and organized brain circuits such as those controlling fear, rage, euphoria, or disgust (Calder, Lawrence, and Young 2001). Others appear to relate to emotional states that may altogether lie below the threshold of awareness. These last, revealed by such pathological states as Capgras syndrome—about which more below—are particularly interesting because they testify to the existence of emotional states that are not experienced as passions, indeed that may not be consciously *experienced* at all.

## The Power of Teleology in Emotions

A further supposedly established fact about the emotions is that they appear to *motivate*. I say *appear* to motivate, because this apparently obvious fact is infected by the obscurity of the notion of motivation. What is now increasingly clear is that despite their reputation for motivating bad behavior, emotions are essentially implicated in our capacity to live a coherent and reasonably well-regulated life: unless you *care*, your life will be a mess; and whether or not you care is surprisingly neatly attested by your galvanic skin response (de Sousa 1987; Damasio 1994).

As amply illustrated in the works just cited, emotions are hugely variable in the ways in which they determine the course of our behavior. But if, as I believe, complex emotions are learned in the context of “paradigm scenarios” (de Sousa 1990) from which they derive a dramatic or narrative structure, it is misleading to think of them as motivating in the same sense as desires motivate. Desire and belief are concepts tailor-made to enter into a certain neat theory of action, in which they constitute the two ready parameters of motivation. But emotions are more complicated and play a more subversive and ambiguous role. Most complex emotions prescribe no specific behavior. They also affect us at many levels, of which the body’s readiness to undertake certain sorts of behavior rather than others is only the simplest. More subtly, they influence decisions not so much by “motivating” as by orienting attention toward this or that among the plethora of considerations that might be thought relevant at any particular juncture.

At all these levels beyond the first (and perhaps even there), the emotions play a determining role by virtue of the narrative structures of paradigm scenarios. These are elaborated mostly in the light of early experience, but also perhaps by exposure to art and literature, and they serve to define and differentiate each person's idiosyncratic repertoire of emotions. The resulting complexity means that the teleological structure of emotions often looks more like *fatalism* than like ordinary *determinism*. Rather as the Delphic oracle's predictions were realized largely because of the efforts made to evade them, the emotions often work like agents able to get around almost any attempt to foil their ends. But while the facts here are impressive—and widely described in the literature under the heading of self-deception, *akrasia*, and other pathologies of thought, inference, and action—the mechanisms of this elaborate teleology remain obscure.

### The Axiological Hypothesis

The role of emotions in ethics has long been debated, and the history of philosophy has oscillated between two extreme views. One holds that emotions constitute a disruptive factor that stands between us and any possible redemption of our bestial nature. At the other extreme is the idea that emotions lie at the core of rationality and ethics. This latter view rests on three considerations.

First, by defining what we *care* about, the emotions set the ultimate ends of all deliberation.

Second, there is a whole range of emotions that are themselves intrinsically epistemic (Hookway 1998). Most notable among these are the feelings of certainty or doubt, as well as a more general feeling of “rightness” that validates our inferences and our conclusions (Mangan 2001). Such feelings, like the feeling of spiritual conviction frequently associated with temporal-lobe seizures (Ramachandran and Blakeslee 1998, 177ff.) are manifestly emotional experiences that come, as it were, bearing the marks of epistemic import.

Third, by controlling the salience of different considerations, the emotions also determine not just ends but also much about the means by which we choose to attain our ends.

On this last point the debate often crystallizes around the issue of *akrasia* or weakness of will. On one view, which I have defended in the past on the basis of an elaboration of the framework laid out in Davidson 1970, *akrasia* consists in a failure to apply to a current decision the most comprehensive set of available reasons. Emotion is uniquely qualified to assume responsibility for these failures, because while it disrupts the path meant to lead from the “best” reasons to action, it is not the sort of brute cause that would remove the event altogether from the sphere of responsible action.

Instead it brings about the failure of the best reasons to be channeled into action by virtue of its control of the agent's allocation of resources.

The work of McIntyre 1990 and Bovens 1999, however, has convinced me that my earlier view was tainted by a common prejudice. Many cases of akrasia can be viewed more benignly. The accusation of weakness of will is most often leveled on the basis of a clash between the operative reason and the vectorial sum of all available reasons as explicitly declared. But why assume that our best selves are to be found on the side of the most explicit of our reasons? An accumulation of philosophical, psychological, and even neurological research can be mustered to support the view that our explicit declarations are greatly contaminated by confabulation, and that much of what passes for rationality is actually rationalization. Thus, McIntyre and Bovens point out, if Huck Finn passes for akratic because he acts against his explicit principles, so much the worse for his explicit principles. Look instead to the deep emotional commitments that actually govern his inability to be ruled by his principles, and you will see that he was being faithful to his best self.

### Emotional Cognition as Exaptation

The contrast between rationality and rationalization presupposes a real difference between cognition and illusion, between objective representation and projection—in short, between truth and illusion. I will therefore shortly need to say something about *emotional truth*. But the role of emotion in cognition may also be an auxiliary one, which we can speculate is the result of an “exaptation.”<sup>1</sup>

The role of emotions in learning, memory, and recognition provides a good example. It is well known that memory best retains (or perhaps retains only) what is emotionally significant. (There may have been something pedagogically sound, if in other ways deplorable, in old-fashioned methods of instruction based on blows, humiliation, and ridicule.) But let me focus on the more specific example of Capgras syndrome (Ramachandran and Blakeslee 1998, 158–73). Subjects affected by Capgras syndrome insist when they see parents or loved ones that they are seeing an impostor. They have no problems of memory, nor do they suffer from prosopagnosia: on the contrary, Ramachandran's patient Arthur conceded that the person before him looked exactly like his mother, and wondered about his mother's motivation in hiring an impersonator. What then could be going on? Ramachandran's suggestion is that a direct link normally exists between the facial recognition mechanism and the areas controlling the appropriate emotional responses (particularly the amygdala). The sight of a parent normally triggers a not necessarily conscious affective response, which is itself subject to an evaluation for familiarity. Think of this on the analogy of an

ID card that carries both a signature and a picture. Both have to match: if one of them fails to match, the other is automatically suspect. Thus the recognition of those close to us rests on two marks: the cognitive or visual representation and the emotional “signature.” If the signature fails, the visual presentation is deemed fraudulent. In Arthur’s case, the affective response is missing. This sets up a discrepancy between the strictly cognitive familiarity check that applies to the face and the missing familiarity check applied to the expected affective response. The Capgras delusion is then no more than a perfectly reasonable inference: the person before me is not my mother, since I get a characteristic emotional marker when my mother appears and I’m not getting it now. On the other hand, she looks exactly like my mother. Therefore she is an impersonator.

The example is of more than merely anecdotal interest. It suggests that the emotions present us with an entire information processing system on its own, a parallel representational system for understanding the world. The questions raised by this possibility, including the question of how emotions might provide us with something we might call *objective representations* of values, belong firmly under the third of my headings.

### 3. WHAT I’D LIKE TO THINK

#### Parallel Systems?

In the light of evidence that the brain’s older systems of control continue to work in parallel—and not infrequently in conflict—with those that have evolved in the cortex, Paul MacLean elaborated a hypothesis consistent with the idea that emotions form not so much a set of functions that *contribute* to cognition and behavior, but actually *a parallel system of control*. This fits in all too well with the evidence cited above from politics, chimpanzees, and day care dynamics. MacLean puts it thus: “We might imagine that when a psychiatrist bids the patient to lie on the couch, he is asking him to stretch out alongside a horse and a crocodile” (1960, 300).

Indeed there is reason to believe that the circuits of the limbic system—chiefly associated with emotion—constitute the *original* system of behavioral control. If so, then we are again faced with the possibility that most of our rational thought is rationalization, and that it works well enough in practice because the robots that we are have been well conditioned to survive regardless of the implausibility of our confabulations. But that, in turn, raises again the question of the nature of the compact between explicit reason and the ends and means of the emotional system. It has become fashionable to claim that there is not *really* any opposition between reason and emotion, but that may be nothing but a comforting myth. On the contrary: there is a deadly opposition between emotion and reason: it’s just that “reason” can’t set its own goals or do anything much about them without the connivance of its adversary.

If anything like this is right, it raises once again the question of the nature of our emotional representations and their claim to be giving us information about the world outside our own minds and bodies. Descartes warned that the representations of the senses might be fulfilling their natural purpose without giving us any information about the world as it is: “The proper purpose of the sensory perceptions given me by nature is simply to inform the mind of what is beneficial or harmful. . . . But I misuse them by treating them as reliable touchstones . . . about the essential nature of the bodies located outside us.” In particular, Descartes continues, “there is no convincing argument for supposing that there is something in the fire which resembles the heat, any more than for supposing that there is something which resembles the pain” ([1641] 1986, 83).

Actually this last claim can be pretty confidently rejected, thanks to the existence of cross-modal confirmation of the reality of heat. What can be seen, heard, touched, measured with various instruments, and observed by means of different senses and by different observers to undergo the same constancies and changes must indeed be objective if anything is (Nozick 2001, 75ff.). The point is especially pertinent to emotions: Are they merely motivational guides to efficient living, or are they actually “cognitive” in the sense that they represent something that can be said to have objective reality? On this criterion, there will indeed be objectivity to its deliverances.

The speculative axiological hypothesis that currently engages me is designed to explore the prospects for emotional objectivity along these lines. *Axiological* is a term that has fallen into disuse. (There is no entry for *axiology* in the recent *Encyclopedia of Ethics* [Becker 2001].) It derives from the Greek word for *worth*, or better, *worthiness*. I revive it to express my conviction that, just as there is a gap between what we *prefer* and what we *decide* to do and another between what we *decide* to do and what we *actually* do (leaving room for two of many varieties of weakness of will), so the gap between what we *know or believe* and what we *prefer* allows for a level of *ascription of value* that is not yet preference, yet is more than mere assessment of factual truth. Poised as it is in that gap, the axiological shares some features with both epistemic and preference assignments. Yet axiology is distinct from both the epistemic and the level of preferences. We might say it is the locus of our capacity to be *moved*, and yet not *moved to any specific sort of behavior*. Like preferences, axiological assessments involve a necessary component of subjective response, without, however, being committed to any unique dimension of valence such as attraction/repulsion. But like facts, the axiological aspires to a high degree of *objectivity*, which for present purposes might be roughly defined—in accordance with the remarks above—as the recognition of a need for and a possibility of *corroboration*.

This last trait suffices, however, to think of the axiological hypothesis as positing a special kind of “cognitivism” about emotions. The hypothesis involves three claims:

First, at least some emotions constitute *perceptions of value*, the content of which may at least sometimes meet something like the multimodal requirement. If this can be made out, we should be able to make sense of a notion of *emotional truth* (de Sousa 2002b).

Second, the values in question are not limited to those valences that determine, along a continuum going from attraction to repulsion, the motivating vectors that drive our behavior. Instead they are multidimensional or, one might say, in contrast to the black-and-white unidimensionality of behavioral motivation, *multicolored*.

Third, most emotions are not intrinsically moral: it follows from the multidimensionality of value that not all value can be moral value. Yet emotions are, in the end, inevitably the ultimate arbiters of all value including ethical value. At the metaethical level, this hegemony of emotion justifies the Wildean adage that ethics is a branch of aesthetics, but it could equally well be formulated by saying that ethics cannot exclude aesthetics from the scope of its vision.

In the rest of this essay, I will say a little more about each of these three points in turn.

## The Concept of Emotional Truth

Standard accounts of truth assume that truth bearers are propositions, however explicated. By adopting a perceptual model of emotion, I discard the idea that emotions are invariably propositional attitudes. (This does not imply that propositions can't be among the many sorts of objects to which emotions are, as I noted, variously related.) This immediately raises two related logical problems: if something can be true, it can be false, and we want to be able to say that its negation is true. We also want to be able to say that a set is *consistent*. How then can the notions of negation and consistency apply to emotions?

Although some named emotions seem to be related as polar opposites (love and hate, hope and despair, admiration and contempt, gratitude and resentment), there is no obvious criterion of contrariness. Two candidates suggest themselves. One is neurological; the other is phenomenological. The phenomenological criterion, however, suffers from extreme subjectivity. While many people would claim it is impossible simultaneously to feel certain pairs of emotions in regard to the very same objects, aspects, and situations, disagreements are hard to settle. "Odi et amo," lamented Horace, and indeed love and hate are at once a paradigm case of opposites and a typical example of emotions that are frequently mingled (Neu 2000; see Greenspan 1978). On the neurological criterion, we might take some inspiration from some elegant work by Marcel Kinsbourne, in which he explored the likely contribution of areas of the brain to different tasks by measuring the degree of interference. He found, for example, that subjects

could easily learn to balance a pole on the tip of a finger on the right or left hand. When asked to speak at the same time, however, the pole tended to fall off if balanced on the right index finger, while it had no disruptive effect on the left. Being asked instead to sing produced the opposite effect (Kinsbourne and Hicks 1978). That sort of interference might yield one measure of consistency: those emotions that inhibit one another, like the activation of the sympathetic and parasympathetic nervous system, would then be rated opposites. Anxiety and sexual excitement may qualify on that score. But the criterion is feeble, as it may relate only to competition for the use of a given brain resource.

Anxiety and sexual excitement may not qualify on the phenomenological criterion. Some people report finding anxiety or fear to enhance sexual excitement. Does the physiology of those emotions work differently for them, or are there simply two incompatible methods of sorting emotions into compatible ones and contraries? However that may be, the phenomenological may, after all, be no more than a reflection of the neurophysiological, as it does in the case of the phenomenological color cone. The structure of the phenomenological color cone conforms perfectly to what is predictable on the basis of the opponent process theory of trichromatic color sensation (Churchland and Churchland 1998, 166–72), which strongly suggests that there is *nothing more* to that phenomenological structure than the reflection of an underlying neurophysiological mechanism. Similarly we commonly assume that the qualities of warmth and coolness attributed respectively to red-yellow and blue-green are so experienced by association. But in fact there is evidence that our perception of colors as warm or cold is a direct reflection of the amount of activation of the low-level opponent channels that give rise to hue, rather than reflecting any learned association between colors and emotions (Hardin 2000, 120).

For all that, the neurological criterion may often be inaccessible, and in many cases perhaps altogether unavailable. In the more subtle emotions, the “formal object” of the emotion—the attribution to its target of whatever features it is that *make* it that emotion rather than another—is the only thing capable of defining a relation of contrariety for that emotion. How many such formal objects can there be, defining how many continua of appropriateness and contrariety? The answer hangs, I believe, on the emancipation of elaborate emotions from the practical purposes for which we can assume that many of them were originally selected (Neese 1990). Let me explain.

I argued in de Sousa 1974 that an important confusion generally mars discussions of consistency of desire, resulting from the false presupposition that a single criterion of consistency is appropriate to both beliefs and desires. That presupposition confuses the *satisfaction* conditions of desire—the condition under which the desired entity or state of affairs exists<sup>2</sup>—with its condition of *success*, the condition under which the desire is *warranted*.

For any two beliefs, compatibility coincides with consistency: satisfaction conditions are just success conditions. But for two desires to be consistent, it is not necessary that their contents be jointly satisfiable, but only that their contents be jointly *desirable*. So while a desire for  $p$  and a desire for  $q$  (where  $q$  implies  $\sim p$ ) are clearly incompatible, it does not follow that they should be regarded as inconsistent. And while this raises difficult questions about how to cash in the claim that two desires are inconsistent, it makes room for inconsistent desires without requiring that consistent desires also be for compatible objects.

Emotions are similar, but messier. The reason is the one just noted: each emotion is linked to its own specific evaluative continuum, and so defines its own proper object, and thereby the dimension along which contrariety might be defined for that emotion. There is no single proper object of all emotions.

Consider a simple example. Allow that the formal object of fear—the norm of appropriateness that is distinctive of fear—is the dangerous. *Fear that  $p$*  is satisfied iff  $p$  is true, but it is *successful* iff  $p$  is *actually dangerous*. Confining ourselves for simplicity to emotions admitting of a propositional object:

$E(p)$  is *satisfied* iff  $p$  is true  
 $E(p)$  is *successful* iff  $p$  actually fits  $E$ 's formal object.

In all cases, the emotion's success is independent of its satisfaction. Fear of monsters is not semantically satisfied, but it may be successful. The converse may be the case in fear of spiders.

Emotional truth, then, refers not to semantic satisfaction, but to success. Fear's assessment of  $p$  as dangerous consists in some sort of *evaluation* of  $p$ . Success is tied to the correctness of that evaluation. The notion of opposition, if any, appropriate to a given emotion is internal to that emotion and follows from the nature of its criterion of success. (Thus trust might be the polar opposite of fear, insofar as *dangerous* and *trustworthy* define a continuum.) And while this provides no general criterion of emotional contrariety, it at least suggests a way in which such a concept might have application, as well as explaining why it is difficult to cash out in practice.

## The Full-Color Content of Emotional Experience

Our experience of ourselves and the world gives rise to feelings, perceptions, desires, and beliefs. These get funneled into a sequence of single decisions, where each requires acting or not acting, yes or no. Desires and preferences, then, form a single black-and-white dimension. Experienced emotions, on the other hand, are so diverse as to constitute no single *kind* of thing at all. Each carries a wealth of specific meanings enriched by a vast class of

contrasts; each involves an evaluative response that may, but need not, exhibit a positive or negative valence or inclination to prefer some behavior. On this, the full-color view, there are no practical limits to the number of distinct emotions that can be experienced, any more than there are limits to the number of thoughts one can have. (See Campbell 1998.)

Consider the emotions aroused by aesthetic experience—watching dance or listening to music, for example. It seems obvious that emotions are involved in some way. But if the point were to express “the great emotions”—those we can list on demand: anger, fear, love, awe, jealousy, sadness, desire—then why go to all the trouble of creating and appreciating the subtleties of great art? It hardly seems likely that works of art in all their diversity should be sustaining our interest for their representation of *emotion*, if there are so few of them. The alternative view is that each different moment in art evokes a *sui generis* emotion. The experience of life affords an indefinite range of possible emotional qualia.

But why call such a view *cognitivist*? If cognition is defined to exclude nonpropositional content, then this view is noncognitivist. If, on the other hand, the term is extended to include any acquired insight into the nature of an external reality, then perception, even where it admits of nonpropositional content, becomes a paradigm case of cognition. Emotions, on the axiological view, can then claim an analogous place in cognition, broadly understood.

### The Multivalence of Narrative and the Hegemony of Emotion

The idea that (some) emotions constitute apprehensions of a *sui generis* realm of values implies that they transcend any of the scripts for behavioral expression some of them might originally have been selected to enact. Like other perceptions, they are not reducible to either beliefs or desires, and I have urged that the world of values they reveal does not reduce to any single measure of positive or negative valence. That feature, I surmise, stems from the linguistic difference of humans: it reflects the possibilities for enlargement (and also the capacity for self-deception) entailed by the ability to elaborate our emotions in terms of a narrative framework.

One consequence has just been noted: although emotions are said to motivate, it might be better to say they channel and construct rationales for behavior, in accordance with the characteristic narratives that define them. Another consequence is that while it is useful to recognize a limited number of named emotions for the purposes of social communication and classification, the actual experience of emotion is better modeled by aesthetic experience, and the range of values they apprehend better assimilated to the indefinitely many dimensions in which aesthetic experience can take us. This means that axiology should be as sharply distinguished from de-

ontology and preference—classes of states concerned with choices and behavior—as from value-free facts.

But now if every emotion sets its own standards of appropriateness or correctness, what can emotions contribute to ethics? It is tempting to appeal to human nature to set a standard of emotional correctness. Given any quale, an evaluative response that falls foul of the norm will lack appropriateness, and on that basis we can call it perverted, abnormal, or *false*.

The problem with human nature, however, is that there probably is no such thing (de Sousa 2000). Rather than taking human nature as a standard against which to judge emotion, I propose to do precisely the reverse: take emotions, rightly understood, as the ultimate arbiters of ethics.

Consider the classic thought experiment of Mencius: you see a child about to fall into a well, and your apprehension of the situation immediately moves you, and you want to save the child. In this instance, what is apprehended is the *need to intervene*. Or better, it is the nature of the total situation, in which the need to intervene roughly sums up the supervenient valence. It is not impossible to witness the scene without being moved thus. Anyone who does so, however, may plausibly be said to lack an objectively appropriate emotion.

This way of describing the situation avoids simple projectionism: what I perceive is not merely the shadow of my own response, but something about the character of a situation as a whole in the context not only of my own singular responses but also of the feelings and interests of others. I call this view *axiological holism*.

We do not apprehend value in discrete units but only in the light of a complex of factors that transcend individual experience. No single range of facts suffices for the overall fittingness or “success” of an emotional response. *Biological facts* will speak to its origins and may thereby assign it a proper function in the sense of Millikan 1989, but they will not determine its relation to currently relevant norms. *Social norms*, in turn, are every bit as likely to be irredeemably nasty as biological ones. (To endorse social norms as the touchstone of normativity would be to condemn all social reformers.) *Individual biography* sets up paradigm scenarios in terms of which each individual understands the world, but this defines only a narrow sense of fit between a current response and a present situation. That fit cannot be identified with value in any comprehensive sense, still less determine what is morally right (D’Arms and Jacobson 2000). Instead it is the totality of all of these factors—biological, social, personal, and more—that may properly be confronted with one another in the hope of arriving at something like reflective equilibrium. That holistic equilibrium of emotional responses is as close as we can come to reconstructing a notion of normative human nature: it will have to do, I suggest, as a substitute for that bankrupt notion. Vision provides distal information about our surroundings, despite the possibility of visual illusions. Just so, despite the

treachery of our emotional dispositions, emotions in general constitute apprehensions of axiological reality. We tell which is right and which is wrong much as we test the veracity of perceptual information: by appealing to corroborating evidence. This is, I believe, as close as we can get to the multimodal invariance that signals objectivity. It is not all that close, because we are looking for a coherent summing up without much hope of finding neat consilience along the way. But perhaps it is close enough.

#### NOTES

1. The term was coined in Gould and Vrba 1982 and refers to “a structure coopted for utility from different sources of origin . . . and not directly built as adaptations for their current functions” (Gould 2002, 41).

2. The word *satisfaction* is used here in the sense in which it is used in traditional formal semantics. See, e.g., Gupta 1998.

## Emotions, Thoughts, and Feelings

### *Emotions as Engagements with the World*

ROBERT C. SOLOMON

I have been arguing, for thirty years now, that emotions are a kind of judgment—or rather, a complex of interlocking judgments, desires, and intentions. Back in the seventies, it was an argument that attracted some attention but little sympathy. What provoked me then—and still provokes me now—is what I call the primitivist conception of emotions, the idea that emotions are basically physiological or (now) neurological syndromes conjoined with feelings that have only marginally to do with cognition or our engagements in the world. William James’s well-known theory, or at least that aspect of it he famously italicized, stated this view concisely: “*Our feeling of [bodily changes] as they occur IS the emotion.*” An emotion might be “triggered” by a disturbing perception, but its essence was physiology plus feeling. Thus my own analysis of emotion challenged this primitivism and maintained that emotions were, to employ a precocious word, *intelligent*. This meant also challenging the sharp divide between emotions and rationality, and I further insisted (inspired by Jean-Paul Sartre) that we question the established notion that the emotions *happen to us*, render us passive, and are essentially involuntary. Some aspects of this theory, notably the idea that emotions are in some sense “cognitive” and consist (at least in part) of evaluative judgments and thus display a kind of intelligence have become mainstream, even popular.<sup>1</sup> Other aspects of the theory, my voluntarist thesis in particular, continue to attract incredulosity.<sup>2</sup> But it is the “cognitivist” thesis that I would like to pursue here.

My bumper-sticker slogan, my walk-away-from-the-lecture catch phrase, has always been “Emotions are judgments.” The abbreviated slogan has too often been mistaken for a full-blown theory in itself, as if the emotions are *nothing but* judgments, and the particular term *judgment* has often

been misunderstood as overly detached and intellectual. But I have long argued that emotions as judgments should not be confused with singular summary judgments (such as might be used to briefly define them or distinguish one emotion from another), nor should a judgment be thought to be something deliberative, articulate, or fully conscious (Solomon 1988). To say that emotions are intelligent is not to say that an emotion is an aspect of intellect, and to insist that emotions are judgments is not to say that emotions are what some philosophers call “propositional attitudes.” Thus when Paul Griffiths took on what he misleadingly characterized as “propositional attitude” theories of emotion as the enemy of all that was true and scientific, I was both flattered and appalled—appalled as several of us had long insisted that propositions (or propositional *contents*) were not what emotions are essentially about, flattered because Griffiths’s abuse was surely a sign that our collective view of emotions had shifted, in Kuhnian terms, from being revolutionary to becoming the “normal” paradigm.<sup>3</sup> The current counterrevolution of affect programs and neuroreductionism says a lot about who we are and how far we have come. The cognitive theory is now the target—high praise for research that was once considered marginal. Progress in philosophy, I hesitate to say, is furthered more by the “dialectic” of one outrageous thesis after another than by cautious, careful, “normal” argument.

Emotions, I have always insisted, are *about the world*. With my concept of judgment I had tried to make clear that this was not a marginal fact about (some) emotions but the essence of all of them (including oneself, of course, as part of the world). Jamesian perceptions of the world are not mere triggers for emotion but are already shot through with the concepts and perspectives that constitute the emotion itself. The scholastic concept of “intentionality” was also an attempt to make this explicit, to insist that the emotions are always “about” something (their intentional *object*). Thus, judgments have intentionality, but I think that the traditional notion of intentionality—and, I now suspect, the concept of judgment, too—still lacks the keen sense of *engagement* that I see as essential to emotions, keeping in mind that thwarted or frustrated engagements characterize many emotions. Emotions are not just *about* (or “directed to”) the world but actively entangled in it. So I now want to improve my analysis by making this point central, that *emotions are subjective engagements in the world*. I still favor the use of “judgment” to make this point, but I now want to stress even more than I have before the idea that a judgment is not a detached intellectual act but a way of cognitively *grappling* with the world. It has at its very basis and as background a complex set of aspirations, expectations, evaluations (“appraisals”), needs, demands, and desires (which says something about why the reigning “belief-desire” analysis of emotions and intentions is so hopelessly impoverished).

The emotions-as-judgments view that I have defended is too often re-

ferred to as a “cognitive theory of emotions,” a borrowing from psychology and “cognitive science.” It is not a happy term. For one thing it would seem to leave out any essential role for desire or active engagement in the world in favor of the intellect or affectless “information.” I try to be very clear in my judgment-view that desire and engagement in the world (or the thwarting of our engagements) are essential to emotion. But the cognitive theory has become the touchstone of all philosophical theorizing about emotion, for or against. It used to be the battering ram (in its various guises) against the primitivist theories of James and his successors. Now it is the target of neurologically based “precognitivist” theories of emotion (as in the new slogan “Emotion precedes cognition”). But what exactly is a “cognitive” theory of emotions? The label “cognitive theory” is not mine, and I fought it for years, not just because it was misleading but also because “cognition” is so variously or ill defined. But it seems as if we’re stuck with it, so in self-defense I would like to take on “cognition” directly and try to say what I think it is and what it isn’t. What is a cognitive theory if one thinks of emotions as engagements in the world? In answering this question, I would also like to look once again at the role of feelings or “affect” in emotions and suggest in passing why I think that emotions are, at least sometimes, “chosen” and voluntary. As engagements (as opposed to things that happen to us), emotions would certainly seem to be good candidates for the role of “existentialia” through which we define ourselves.

WHAT IS AN EMOTION? EMOTIONS AS  
“THOUGHTS” (AND RELATED VIEWS)

I do not want to rehearse the long and fascinating history of James’s classic question, but it seems to me that there is a disagreement right from the beginning of our discussion as to what sort of a phenomenon an emotion might be. I want to reject, or at any rate call into question, the very *dimensions* of the emotional phenomena that are now under investigation and the subject of a great many contemporary theories. In his *Encyclopedia Britannica* article on emotions (1989), Carroll Izard begins by defining them as “brief” responses. In recent work by Joe LeDoux, Jaak Panksepp, and Antonio Damasio, an emotion is sometimes presented as if it is more or less over and done with in 120 milliseconds, the rest being mere aftermath and cerebral embellishment (LeDoux 1996, Panksepp 1992, Damasio 1999). An emotion, so understood, is a brief, preconscious, precognitive, more or less automatic excitation of an affect program. Now, I do not deny for a moment the fascinating work that these researchers have done and are doing, but I am more interested in substantial processes that last a long time—lifelong love, for instance. I am interested, in other words, not in those brief “irruptive” reactions or responses 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.

Those bold and intriguing discoveries in the neurobiology of emotion have stimulated a mantra of sorts, “Emotion *before* cognition,” which rather leaves the cognitive theory with its pants down, so to speak. (A fair turn, one might argue, from my old slogan, “Emotions are judgments,” i.e., “No emotion *without* cognition.”) But the very statement of the new mantra provokes a cognitivist rejoinder: Surely the very fact of a *response* indicates some form of recognition, and (just to say the obvious) recognition is a form of cognition. What gets thrown into question, therefore, is not the intimate connection between emotion and cognition but the nature of cognition itself. Cognition is not to be understood only as conscious and articulate. There are primitive preconceptual forms of cognition, “a cognitive neuroscience of emotion” (Lane 1999). These are not the forms of cognition or emotion that primarily interest me, perhaps, but they are extremely important in understanding not only the very brief phenomena studied by the neuroscientists but also the long-term emotional psycho-dramas that do interest me. Whatever else I may have meant or implied by my slogan “Emotions are judgments,” I was not thinking of necessarily conscious—or self-conscious—reflective, articulate judgments.

So, emotions are cognitive and they are processes, often long-term processes (and not merely “dispositions” to have brief emotional “episodes”). But “cognition” is a not very informative technical term, and there has been a lively debate (within the “cognitivist” camp) whether the type of cognition in question is best thought of as a belief, a thought, a judgment, or as something else. Many authors, Jeffrey Murphy and Kendall Walton, for example, suggest *beliefs*. Jerome Neu (1977) suggests that the cognitive elements that matter most are *thoughts*, a view that (at least nominally) goes back to Descartes and Spinoza. Several philosophers join me in defending the theory that emotions are *evaluative judgments*, a view that can be traced back to the Stoics. Cheshire Calhoun has suggested “seeing as” and Robert Roberts has offered us “construal” as alternative, more perceptual ways of understanding cognition in emotion (Calhoun 1984, Roberts 1988). Other theorists, especially in psychology and cognitive science, play it safe with “cognitive elements” or “cognitive structures” (e.g. Ortony, Clore, and Collins 1988; Gordon 1987). Some psychologists have argued that “appraisals” are “cognitions” (Lazarus 1970, 1994).

Many philosophers still prefer the technical term *intentionality*, although interpretations of this technical concept are often even less helpful than “cognition” (Kenny 1963). Patricia Greenspan has employed “belief warrant” while rejecting the “cognitive” theory in its more committal forms (Greenspan 1988). Michael Stocker (1996) is more directly combative when he rejects all of this in the defense of “affect” and “affective states,” although I have always suspected and will again here that Stocker’s “affect”

sneaks in a lot of what others portray as cognition. Ronald de Sousa (1987) suggests “paradigm scenarios,” an intriguing and more contextual and behavioral conception that is intended (among other things) to undermine the cognitive theory.

I think the common linkage between emotion and belief is misleading. Beliefs and emotions are related in many important ways: belief as precondition or presupposition of emotion, and belief as brought about by emotion (say, by way of wishful thinking or rationalization). But belief isn’t the right sort of psychological entity to *constitute* emotion. Beliefs are necessarily dispositions, but an emotion is, at least in part, an *experience*. A belief as such is not ever experienced. Belief may be perfectly appropriate in *explaining* emotion but it is inappropriate in the *analysis* of emotion. Moreover, belief is too loosely tied to perception to account for those cases where one has an emotion immediately upon coming into a situation, and it is too tightly tied to the logic of propositions to explain, for example, how it is that we can often hold conflicting (but not literally contradictory) emotions at the same time. Moreover, beliefs are propositional attitudes while many emotions are not (which is what’s wrong with Griffiths’s characterization). If Fred loves Mary and hates spinach, the objects of his emotions are Mary and spinach, respectively, not propositions. If Mary believes that spinach is good for you (and that, perhaps, is *why* she loves it) the object of her belief (but not her emotion) is the proposition that spinach is good for you.

In his early work—and I see little evidence of radical change since—Jerome Neu took the defining element of emotion to be the very Spinozistic notion of a “thought.” He makes it quite clear that one cannot have an emotion (or a particular kind of emotion) without certain types of thoughts. Emotions, simply stated, *are* thoughts, or dispositions to have thoughts, or defined by thoughts. (I am not considering here the very general Cartesian sense of *cogitationes* that would include virtually any mental process, state, or event, making the claim that emotions are thoughts utterly uninformative.) At the very least, Neu is correct when he says that thoughts are indicative of emotions and are typically produced during emotions. But I think that the notion of a “thought” is too specific and involves too much intellect to provide a general account of the emotions. To be sure, a person with an emotion will have thoughts appropriate to the emotion and the context shaped and constrained by his or her language and culture. In the case of adult human emotions, I think that this may necessarily be so. But if belief is too dispositional to capture the essence of emotion, thoughts are too episodic for emotions, which often turn out to be enduring processes rather than mere episodes. Thus a thought may punctuate and manifest an emotion, but it is in itself not a process. (*Thinking*, of course, is a process, but thinking is clearly too cerebral to characterize most emotions.)

I would thus challenge Jerome Neu's Blake-inspired title, "A Tear Is an Intellectual Thing" (Neu 2000), on the grounds that it is not the *intellect* that is typically engaged in emotion. But it is often charged (from the other side) that cognitive theory—which is conflated with the view that emotions are products of the intellect—*excludes affect*. The fact that many if not most emotions are nonreflective, however, has no bearing on the question whether (so-called) affect might be an essential part of the cognitive aspect of emotional experience. Thus what continues to exercise me are the ambiguity and confusion sown by both the overly vague notion of a "cognitive" theory of emotion and the accusation (made, for instance, by Michael Stocker [1976] and more recently by Peter Goldie [2000]) that such theories are pathologically dissociative insofar as they deny or neglect affective feelings. I want to argue that a cognitive theory can include affect, or much of what is intended by that misleading term.

Doubts about belief and thoughts explain the appeal of "perception" as the "cognitive element" most appropriate to the analysis of emotion. Ronnie de Sousa makes this case, as did John Dewey years ago, and I think that perception does indeed capture the heart of one kind of emotional experience, that which I would call "immediate" (though without bringing in the heavy philosophical baggage that term conjures up in the history of epistemology)—that is, those examples in which I have an emotional reaction to a situation unfolding right in front of my eyes, as in the sorts of examples employed (for obvious reasons) by William James in his classic analysis of emotion. Pointing out the close link between emotion and perception seems to me a plausible way of proceeding. Indeed one of its virtues is that it blocks the insidious distinction (still favored by some positivistic psychologists) that perception is one thing, while appraisal, evaluation, interpretation, and emotional response are all something else. Again, I prefer the concept of judgment precisely because it maintains these close ties to perception but at the same time is fully conceivable apart from perception.

De Sousa also introduces the tantalizing and very fruitful idea of a "paradigm scenario." In his book *The Rationality of Emotion* (1987), de Sousa does not take this as a specification of cognition so much as an alternative to cognition. His notion has an explicitly developmental and evolutionary bent, and his shift from emotional content to emotion context and behavioral training has always seemed to me a huge step forward in philosophical emotions research. De Sousa's theory also has two more virtues: its explicit bringing in the body in a behavioral (not physiological) mode and its explicitly social nature, where other people are not just objects of our emotions or those who (in some sense) share our emotions but, in a critical sense, coconspirators in the cultivation of our emotions.

Back there in ancient history, in “Emotions and Choice” (1973) and *The Passions* (1976), I suggested “judgment” to capture the close but not essential link between emotions and perception, leaving lots of room for emotions that were evoked by memories or associations and even emotions that could be cultivated through thought and reflection. If Neu had the camaraderie of the neo-Stoical Spinoza, I could claim a linkage with the original Stoics, although I obviously rejected their conclusion that emotions as judgments are as such irrational. I confess that I became wedded to the term *judgment*, and I spent a good many years insisting on all of the things I *did not* mean by that. In particular, I denied that judgments had to be magisterial or detached, that they were essentially episodic as opposed to processes developed over time (although, to be sure, one can *make* a judgment at a particular moment), that they were necessarily articulate or for that matter conscious. I took it as uncontroversial that animals make all sorts of judgments (e.g., whether something is worth eating, or worth chasing, or worth courting), but none of these are articulated or “spelled out,” nor are they subject to reflection. And I insisted that we make nonreflective, nondeliberative, inarticulate judgments, for instance, kinesthetic judgments, all the time. Kinesthetic judgments are rarely deliberative and rarely merit conscious attention, but they characterize an essential aspect of our ongoing engagement in the world. Michael Stocker has a poignant story about his falling on the ice, thus making both his fear and his bodily awareness painfully conscious. But the example only illuminates the fact that such judgments are not usually conscious at all.

Judgments, unlike thoughts, are geared to perception and may apply directly to the situation we are in, but we can also make all sorts of judgments in the utter absence of any object of perception. Thus while I find the language of “thought” just too intellectual, too sophisticated, and too demanding in terms of linguistic ability, articulation, and reflection to apply to all emotions, “judgment” seems to me to have the range and flexibility to apply to everything from animal and infant emotions to the most sophisticated and complex adult human emotions such as jealousy, resentment, and moral indignation. In other words, I argued the following to be essential features of emotion and judgment: they are about the world (including oneself in the world). They are episodic but possibly long-term processes as well. They must span conscious and nonconscious awareness. (In fact, I would argue, these are ill conceived as a simple polarity. There are many “levels” of consciousness.) Emotions as judgments must accept as their “objects” both propositions and ordinary objects of perception (imagination, memory, etc.). They must be appropriate both in the presence of their objects and in their absence. They must involve appraisals and evaluations without necessarily involving (or excluding) reflective appraisals

and evaluations. They must stimulate thoughts and encourage beliefs (as well as being founded on beliefs) without being nothing more than thoughts or beliefs themselves. And, of considerable importance to my larger view, they must artfully bridge the categories of the voluntary and the involuntary.

Thus emotions are like judgments, and emotions necessarily involve judgments. Does this entitle me to say that emotions *are* judgments? Well, not by logic alone, needless to say, and not if saying that emotions are judgments is intended to preclude emotions involving desires, intentions, feelings, and so on. But as a heuristic analysis and a way of understanding the peculiarities of emotion, I think it is of value. But, of course, an emotion is not a single judgment. An emotion is rather a complex of judgments and, sometimes, quite sophisticated judgments, such as judgments of responsibility (in shame, anger, and embarrassment) or judgments of comparative status (as in contempt and resentment). Emotions as judgments are not necessarily (or usually) conscious or deliberative or even articulate, but we certainly *can* articulate, attend to, and deliberate regarding our emotions and emotion-judgments, and we do so whenever we think our way into an emotion, “work ourselves up” to anger, or jealousy, or love. But the idea that we can and do “work ourselves” into an emotional state suggests a way that “judgment” is still inadequate to capture the nature of (at least some) emotions. When we make ourselves angry or “fall” (leap) into love it is not as if we simply come to “see” the world differently. We become deeply engaged in it. We produce in ourselves powerful demands and expectations and we project onto the “object” (the other person) the possible fulfillment or frustration of those demands and expectations. “Judgment” is less than adequate not because it is too detached or cerebral but because it fails to make fully explicit our active engagement in the world.

What has always attracted me to de Sousa’s notion of a “paradigm scenario” is the fact that it embeds an emotion and its cultivation in a social context. It makes clear that an emotion is not just an individual creation but is in essence “political”—that is, it has to do with our relations with other people (Solomon 1998). So, too, I excitedly read Jonathan Lear’s admirable interpretation of Freud (Lear 1990), which rightly excoriates Freud for his ill-defined conviction that emotions are biological “drives” and suggests instead the thesis that emotions are engagements in the world. I think that this is essentially right, although when Lear goes on to insist that the essence of all such engagements is “love,” I find his argument considerably less compelling. But I continued to realize that the core of the theory I had been developing for decades was the *dynamic* nature of emotions. It was always driven by my existentialism. That is why I have been so adamantly opposed to “primitivism” and insistent on the degree to which we should think of our emotions as our “doing” and as our responsibility. The idea that emotions are judgments supports that thesis, but meekly. (The

fact that we “make” judgments does not yield the conclusion that we choose or are responsible for our emotions.) But the existential idea that through our emotions we are engaged in the world (or as Heidegger wrote, we are “tuned into” the world) not only supports the possibility of choice and responsibility but also has built into it just that ambiguity between willfully engaging and “getting caught up in” that captures the fundamental ambiguity of the emotions themselves.

I am willing to acknowledge that different cognitive candidates may work better or worse for different emotions, and here I see further reason to heed and embellish the warning that Amelie Rorty and Paul Griffiths (for very different reasons) have issued—that “emotion” is not an adequate category for across-the-board analysis. Different emotions employ different kinds of cognition and require different kinds of analysis. There are more or less active emotions and more or less passive emotions (e.g., anger and grief, respectively). This is the virtue, perhaps, of such noncommittal notions as “cognitive elements” or “cognitive structures.” They are elastic enough to cover just about anything vaguely conceptual, evaluative, or perceptual. But while these seem to me to be useful conceptual tools for working out the general framework of cognitive theory, they clearly lack the phenomenological specificity that I am calling for here. Judgment seems to me to be, all in all, the most versatile candidate in the cognitive analysis of emotion. Engagement, however, seems to me an apt interpretation of the kind of judgment I want to pursue in a dynamic analysis of emotions. But by embracing the whole host of cognitive candidates, it is left open whether some emotions might be better analyzed in terms of perception, others in terms of thoughts or judgments, others in terms of construals, still others in yet more dynamic terms. The real work will continue to be with regard to particular emotions, and often with specific regard for the particular instance of a particular sort of emotion. Thus a “theory of emotion” might be too much to expect, as Amelie Rorty has argued (and argues in this volume), but it is always a good place to start, if only to appreciate the fascinating variety of phenomena we want to understand.

#### WHAT IS AFFECT? EMOTIONAL EXPERIENCE, FEELINGS, AND THE BODY

Michael Stocker and, more recently, Peter Goldie have accused the cognitive theory of neglecting feelings, or “affect.” I have said that in *The Passions* I was dismissive of the “feeling theory” that then seemed to rule what passing interest there was in the emotions (particularly in the work of William James and his successors). I argued that whatever else it might be, an emotion was no mere feeling (interpreting this, as James did, as a bodily set of sensations). But what has increasingly concerned me ever since and

brought me back to James is the role of the body in emotion, and not only the brain. In my original theory, it was by no means clear that the body had *any* essential role in emotion. I presumed, of course, that all emotional experience had as its causal substratum various processes in the brain, but this had little to do with the nature of emotion as such, as experienced. But as for the various physiological disturbances and disruptions that serve such a central purpose in James's analysis and in later accounts of emotion as "arousal," I was as dismissive as could be, relegating all such phenomena to the causal margins of emotion, as merely accompaniments or secondary effects.

Much of what makes up emotional experience, of course, are the complexes of our experiences of the world (including ourselves), shaped and colored by appraisals and judgments and the peculiarities of this or that particular emotional perspective. I used to think that this was *all* that was essential to emotional experience, and again, I treated the feelings of arousal and the like as experiential marginalia, of little importance to the phenomenological experience that could be understood only via the cognitive complexes that shaped emotional experience as such. But what led me to an increasing concern about both the role of the body and the nature and role of bodily feelings in emotion was the suspicion that my judgment theory had been cut too thin, that in the pursuit of an alternative to the feeling theory I had veered too far in the other direction. I am now coming to appreciate that accounting for the bodily feelings (not just sensations) in emotion is not a secondary concern and not independent of appreciating the essential role of the body in emotional experience. By this I do not mean anything having to do with neurology or the tricky mind-body relationship linked with Descartes and Cartesianism but rather the concern about the kinds of *bodily experience* that typify emotion and the bodily manifestations of emotion in immediate expression. These are not mere incidentals, and understanding them will provide a concrete and phenomenologically rich account of emotional feelings in place of the fuzzy and ultimately content-free notion of "affect."

The role of physiology in feeling is not straightforward. On the one hand, many physiological changes (including autonomic nervous system responses) have clearly experiential consequences—for instance flushing and the quickening of the heartbeat. Many others (including most neurological activities) do not. James was rather indiscriminate in his specification of bodily and "visceral" disturbances, but when he clearly referred to just those bodily processes (not necessarily disturbances) that had clear experiential or phenomenological effects, he did indeed capture something of importance even if he short-changed the nature of the emotion itself. I now agree that bodily feelings have been "left out" of the cognitive account, but I also believe that "cognition" or "judgment" properly construed captures

that missing ingredient. The analogy with kinesthetic judgments suggests the possibility of bringing feelings of the body into the analysis of emotion in a straightforward way.

What are the bodily feelings in emotion (though, to be sure, an emotion may last much longer than any given feelings, and feelings may outlast an emotion by several minutes or more)? The workings of the autonomic nervous system (quicken pulse, galvanic skin response, release of hormones, sweating) have obvious phenomenological manifestations (feeling excited, “tingly,” feeling flushed). Moreover, the whole range of bodily preparations and postures, many of them but not all of them within the realm of the voluntary, have phenomenological manifestations. Here too the well-cataloged realm of facial expression in emotion plays an important role. The feelings of our “making a face” in anger or disgust constitute an important element in our felt experiences of those emotions. So do other forms of emotional expression. The category of “action readiness” defended by Nico Frijda and others seems to me to be particularly significant here, not only in terms of a dispositional analysis of emotional behavior but rather in an account of emotion feelings. What is commonly called “arousal” might much better be considered as an aspect of “action readiness.” Anger involves taking up a defensive posture, and some of the distinctive sensations of getting angry have to do with the tensing of the various muscles of the body and bodily preparation for physical aggression. All of these are obviously akin to kinesthetic feelings, the feelings through which we navigate and “keep in touch with” our bodies. But these are not just sensations or perceptions of goings-on in the body. Both arousal and action readiness should be subsumed under the more general phenomenological rubric of *getting engaged in the world*.

The voluntary status of these various emotion preparations and expressions is intriguing. Many gestures are obviously voluntary and the feelings that go along with them are the feelings of activity and not passivity. Some bodily preparations, even those that are not autonomic nervous system responses, are not voluntary, and our feelings are more of “what’s happening” than of “I’m doing this.” Facial expressions are an especially intriguing category in this regard. Paul Ekman and others have analyzed what most of us have recognized: the difference between (for example) smiles that are genuine (that is, to a certain extent involuntary) and smiles that are “forced” (that is, voluntary but to some extent incompetent). Action-readiness includes both autonomic (involuntary) as well as quite conscious and reflective posturing—for example, adopting a face and stance fit for the occasion, a darkened frown and threatening gesture in anger, a “shame-faced” expression and a gesture of withdrawal or hiding in shame, a sentimental, even teary-eyed smile and a tender gesture in love. And each of these has its phenomenological manifestations, its characteristic sensa-

tions or feelings that are part and parcel of emotional experience (whether noticed or recognized as such or not).

To put my current thinking in a nutshell, I think that a great deal of what is unhelpfully called “affect” and “affectivity” and is supposedly missing from cognitive accounts can be identified with the body, or what I will call (no doubt to howls of indignation) *the judgments of the body*. George Downing has put the matter quite beautifully in some of his recent work. He writes of “bodily micro-practices” and suggests that emotions are to a very extent constituted by these (Downing 2001). This could, of course, be taken as just another attempt at behavioral reductionism, but Downing also insists that an emotion is essentially an experience. He also is quite happy to insist that cognitions (judgments) are also an essential part of any emotional experience. But he adds, and I agree, that a good deal of cognition is of a radically prelinguistic (very misleadingly called “precognitive”) nature. Building on the work of Hubert Dreyfus and suggestions in Heidegger and Bourdieu, Downing insists that a good deal of emotional experience and even emotional knowledge can be identified in the development of these bodily micropractices.

Does it make sense to call these judgments? I am sure the answer is yes, and I would defend this in two steps. First, I have already insisted that judgments are not necessarily articulate or conscious and so the sorts of discriminations we make and the construals that we perform are sometimes (often) made without our awareness of, much less reflection on, our doing so. Second, a relatively small store of human knowledge is of the form “knowing that.” Philosophers, of course, are naturally concerned with such knowledge, and that leads them not unnaturally to the prejudice that only such propositional knowledge is important. Not that they deny the need for all sorts of nonverbal skills of the “knowing how” variety, but these are hardly the stuff of philosophical analysis. First, perhaps, because there may be nothing distinctively human about them (animals display such nonverbal skills at least as impressively as humans) and second, it is well known that “knowing how” cannot be reduced to any number of “knowing that”-type propositions. But it is a distortion of cognition and consciousness to suggest that “knowing that”-type propositional knowledge is in any way primary or independent of “knowing how.” The thesis here obviously takes us back to Heidegger and Merleau-Ponty (and to a lesser extent, to Heidegger’s onetime disciple Gilbert Ryle). But since I have already insisted that emotional judgments are not necessarily propositional but rather ways of engaging the world, the way is open to make the further claim that they are not necessarily “knowing that”-type cognitions either.

It goes without saying that many of our most “knowing” responses to the world and the ways in which we bring meaning to our world may have more to do with the habits and practices we perform than with the ways

in which we think about and describe the world. Feelings of comfort (and discomfort) have a great deal to do with doing the familiar and finding ourselves acting in familiar ways with familiar responses. Feelings of comfort and discomfort range from satisfaction, frustration, and low-level anxiety to exuberant joy, full-blown anxiety, rage, and panic. Anger often involves feelings of discomfort, but to be anger (and not just frustration or irritation), the emotion must be further directed by way of some sort of *blame*, which in turn involves feelings of aggression and hostility that may themselves be readily traced (as James did) to specific modes of arousal in the body (tensing of muscles, etc). So, too, shame is at least in part a feeling of discomfort with other people, a feeling of rejection, as love is (in part) a feeling of unusual oneness with another. Feelings are not just sensations, nor are they mysterious “affects,” but felt bodily engagements with the world.

Thus 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. They may become reflective and self-conscious. What is cognition? I would still insist that it is basically judgment, both reflective and prereflective, both knowing how (as skills and practices) and knowing that (as propositional knowledge). A cognitive theory of emotion thus embodies what is often referred to as “affect” and “feeling” without rendering these unanalyzable. They may not be analyzable in the mode of propositional analysis, but neither are they simply manifestations of the biological substratum, as James and Griffiths suggest. There are feelings, “affects” if you like, critical to emotion, but they are not distinct from cognition or judgment and they are not mere “read-outs” of processes going on in the body. They are judgments *of* the body, and this is the “missing” element in the cognitive theory of emotions. They are profound manifestations of our many ways of emotionally engaging with the world.

#### NOTES

1. See, e.g., Coleman 1990.
2. See Solomon 2002.
3. See Griffiths 1997; Baier 1977; Solomon 1988.

### III

## Emotions and Feelings

*This page intentionally left blank*

## Emotion, Feeling, and Knowledge of the World

PETER GOLDIE

There is a view of the emotions (I might tentatively call it “cognitivism”) that has at present a certain currency. This view is of the emotions as playing an essential role in our gaining evaluative knowledge of the world. When we are angry at an insult, or afraid of a burglar, our emotions involve evaluative perceptions and thoughts directed toward the way something is in the world that impinges on our well-being, or on the well-being of those that matter to us. Without emotions, we would be worse off, prudentially and morally: we would not see things as they are, and accordingly we would not act as we should. Emotions are, according to this view, a Good Thing. No wonder we have evolved as creatures capable of emotion.<sup>1</sup>

So far as it goes, I myself favor this view. But I think that, as I have just put the view, it leaves out two important things about the emotions (and neglects other things I will not mention), each of which is utterly familiar to all of us. The first omission is feelings: feelings of the condition of one’s body, such as the feeling of the hairs going up on the back of one’s neck; and feelings directed toward the object of one’s emotion, such as feelings of fear directed toward the strange man approaching one in the dark alley. The second omission is that there is no mention of how profoundly and systematically our emotional feelings can mislead us—of how the emotions can distort perception and reason.

It might be objected at this point that, even if there is to be a place for feelings in an account of emotion, feelings are surely not going to be the sort of thing that could do anything like *mislead* us about the way things are—they are just not that sort of thing. *Thoughts* might mislead us, but

they, according to this objection, have already been included into the view of the emotions. I want to show that this objection is founded on a mistaken conception of what feelings are. The misconception is that feelings are *brute*: they can tell us nothing about the world and how to act in the world, and this is because feelings are not *about* anything (or if they are about anything, they are about only the condition of one's body). Moreover, feelings are inessential and peripheral to an account of what emotions are, although, of course, one might admit that they do sometimes occur. This misconception of the place of feelings—I have called it the *add-on* view (Goldie 2000)—overintellectualizes emotional life.<sup>2</sup>

The correct conception of emotional experience, which I want to put forward here, is one where the feelings involved are at center stage, playing a centrally important epistemic role in revealing things about the world.<sup>3</sup> Once feelings gain (or perhaps have restored to them) their rightful place in an account of emotion, and in an account of how emotions can give us knowledge of the world, I can then address the second omission: not only are emotional feelings a potential source of knowledge, they also have a tendency to mislead us, and to do so in a systematic way that cannot be dismissed as merely the tendency to throw up a few “false positives.” There is something especially troubling about the emotions here, which the view first canvassed cannot account for. It is only when feelings have their rightful and proper place in emotional experience that we may see how emotions can mislead us about the way the world is.

I will proceed as follows. To begin with, I will give an account of how the mind can be directed toward things in the world. This is the phenomenon of *intentionality*.<sup>4</sup> Then I will be able to show where emotional feelings fit into this account of intentionality: as bodily feelings and as feelings directed toward the object of the emotion. Both kinds of feelings can reveal things about the world: things about ourselves—what I will sometimes call *introspective knowledge*—including our thoughts, emotions, and the condition of our bodies; and things about the world beyond the bounds of our bodies—what I will sometimes call *extraspective knowledge*. And both kinds of feelings can mislead us in respect of our efforts to gain both introspective and extraspective knowledge. They can mislead us not only when we are in emotional turmoil, such as when in anger the red mist comes down over the eyes; they can also mislead us when we are ignorant of our emotions, such as when a deeply suppressed envy is quietly lurking in the background. If we do not have the right emotional dispositions, prudential and moral, that properly attune us to the world, then, I will argue, our emotions can distort perception and reason so that the world seems to us other than it really is: as I will put it, the emotions *skew the epistemic landscape*. Emotions may be a Good Thing, but we should not be too optimistic: they come at a certain epi-

stemic cost, which should not be ignored if one wants to be faithful to emotional life as we all live it.

Intentionality is the mind's capability of being directed onto things in the world. When you think about your spouse or partner, and about what he or she is doing at this very moment, or when you remember the tree house that you played in that summer when you were twelve, your thoughts and memories are directed toward these people or things as being a certain way; they are presented to you under a certain aspect.

A bodily feeling or sensation, the feeling from the inside of the condition of one's body, is intentional in just this sense: the feeling is directed toward an object, one's body, as being a certain way or as undergoing certain changes.<sup>5</sup> For example, when you feel an agonizing pain in your elbow, the object of the sensation is your elbow, which feels a certain way: agonizingly painful.<sup>6</sup>

Many emotions, especially short-term emotions such as fear, anger, and disgust, involve characteristic involuntary bodily changes—muscular reactions, hormonal changes, changes to the autonomic nervous system, and so on; their precise characterization is not my concern here. Such emotions have what Paul Ekman (1994) has called a “distinctive physiology.”<sup>7</sup> These bodily changes can be felt. For example, when you are afraid, you might feel the prickly sensation of the hairs going up on the back of your neck, and here the object of the feeling is the hairs on the back of your neck that feel a certain way: prickly, as if they were rising.<sup>8</sup>

A bodily feeling of this sort can provide a *prima facie* reason for one's believing that one is experiencing an emotion of a certain type.<sup>9</sup> It is only a *prima facie* reason because one can be mistaken about whether the feeling is part of an emotional experience.<sup>10</sup> You might, for example, feel your face going red and think that this is because you are embarrassed (that you blushed *in* embarrassment), while in fact your face is red because you have just come in to a warm room on a frosty day. (The converse is also possible: you can think that the feeling is *not* part of an emotion when it really is.) Moreover, even if the feeling of your bodily condition does truly reveal that you are experiencing some emotion or other, it may mislead you as to just which sort of emotion it is. For example, the tense feeling in your stomach as you get on the roller coaster might be one of fear rather than excitement. And as we move further away from the relatively short-term emotional responses such as immediate fear and anger, which have tended to be the central concern of Paul Ekman, we tend at the same time to move further away from there being a distinctive physiology that one can feel and that can provide a reason for believing that one is experiencing an emotion of a certain type.

So far, then, we have seen that bodily feelings can yield introspective knowledge about the condition of your body and about the type of emotion

that you are experiencing. But they can tell you more than that. They can also yield extraspective knowledge about the world beyond the bounds of your body. This may at first seem surprising. But consider an example from outside emotional experience. You experience a feeling of cold: this feeling could give you *prima facie* reason to believe that the ambient temperature in the room has fallen, and that the central heating has turned itself off. You might be wrong, however: your feelings might reveal something about yourself and not about the world beyond the bounds of your body: perhaps you are experiencing the first signs of flu. And the same principles apply with emotional experience. A feeling of the hairs going up on the back of your neck can give you *prima facie* reason to believe not only that you are afraid, but also that *there is something frightening nearby*. And if it is in fact *true* that there is something frightening nearby, then your bodily feelings will have yielded extraspective knowledge. But again, perhaps, the feeling might in fact reveal something about yourself and not about the world beyond the bounds of your body: for example, it might reveal that you are of an unduly nervous disposition, and in fact there is nothing frightening nearby. (I will return to this important point later.)

These principles do not go so far as to yield up an epistemic route from a bodily feeling to a belief about the object of your emotion *as such*; the most the bodily feeling can reveal is that there is *something* in the environment (you know not what) that has a certain property, such as the property of *being frightening*.<sup>11</sup> Let me here introduce a term for properties such as being frightening: I will call them *emotion-proper* properties, to capture the idea (borrowed from the ancient scholastics) that a property can belong to, or be proper to, an emotion. Other examples of emotion-proper properties are being disgusting (proper to disgust), being shameful (proper to shame), being enviable (proper to envy), and being worthy of pride (proper to pride).

This epistemic route (a route from a bodily feeling to an introspective belief that one is experiencing an emotion of a certain type, and from there to the extraspective belief that there is something in the environment that has the emotion-proper property) seems to me to be important and to capture a sense in which we are right to say that we should pay attention to, or “listen to,” our feelings. For example, you might wake up in the middle of the night feeling frightened. You are aware of your bodily condition as being characteristic of fear; you feel the hairs going up on the back of your neck and your heart racing. In such circumstances, it is not just intelligible, it is also *sensible*, to look around fearfully, ask yourself whether there was a strange noise from downstairs that woke you, whether there is a burglar in the house, and so forth. It might have all been a dream, but it makes sense to be sure before going back to sleep.

Of course, this epistemic route, beginning as it does with bodily feelings, is only available first-personally or from “the inside”: one cannot feel in *this* way the condition of someone else’s body.<sup>12</sup> However, this is not in any way

to suggest that bodily feelings are essentially private, or that we cannot often gain a grasp of other people's bodily feelings in different sorts of ways. When we think of feelings and recognize them as what they are, we are deploying, in our everyday thought and talk, a common set of concepts shared with others. We can as well think and talk of how another is feeling as we can think and talk of how we are ourselves feeling; we are speaking third-personally, but still personally. Because being able to think about and talk about one's own feelings requires a shared set of concepts, the child must come to learn the use of the concept from its caregivers and learn to apply the concept to herself; after all, there is nothing *intrinsic* to the experience of, for example, the hairs going up on the back of your neck to suggest that it is characteristic of a feeling of fear. And, at the same time, and without the priority of the first-personal over the second- and third-personal, or vice versa, the child must learn to apply the concept to others, on the basis of their bodily condition, or their behavior, including what others say about their own feelings.<sup>13</sup> So, to know what someone is feeling, one does not need—which is impossible—somehow to share his unique, immediate, and “privileged” method of access; all one needs to be able to do is answer the question, “What is he feeling?”<sup>14</sup> It is true that there is a unique, immediate, and privileged route to knowledge of one's own feelings, namely introspection “from the inside,” but this should not be taken to imply either that introspection is an indefeasible route to introspective knowledge (for it is not), or that knowledge of others' feelings is impossible (for it is not).

I emphasize that our everyday thought and talk of feelings is personal (first, second, and third, singular and plural) partly in order to contrast the personal perspective with the impersonal perspective of the sciences.<sup>15</sup> The two sorts of perspective, and the two ways of thinking and talking, are in different businesses, deploying different kinds of concepts—call them respectively *phenomenal* and *theoretical* concepts. When we use a phenomenal concept to think or talk about, for example, the experience of being afraid (my experience, your experience, his experience), we are thinking partly in terms of *what it is like* to be afraid. On the other hand, a purely theoretical concept of being afraid would be one which, roughly, picks out the emotional experience by its causal role, and which leaves out entirely what it is like to be afraid.<sup>16</sup> It might be the case, as some would argue, that our thoughts, feelings, and emotions can be fully described using these purely theoretical concepts.<sup>17</sup> If this is the case, then, from the impersonal perspective there will be nothing left out; a Martian, incapable of emotion, would be satisfied with it as a *complete* account of the workings of human beings. Yet, when we compare this perspective with the personal perspective, there is much that is left out: our Martian, in possession of a complete scientific account of the workings of a human being, would still have *no conception* of *what it is like* to have the experiences that the impersonal perspective picks out using its theoretical concepts. Scientific investigation of the emotions,

from a purely impersonal perspective, deploying purely impersonal theoretical concepts, inevitably—and quite appropriately from this perspective—makes no *use* of phenomenal concepts, which are available only from the personal perspective, whereas our everyday thought and talk is *essentially* personal and makes *essential* use of phenomenal concepts.<sup>18</sup> In one sense, then, the impersonal stance of the sciences leaves nothing out; in another sense, it leaves much out, for it leaves out our ordinary, everyday way of thinking of our emotional experiences from the personal perspective.<sup>19</sup>

To sum up where we are so far, then, the position is as follows. Many emotional experiences involve characteristic bodily feelings. These are intentional, being directed toward the condition of one's body. Such bodily feelings can provide *prima facie* reasons for believing that one is experiencing a certain sort of emotion (introspective knowledge), and for believing that there is something in the environment that has the related emotion-proper property (extraspective knowledge). But bodily feelings alone cannot reveal to you what your emotion is about. The feeling of the hairs on the back of your neck going up can tell you that there is something frightening nearby, but it cannot tell you that this something is a burglar. The other kind of emotional feeling, on the other hand, is directed toward the object of one's emotion as such—for example, your feeling of fear that is directed toward the burglar.

When an emotion is directed toward its object, then this is a sort of *feeling toward* the object. The object can be a thing or a person, a state of affairs, or an action or event: when you fear a burglar, the object of your fear is a person; when you are angry about the level of unemployment, the object is a state of affairs (or a fact); and when you are disgusted at the drunken behavior of a man on the train, the object of your emotion is an action.

Feeling toward is unreflective extraspective emotional engagement with the world beyond the body; it is not a consciousness of oneself, either of one's bodily condition or of oneself as experiencing an emotion.<sup>20</sup> Such feelings are thus something that a creature incapable of self-reflective thought—a dog or a toddler, for example—could achieve. We adult humans, however, are capable of a turn of reflectiveness: we are capable of noticing through introspection that we have feelings toward something. For example, you are in an audience at a conference and a new speaker takes the stand. A friend next to you observes that you are becoming increasingly restless; your fingers are drumming on your notepad, your foot is tapping, and your lips and jaw are tense. Your friend surmises, rightly, that you are becoming irritated by something about the speaker: his manner, what he is saying, or something. But you are not aware of this. You have not noticed that you are feeling irritated by the speaker, yet you do have feelings of irritation toward him. Then your friend passes you a note, asking what is irritating you; and *then* you notice, or become aware, that you are feeling

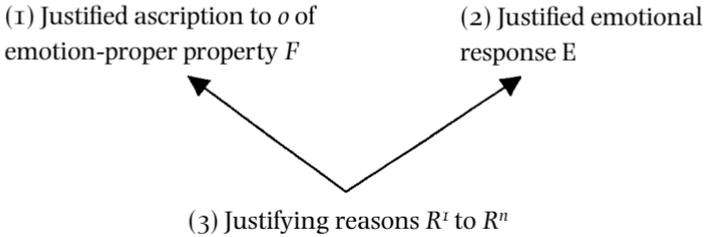
this emotion. Before seeing the note, you had feelings of irritation toward the speaker but were not aware that this was so.<sup>21</sup>

The notion of having feeling toward things in the world may seem to be a puzzling one: it is not a familiar sort of “attitude” in the philosopher’s armory, unlike, for example, perception, belief, desire, memory, or imagination. There is, accordingly, a philosophical approach that seeks to give an account of the emotions by a sort of divide and rule: first, to capture their intentionality in terms of these familiar unemotional attitudes—that is to say, attitudes that we can have when we are not experiencing an emotion; and second, as an afterthought, so to speak, to capture what emotional experience is like—its phenomenology—by reference to feelings as nonintentional states or as intentional states that are merely bodily feelings, not directed toward objects in the world beyond the body. This is a version of the “add-on theory” of emotions.<sup>22</sup> Rather, emotional feelings are inextricably intertwined with the world-directed aspect of emotion, so that an adequate account of an emotion’s intentionality, of its extraspective directedness toward the world outside one’s body, will at the same time capture an important aspect of its phenomenology. Intentionality and phenomenology are inextricably linked. Your feeling afraid of a burglar, or your feeling angry about the level of unemployment, involves having feelings (of fear, of anger) toward the object of your emotion, and this sort of intentional attitude cannot be identified with, or analyzed into, terms that refer only to unemotional attitudes.<sup>23</sup> But this is not to suggest that perception, belief, and reason are not involved in, or closely related to, emotional experience. Indeed, they are. And it just here that the emotions become epistemologically problematic.

When we respond emotionally to things in the environment, we also, as part of the same experience, typically *perceive* those things as having the emotion-proper property. For example, as a caring parent, you see the out-of-control toboggan hurtling straight for your child, you feel fear, and you see the toboggan as *being frightening*. Or you feel disgust at a maggot-infested piece of meat, and you see the meat as *being disgusting*. Moreover, in the typical case, the emotional response, combined in phenomenology with the perception of the object as having the emotion-proper property, will involve the experience of the emotion as being *reasonable* or *justified*. One might put the idea like this: an emotional experience, in the sorts of cases I am considering here, typically involves an extraspective (typically perceptual) judgment, about something in the world as having an emotion-proper property (for example, the judgment that the meat is disgusting), as well as an emotional feeling, which is experienced as reasonable, directed toward that thing (for example, a feeling of disgust at the meat).<sup>24</sup>

So an emotional experience typically *seems* to one to be reasonable or justified. But what makes it, in fact, justified? A possible reply is that the

emotion (disgust at the meat) is justified by the perceptual judgment (the judgment that the meat is disgusting). But this reply is not right. Rather, an emotion, if it is, in fact, justified, will be justified by something else external to the emotion itself and the perception: it will be justified by reasons, and these reasons will also justify the ascription to the object of the emotion-proper property involved in the perceptual judgment. Thus the fact that the meat is maggot infested is a reason that justifies your perceptual judgment that the meat is disgusting, and the fact that it is maggot infested will also justify your feeling of disgust. This relationship between (1) justified ascription of emotion-proper properties to the object of the emotion (the meat's being disgusting), (2) justified emotional response (your feelings of disgust directed toward the meat), and (3) justifying reason or reasons (such as the fact that the meat is maggot-infested) can be shown diagrammatically, where the lines represent justifying relations:



It can be seen that (continuing with the disgusting piece of meat as an example) the reasons that justify the ascription of disgustingness to the piece of meat (the fact that it is maggot infested, etc.) are the *very same* reasons that make feeling disgust justified on this occasion. It is neither one's perceiving it to be disgusting that justifies one's disgust, nor is it one's feeling disgust that justifies one's perceiving it to be disgusting; the justifying route is only from the bottom up.<sup>25</sup>

The epistemology of the emotions, on the other hand, often begins at the top: one often first either notices that one is experiencing the emotion (top right in the above diagram), or one perceives the object as having the emotion-proper property (top left); only later does one become conscious of the reasons that justify both one's emotional experience and the content of one's perception.<sup>26</sup>

Now, part of what lies behind the intuition that our emotions are a Good Thing and that they should be "listened to" is that they can play this epistemic role: they can enable us to see things in their true light and to make justified perceptual judgments in ways that we would not otherwise be able to do: emotions can reveal saliences that we might not otherwise recognize with the same speed and reliability. For example, we can immediately see that something is disgusting in a way that we would not be capable of if we were not capable of feeling disgust. Our emotional dispo-

sitions can, so to speak, *attune* us to the world around us, enabling us quickly and reliably to see things as they really are, and thus to respond as we should. In short, emotions enable us to *get things right*.<sup>27</sup>

To have the right emotional disposition is not, however, *sufficient* for getting things right. Other factors can also unduly interfere with one's emotional response on an occasion, leading one to fail to get things right. I will mention two notable ones. First, one's mood can affect one's emotional response: for example, if one is in an irritable mood (perhaps through drinking too much coffee), then one is more likely to find a remark insulting and to get angry. Second, a recent emotional experience in relation to one thing can resonate across to some other, unrelated thing: for example, if one has just had the terrifying experience of being mugged in an alleyway, then one may be especially likely to be jumpy every time there is a knock at the door; your emotional disposition gets temporarily put "out of tune."

In short, then, the picture looks like this: if one is of the right emotional disposition, and if there are no other undue influences, then one will feel the right emotions, and one will perceive things as having emotion-proper properties when and only when they do have such properties. One's emotions will then help one to find one's way around the world and to gain extraspective knowledge, so one will be right to "listen to" them. But if one is not properly disposed, or if there is some undue interference with one's emotional response, then there is a significant risk of getting things wrong. Not only that; one's emotions can also *distort* perception and reason.

As I have already said, it is typical of emotional experience to consider one's emotional feeling to be justified and to perceive the object of one's emotion as having the emotion-proper property. So far so good. But what if, without one's knowing it, one's emotional response is *unjustified*, and the object of your emotion does *not* have the emotion-proper property that it seems to have? (Perhaps you think you have the right emotional disposition but you do not, or perhaps your mind is subject to other undue influences that you are not aware of.) In such cases (and here is the worry), one's emotional feelings tend to *skew the epistemic landscape* to make it cohere with the emotional experience: referring back to the diagram, the epistemic landscape tends to be skewed *downward*, so to speak: we seek out and "find" reasons—reasons that are supposed to justify what is in reality the unjustified ascription of the emotion-proper property, and that, at the same time, are also supposed to justify the emotional feeling. The feeling directed toward the object of the emotion, and the related perception of the object as having the emotion-proper property, tend to be *idées fixes* to which reason has to cohere. The phenomenon is a familiar one: when we are afraid, we tend unknowingly to seek out features of the object of our fear that will justify the fear—features that would otherwise (that is, if we

were not already afraid) seem relatively harmless.<sup>28</sup> This is surely part of what is behind the commonsense intuition that our emotions can mislead us: they are *passions*, which, like *idées fixes*, we can be in the grip of.<sup>29</sup>

The skewing process can be continuous while the emotion is in place, operating on new information as it comes in. One's emotions and emotionally held perceptual judgments *ought* to be open to be shown to be wrong by new evidence, but when new evidence does emerge, one tends not only to be insensitive to that evidence, but also, for the sake of internal coherence, to doubt the reliability of the source of that new evidence.

An extreme case is Leontes in Shakespeare's *A Winter's Tale*, who becomes jealous of his wife Hermione and is convinced that he has been cuckolded by his boyhood friend Polixenes. Although his jealousy is not justified, everything now *seems to him* to justify his jealousy in what has suddenly become an emotionally skewed epistemic landscape: the way Hermione and Polixenes behave together, the sudden uncertainty about whether his daughter looks like him, the disappearance of his previously trusted friend Camillo, who is now a "false villain." He even rejects the evidence of the oracle of Apollo, that "Hermione is chaste; Polixenes blameless; Camillo a true subject; Leontes a jealous tyrant; his innocent babe truly begotten." Apollo, angry at having his word doubted, immediately wreaks his terrible revenge by bringing about the death of Leontes' son and wife. Only then does Leontes finally come to recognize that he has "too much believ'd his own suspicion"; and then it is too late.

A possible objection to my position here is that there is nothing special about the emotional case: people are generally subject to all sorts of well-documented cognitive deficiencies, such as the confirmatory bias,<sup>30</sup> and the emotional case is just an instance of this. One response to this objection, which I find independently attractive but will not pursue here, is that perhaps more of these cognitive deficiencies can be traced back to the emotions than might at first be thought. The other response, which I will put forward here, is that there is something special about the emotional case: emotional feelings, and emotionally held perceptual judgments about things as having emotion-proper properties, are *more intransigent* than are their non-emotional counterparts, and thus the skewing of the epistemic landscape (for the sake of internal coherence) tends to be toward the preservation of the emotionally held *idées fixes* at the cost of the unemotional thoughts.

Now, it is surely a reasonable and quite general epistemic requirement that one be willing and able to "stand back" to reflect on, criticize, and if necessary change our way of thinking of things. And this general requirement surely rightly ought to include critical reflection on the way that one's emotions can have this skewing effect. This is *obviously* the case when one knows that one's emotional responses are not as they should be—those atypical occasions when at the time one knows that one's emotion is not justified. But it is also the case when one has no particular reason to doubt

one's emotional responses: even then one should try to be especially watchful and reflect dispassionately on the evidential support for one's emotional feelings and for the related emotionally held perceptual judgments.

But doing this is not so easy, largely because one's epistemic landscape has *already* been skewed; so, like Leontes, one is not in a position, from the here and now of emotional experience, to take the dispassionate view of the evidence that the epistemic requirement demands. The problem is very familiar to everyday life: how to satisfy this epistemic requirement when one is in the swim of emotional experience. Consider this example. You feel in despair about your job. The job seems hopeless, and it seems to be hopeless for all sorts of reasons that seem to *justify* your feelings of despair: there are no decent prospects for promotion; most of your colleagues are people with whom you really have very little in common; you do not seem to be able to get the work done properly; the journey to and from home is a nightmare; and so on. Your friends, not in the here and now of this emotional experience, assure you that things seem this black only *because* you are feeling so despairing (you used not to be like this; perhaps some Prozac might help?). You try to stand back and see things as others do (maybe things will look a bit brighter in the morning). And you might succeed in doing this to some extent. But you could still think that it is your friends who are wrong: they believe these things because they do not see that things really *are* hopeless and how *right* you are to be in despair (Prozac might lift the despair, but the job will still be hopeless). The question remains: Is it you, or is it the job?

This leads me directly to a further, deeper worry about how emotion can distort perception and reason by skewing the epistemic landscape. So far, my focus has been on cases where one is aware through introspection that one is experiencing a particular sort of emotion; in the example just discussed, you are aware that you are in despair. But it would be a grave mistake to think that our emotional feelings are *always* transparent to introspection in this way: we can be ignorant of our own psychological states.

To begin with, as I have already mentioned, one can sometimes not be sure what emotion it is that one is experiencing—fear or excitement at the roller coaster. Second, as I have also already mentioned, one can have feelings without noticing them—such as that unnoticed irritation at the speaker. (A sort of limiting case here is feelings that are repressed in the Freudian sense.) And then third, emotions can continue to resonate in one's mental economy long after they are, as it might seem, "over." In all these sorts of cases (and others besides), emotion can distort perception and reason in the ways I have been discussing. But now, one is in the worrying position of not knowing what emotions, if any, are at work: one lacks introspective knowledge in this respect. One can therefore be inclined to think that one is being "dispassionate" when one is not, or to think mistakenly that one sort of emotion is at work rather than another. Thus one

has no way of knowing how to direct one's watchfulness in the quest for extraspective knowledge. One is in the position of having an epistemic requirement, which one knows of and acknowledges to be reasonable, but which one does not know how to satisfy.

Let me give an example. A long time ago you were very angry with a colleague at work because he failed to turn up to a meeting that you were chairing, and at which his presence was essential. How could he do this when he *promised* to be there? You thought your anger to be thoroughly justified on the grounds of his being so unreliable and inconsiderate. The following day, though, he came to see you with a full explanation, and was extremely apologetic. His son had been taken suddenly ill and had to be rushed to the hospital, there was no chance of getting to a phone, and so on. You put your anger behind you, as you should do, realizing that your anger, although understandable at the time, was not justified, for he really had a good reason not to be there, and a good reason why he could not give you advance warning. Later still—*much* later—you are asked to provide a reference about this colleague. Without your realizing it, the content of what you say is affected by the residue of your anger, which still lies deep in the recesses of your mind. Of course, you do not go so far as to state outright that he is unreliable and inconsiderate, for your memory of the incident is at best only hazy; and anyway, as it later emerged, he was neither unreliable nor inconsiderate. But still, unknown to you, for you think that you are being fair and dispassionate in what you say, your reference is not as favorable as it would have been if the incident had never taken place. Aware of the epistemic requirement, you ask yourself, "Am I emotionally involved here? Because if I am, I should be especially watchful." But the answer comes back, "No, I am *not* emotionally involved"; moreover, you might sense a certain puzzlement as to what *sort* of emotion might be at work on this occasion. And if you were reminded of the long-past incident, you might insist that any anger that you felt all that time ago is no longer at work, distorting reason.<sup>31</sup>

Where does this discussion leave us? Feelings are restored to their rightful place in emotional experience: intentional, and playing a centrally important role in our finding our way around the world. But then the worries begin to arise: our emotions can systematically mislead us. First, while one is in the swim of life, emotionally engaged with what is going on, one's epistemic landscape is liable to be skewed by one's emotional feelings, *idées fixes* to which perception and reason is forced to cohere. To avoid this as much as possible, one should see oneself as subject to the epistemic requirement to reflect on what one takes to be reasons, to make corrections where necessary, and to be aware that one should be especially watchful when one is emotionally engaged. But then the further worry arises that one can be emotional without knowing it, so one has no way of knowing

that one's perception and reason are being distorted, or in what ways. Even if one were to accept the idea (which I am inclined to endorse) that emotions are always somewhere at work in our psyche, and thus to accept that a special watchfulness is *always* required, one will still be no wiser as to *how* to apply this epistemic requirement at any particular moment. This seems to me to be especially troubling: lack of introspective knowledge impedes the attainment of extraspective knowledge.<sup>32</sup>

So the view with which I began (and which I avoided calling "cognitivism") begins to seem not only incomplete but also unduly *optimistic*,<sup>33</sup> as if we emotional creatures remain firmly governed by reason, and as if emotional feelings do not systematically tend to mislead us and distort perception and reason, in ways that are not always knowable from the here and now of emotional experience. Rather, as we all know from our own experience, emotional life is often messy, confusing, and difficult.

## NOTES

This chapter draws on material from two other papers where these issues are considered and developed in greater detail: Goldie 2002 and forthcoming. Many thanks to the editors of *Phenomenology and Cognitive Science* and *Emotion, Evolution, and Rationality* for allowing me to do this. Thanks also to Tim Crane, David Papineau, and Finn Spicer (I owe the expression "epistemic landscape" to Finn) for their help, and special thanks to Bob Solomon for his comments and suggestions, and for inviting me to contribute to this collection.

1. An example of this sort of view is in Nussbaum 2001. Nussbaum calls her view "cognitive" or "cognitive-evaluative" (23), and she holds that emotions are judgments of value. Other examples are Kenny 1963, Alston 1967, Green 1972, Davidson 1976, Solomon 1993, and Farrell 1980. There is a useful survey in Deigh 1994.

2. Nussbaum is again a good example here. So far as concerns bodily feelings, these, she says, are "without rich intentionality or cognitive content," or even "nonintentional"; and as there is variability in feelings across people and cultures, and as we should admit the possibility of nonconscious emotions, bodily feelings cannot be part of an emotion's identity conditions. So far as concerns what she calls "feelings with a rich intentional content," "the 'feeling' now does not contrast with our cognitive words 'perception' and 'judgment,' it is merely a terminological variant of them" (2001, 60).

3. Nowhere do I insist that bodily feelings are a necessary condition for emotion; whether the other sort of feelings is necessary, I am not so sure. I choose the expressions "center stage" and "centrally important" with care.

4. Unfortunately, even if it were in my powers to do so, I cannot provide here anything like a defense of this particular account of intentionality; I am afraid I will have to ask this much to be taken for granted. Crane 2001 contains an excellent and accessible discussion of the numerous problems that intentionality gives rise to.

5. See Armstrong 1968, Martin 1995, and Crane 1998 and 2001. Crane argues that intentionality is the mark of the mental; nothing that I say is inconsistent with this.

6. One might naturally say (as Bob Solomon has suggested to me) that the object of the sensation is the pain. Agreed. But if we were to take seriously the notion of pain as an *object* of sensation, all sorts of philosophical difficulties would arise. It is better to say, as Crane argues (2001, 78–83), that being in pain needs a part of the body as an object to “complete” it. And this is not contrary to linguistic practice: “I am in pain.” “Where is it?” “In my elbow.”

7. There may be good evolutionary reasons why this is so. See, for example, Griffiths 1997.

8. Of course one need not perceive one’s bodily changes under the description by which they would be picked out by the sciences. For example, you need not perceive an endocrine system change *as such*; perhaps all you need to perceive is what you think of as “that funny feeling in my guts.”

9. Throughout I use the term *reason* in the standard normative sense, in which, if a consideration is a reason, then it is a *good* reason. A *prima facie* reason is a consideration that appears *at first sight* to be a reason, but which may turn out, in fact, not to be a reason. For example, your seeing something as red is a *prima facie* reason for believing it to be red. But if you were wearing contact lenses that made red things look blue and blue things look red, then your seeing something as red is not a reason (that is, not a good reason) for believing it to be red.

10. The experiments by Schachter and Singer (1962) show this. Wollheim (1999, 115–28) has an excellent discussion of the role of feelings in emotion and of these experiments. Feelings can also mislead one about the condition of one’s body, as they do in the phantom limb example.

11. The belief that there is something frightening is thus purely quantificational.

12. For an argument for this, see Martin 1995, and for an alternative view, see Brewer 1995.

13. See the papers by Brewer, Hutto, and Smith in Goldie 2002.

14. See Austin 1946, 96–97.

15. It is a potential source of confusion to speak of the impersonal perspective of the sciences as “third-personal,” as does for example Chalmers (1996). By “the sciences” I have in mind here particularly cognitive science; I am not sure where, for example, empirical psychology stands.

16. Chalmers (1996) makes the same distinction, but he calls them “phenomenal” and “psychological” (i.e., what I call theoretical) concepts. Our ordinary, everyday way of thinking of emotions probably involves an amalgam or fusion of both sorts of concept (see Papineau 2000, 98). If so, the appropriate contrast is between, on the one hand, such fused concepts (a fusion of phenomenal and theoretical), and, on the other hand, *purely* theoretical concepts.

17. For such a materialist view, see, e.g., Papineau 2000.

18. Science, however, will presumably need to *mention* phenomenal concepts in order to explain the new powers and potentialities of thought, feeling, and imagination that arise from being able to use these concepts.

19. I discuss these issues in more detail in Goldie (forthcoming a), where I draw an analogy, in respect of emotional experience, with Frank Jackson's famous thought experiment (Jackson 1982 and 1986) of Mary, the scientist brought up in a black and white world, who knew all the scientific facts about color, but did not know what it was like to see red.

20. One can, however, have feelings toward one's own body that are not bodily feelings as I have characterized them. Here the object of the emotion is the body image (Gallagher 1995). For example, I can feel disgust at my obesity or anger at my useless arthritic fingers.

21. The difference between unreflective engagement with the world and reflective awareness of one's engagement should not be taken to be a stark one: one can be more or less aware of how the world strikes one. See Stocker 1983, 14. This account should be able readily to accommodate repressed feelings not available to be recognized through introspection alone.

22. See my discussion of Nussbaum in n. 2 above.

23. I argue for this in Goldie 2002.

24. As John Skorupski puts it, "The affective response typically carries with it a normative impulse" (2000, 125). The atypical cases are not like this: these are the occasions where one realizes *at the time* that one's emotional response is *not* reasonable or justified. For example, you feel afraid of the mouse in the corner of the room, and yet at the same time you know that your feelings are not justified. In these atypical cases, although the object might still *seem* to have the emotion-proper property (the mouse does *seem to be* frightening), one is not inclined, as one is in the typical case, to consider one's emotional response to be justified, and one withholds the perceptual judgment (that the mouse *is* frightening). There is, thus, the possibility of acknowledging, in one's own case, and at the same time as the emotional experience takes place, that things are not really as they seem: the mouse seems frightening, but you know that it is not, for you know that your fear is not justified.

25. None of the relata can be analyzed in nonnormative terms at pains of falling foul of Moore's open question argument; see Moore 1903. More formally, the relation between (1), (2), and (3) can be put as a schema: "An object *o* has emotion-proper property *F* iff it is possible for *o* to be the object of a justified emotion *E*; and the reasons, *R*<sub>1</sub>. to *R*<sub>*n*</sub>., that justify the ascription of *F* to *o* will be the same reasons as those that justify *E*." Emotion-proper properties that are related one-to-one to emotions will generally be at the "thicker" end (disgusting-disgust; hateful-hate; shameful-shame). Others will be much more complicated in their relations. There are a number of other issues that would have to be dealt with in a fully developed account, but I will have to put these to one side here.

26. One's reasons, then, will not be part of the emotional experience itself. An analogy with aesthetic experience might help here. One might have a certain aesthetic experience on seeing a sculpture (or a human face, or the curve of a valley) as being graceful. But one might not be able to articulate what makes it graceful, and, correlatively, what justifies one's feeling aesthetic pleasure on looking at it. Nevertheless, there will *be* reasons why it is graceful (perhaps it is the particular shape of it), for aesthetic properties depend on nonaesthetic properties (see Sibley 1965).

27. Having the right emotional disposition, the deployment of which will enable one to get things right, is a profoundly normative notion. As Aristotle saw, to have such dispositions is part of what it is to be virtuous (where the virtues are both prudential and ethical and involve virtues of thought as well as ethical virtues of character). As Aristotle put it, the virtuous person will feel—that is, have emotions—and act “at the right times, about the right things, toward the right people, and in the right way . . . this is the intermediate and best condition, and this is proper to virtue” (1985, 1106b20).

28. Remember, I am not concerned here with those atypical occasions (such as the fear of the mouse) when one knows at the time that one’s emotional response is not justified, but the emotion remains; for on those occasions one’s reason stands *opposed to* one’s emotional feelings, and one recognizes that it is one’s feelings that are in error.

29. Overintellectualizing accounts of the emotions struggle to explain how we can be *in the grip of* unemotional attitudes, such as judgments and beliefs.

30. See, for example, Nisbett and Ross 1980.

31. For some related empirical research, see Zillman and Cantor 1976.

32. In Goldie forthcoming, I discuss this difficulty in relation to the intellectual virtues and virtue epistemology. I also discuss the question of whether we can be properly blamed or held accountable for our ignorance.

33. For example, Nussbaum optimistically says that “emotions . . . do go away when the relevant beliefs about the object and about value alter. . . if I am convinced that the wrong did not really take place, or was not really a wrong, my anger will go away” (2001, 131).

## Subjectivity and Emotion

CHESHIRE CALHOUN

We take a stance toward emotional responses. Sometimes it is one of haughty superiority: I can rise above my emotions, cleanse my vision. Sometimes it is one of embarrassed nakedness: my emotions reveal or betray my self. Both are stances toward subjectivity; emotions are subjective. But what does this mean? Here are some arguments: Emotions, when corrupted by base desires (as they often are), distract us from clearly viewing the Forms (Plato). Emotions are often self-interested (Hutcheson, Hume). Emotional reactions cannot be universalized (Kant). They are instinctive (Freud, James). Many emotions resemble blind prejudices (Brentano). Emotional consciousness constitutes objects as dangerous, adorable, or boring (Dewey). It is the concrete, existing individual that feels emotions (Kierkegaard, Buber). Emotional consciousness magically transforms the world (Sartre).

From these meanings, one might distill out two quite different comments about emotional subjectivity. First, emotions are epistemically subjective. Emotional consciousness does not mirror reality. It is prejudiced, partial, constitutive, perhaps even altogether noncognitive. Second, emotions are biographically subjective. Emotions are tied to the individuality of individuals, reflecting their interests, context, and style of living. Though there are very different meanings of emotional subjectivity, philosophers who talk about emotional subjectivity often run them concurrently. If emotions are biographically subjective, if what I feel somehow reflects myself, then, so the reasoning goes, they must also be epistemically subjective. What reflects me can only be biased and partial.

This conviction that emotions are, in the end, epistemically subjective drives the wedge between reason and passion and buttresses the philosoph-

ical devaluation of emotion. This result is reason enough to ask again "What makes emotions subjective?" and to hope for a different answer.

There are other reasons as well for skeptically examining the equation of emotional subjectivity with epistemic subjectivity. One has to do with a condition for constructing a successful cognitive theory of emotion. Cognitive theories, which argue that emotions are or entail beliefs, offer the hope, often as their selling point, of excising the wedge between reason and passion. A problem facing all such theories is to explain what distinguishes emotions from nonemotional beliefs (i.e., to explain how emotions have not simply been reduced to beliefs). One answer is that while emotional and nonemotional beliefs are, *qua* beliefs, the same sorts of thing, emotional beliefs are accompanied by some affective tone.<sup>1</sup> In other words, the distinctively emotional aspect of emotion is something nonrational. But this answer only moves the wedge between reason and passion to a new location. A second, more popular, answer appeals to the greater subjectivity of emotional beliefs. So, for example, Solomon (1973) claims that emotions are judgments but adds that emotional judgments are hasty ones. By insisting on the epistemic subjectivity of emotional belief, this tack also fails to excise the wedge. Both tacks locate the distinctively emotional in the nonrational or irrational. If we are serious about overcoming the reason-passion split, we will need some way of distinguishing emotions from nonemotional beliefs that does not invoke epistemic subjectivity. This might be done by picking up on the idea of biographical subjectivity. What is distinctively emotional in emotional beliefs may be the strong connection of those beliefs with our personal biographies. This, I shall argue, has nothing to do with epistemic subjectivity.

One last reason for this move. Feminist philosophers have argued that neither reason nor emotion is a gender-neutral concept.<sup>2</sup> Philosophical constructions of reason and emotion play into and out of social constructions of masculinity and femininity. It is no accident that the capacity for epistemic objectivity has historically been located in both reason and men, while the defect of epistemic subjectivity has been identified with both emotion and women. If we now think this was a mistake in the case of women, we should also consider whether it was not also a mistake in the case of emotion.

#### DOUBTS ABOUT EPISTEMIC SUBJECTIVITY

What does it mean to say that emotions are epistemically subjective? Why doubt it? Where did this view come from?

The most common uses of "subjective" and "objective" are both epistemic and evaluative. "Subjective," always used pejoratively, indicates a lack of adequate justification or of representativeness. Saying that a belief is subjective is a way of critically implying a lack of good, justifying reasons.

When applied to emotions, this pejorative label implies that emotions are based on false or unjustified beliefs, or that viewing the world emotionally is a biased or myopic way of seeing things, or that emotions have no cognitive content at all. Thus we should take emotional people's judgments with a grain of salt, and in pursuing objective knowledge, we should purge ourselves of the biasing influence of emotion. Labeling emotions "subjective" works to undercut their epistemic significance and, given the premium we place on knowledge, their significance period.

That emotions should so frequently be accused of epistemic subjectivity is remarkable. While emotions sometimes are subjective in this sense, they are not obviously always or even typically epistemically flawed. That emotions always are epistemically flawed is simply false. Emotions do not always rest on false beliefs, nor do they always imbue objects with unreal qualities. Jealousy may be, and often is, well founded on a recognition of real threats to a relationship. Moreover, the distinction between appropriate and inappropriate emotions requires that at least some emotions are epistemically objective. If talk about appropriate versus inappropriate emotions is to be meaningful, not all emotions can be Sartrean ones that transform the world into what it is not.<sup>3</sup>

Nor do all emotions myopically bias our perceptions. The kind of selective attention that typifies emotion can be epistemically fruitful. Jealousy, rather than detrimentally biasing perception, may appropriately draw attention to evidence of a decaying relation that otherwise might have been overlooked. Moreover, we are selectively attentive whenever we adopt a point of view, whether it is emotional, moral, prudential, or scientific. Rigidly adhering to any point of view, emotional or nonemotional, may incur judgment errors; but simply adopting a point of view and being selectively attentive does not entail epistemic subjectivity.

Similarly unreasonable is the charge that emotions are subjective because they create a world of, say, frightening or adorable objects, or heroes and heroines. Any interpretive reconstruction of the world—for example, positing the existence of quarks—constitutes a world. Avoiding subjectivity only requires having good reasons for the interpretation.

Even so, the belief that emotions typically, if not always, produce epistemic subjectivity while "acts of reason" do not is commonplace both in and outside philosophy. Whether this belief is true is an empirical question that could be settled only by comparing the percentage of inappropriate emotions to that of unjustified beliefs and inferences. In lieu of this, I want to suggest some reasons for doubting the relatively greater subjectivity of emotions than "acts of reason." Often the plausibility of this view gets purchased with biased examples. Throughout the seventeenth, eighteenth, and nineteenth centuries, strong emotions (rage, sexual love) and strong variants of emotions that take degrees (pride, shame, joy) became paradigms for theorizing about emotion. And strong emotions (both in terms of

physiological disturbance and motivational drive) are precisely the sorts of emotional responses most likely to exhibit epistemic subjectivity. Thus, taking these rather than milder emotions (embarrassment, affection, irritation) to be paradigmatic made it unnaturally easy, indeed almost a matter of definition, to argue for emotion's epistemic subjectivity. Conversely, those beliefs most liable to falsity—prejudices, cultural taboos, dogmas, convictions produced through brainwashing or hallucinogens—never appeared as paradigms of belief. The comparison of emotion and nonemotional belief was thus biased from the outset. Furthermore, the absence until relatively recently of any cognitive theory of emotion, together with an emphasis on emotion's noncognitive dimensions (feelings, physiological causes, and symptoms) further undergirded the belief in emotion's epistemic subjectivity.

A variety of cultural factors also contributes to the conviction that emotions and "acts of reason" can be sharply distinguished. The association of emotion with women reinforces the predisposition to notice the irrational in both. And in language, metaphors for emotional irrationality abound: "insane with jealousy," "blind rage," "love is blind." The availability of such language legitimates both the belief that strong emotions are paradigmatic and the belief that emotions typically are epistemically flawed. The metaphors available for describing objectivity—"level-headed," "cool," "calm"—have a similar effect by excluding emotions. Such metaphors may be accurate, but their accuracy is difficult to assess when language is already biased against the epistemic objectivity of emotion. Finally, the polarization of the public and personal, and the location of genuine knowledge in the public realm, further divorces emotion from true belief. As I shall argue shortly, emotions are rooted in our personal lives. Thus when genuine knowledge is presumed to be impersonal, emotions can be at best irrelevant to the pursuit of truth, and at worst, obstructive.

In sum, the claim that emotions are epistemically subjective is surely false if this means that emotions always rest on false beliefs or always distort perception. That emotions typically do may reflect philosophical and cultural biases more than the truth.

#### EMOTIONS, BIOGRAPHY, AND KNOWLEDGE

Perhaps you find this conclusion unsettling. You see already its implications. If emotions mirror the world as reliably as other faculties, then we should begin trusting our emotions to deliver truth in the same way we trust our perceptions or our chains of reasoning or the voice of experts. Perhaps we should begin writing philosophy from the gut. But there is a lingering sense that emotions are not trustworthy, are not deliverers of the kind of truth pursued in academia and science. One might also rebel against the implicit overrationalization of emotions that this conclusion involves. Emotions become merely another vehicle for knowledge, and this grates against our

sense of the intimacy of emotional response, the personal investment and revelation of self in emotions. How can emotions be both personal and sources of knowledge? That is, how can they be both biographically subjective and epistemically objective?

The Enlightenment answer was “They cannot.” From Descartes, through the heated controversy between rationalist and moral sense theorists over the relevance of emotion to morality, and up to Kant, one of the most often repeated observations was that emotional responses lack uniformity. Both emotional temperaments (our proclivity for feelings of one emotion-type rather than another) and emotional reactions to any one situation vary between individuals. A good bit of energy went into explaining this variance. For Descartes, it had to do with differing physical constitutions and strengths of will; for Hutcheson, it was in part due to our educators’ having been strict or lax in cultivating nobler feelings in us; and for Kant, it was the inevitable consequence of natural, idiosyncratic liabilities to be pleased or pained by different things. Though disagreeing in their explanations, they all agreed that not only do emotional responses vary from individual to individual, but also that this variance is, within limits, normal. No one assumed we could or should lead identical emotional lives. But they did when it came to our lives as knowers. Knowledge meant theoretical knowledge, particularly in science, mathematics, logic, and ethics. And by definition, anything that was to count as knowledge had to be the sort of claim that was accessible to and could command assent from individuals, irrespective of the particularities of their lives. Given this understanding of knowledge, the successful pursuit of knowledge and the achievement of epistemic objectivity seemed to mean adopting a fully impersonal point of view. The cognitive processes going into moral, mathematical, logical, and scientific knowledge had to have an independence from the individuality of individuals. The fact that emotions are idiosyncratic, personal, and unalterably relative to the peculiar fabric of our lives—what I have been calling their biographical subjectivity—rendered emotions unfit vehicles for knowledge and necessarily epistemically subjective.

What gets set up, then, in the Enlightenment period is a kind of one-to-one correspondence between epistemic objectivity—that is, taking up a viewpoint that will reliably deliver truth—and biographical objectivity—that is, taking up an impersonal, impartial, disinterested viewpoint. There are two different ways one might go about breaking this correspondence and bringing emotions to epistemic respectability. One is to challenge the assumption that it is possible to achieve the kind of complete impartiality, disinterestedness, and impersonality that was the apple of the Enlightenment’s eye. All knowledge seems at the very least to be historically and culturally conditioned. It also seems likely now that what claims are accessible to discovery and understanding and what claims command assent will be partially a function of our personality structure, social location, and life

experiences. The pursuit of knowledge is always to some extent personal (we are already writing philosophy from the gut). If it makes sense to continue talking about knowledge under these circumstances, we will need a revised notion of epistemic objectivity in which truth might reliably be delivered from an at least partially personal point of view. That means emotions could plausibly be used as a resource for knowledge.<sup>4</sup>

I do not want to push this line, although I think it is basically right. The reason for not doing so is that the personal point of view comes in degrees. Even if we cannot achieve the Enlightenment ideal of disinterestedness and impersonality, we can go a long way in trying to bracket our own commitments, interests, and conceptual framework and to sympathize with others'. The pursuit of knowledge in academia and science still requires this modified form of biographical objectivity. And emotions, because they emerge largely out of our unbracketed commitments, interests, and conceptual frameworks, will still be vulnerable to the charge of epistemic subjectivity. They are just too personal to reliably deliver that kind of knowledge.

A more promising approach might be to challenge the identification of the highest form of knowledge, that most worth pursuing, with theoretical knowledge—that is, with the kinds of claims that are most abstracted from the realities and exigencies of everyday life. In ethics, for example, moral knowledge has become equated with the elaboration of highly formalized and stylized decision procedures, with extremely general “first principles,” and with manipulating the logical implications of hypothetical cases whose occurrence in daily life is either impossible or improbable. The result is the production of a kind of knowledge that has no valence, that does not speak to human needs, fears, and aspirations, and thus cannot attract the assent of the biographical self, but only a kind of purely logical assent. The further result may be, as Annette Baier claims, that the morality endorsed by many moral theorists “is seriously endorsed only in their studies, not in the moral education they give their children, nor in their reflective attitude to their own past moral education, nor even in their attitude to how they teach their own courses in moral philosophy” (1986, 541). The pursuit of a kind of knowledge that, because it is abstracted from daily life, cannot engage the biographical self may produce an epistemic schizophrenia, where what we know as theorizers has practically no point of contact with what we know as practically engaged actors. Given that putting as much distance as possible between our knowing lives and our everyday lives seems neither warranted nor healthy, there is no strong reason for presuming that knowledge will not (must not) be emotionally charged.

From an everyday point of view, the knowledge that matters most to us, that gets used the most, and that comes most readily to mind is biographically loaded. That is, in everyday life we do not adopt the viewpoint of the impartial knower pursuing truth. Mere truth is insufficient in daily

life to make knowledge either interesting or worth pursuing. Rather, we ask of knowledge that is worth pursuing that it be relevant to our way of life. The result is that our patterns of knowledge and ignorance tend to reflect our personal past or present biographies. So we can ask conversationally, "How do you come to know so much about that?" and expect to get a personally revealing answer ("I know the history of Panama because I grew up in the Canal Zone"). We also expect a match between what individuals know most about and where their emotional energies are invested.

My point here is that it is only from a peculiar, and questionably desirable, academic point of view that epistemic objectivity appears to require biographical objectivity. If our ideal is of a knowledge purified of its connection to the knower's daily life, then getting the truth will of necessity mean adopting an anonymous, impersonal point of view, and emotions will have no place in the pursuit of knowledge. If, on the other hand, our ideal of knowledge is that it be relevant to living some kind of life, then getting the truth will be compatible with taking a biographically subjective point of view. And we might reasonably expect emotions to come into play in the pursuit of knowledge.<sup>5</sup>

#### BIOGRAPHICAL SUBJECTIVITY

So far, I have referred to "biographical subjectivity" as though it were perfectly obvious what it means. It is not. I hope "biographical" evokes the image of a subject about whom we tell stories. We are, after all, individuals with histories, growing up in Panama or New York, born into middle-class or lower-class families, suffering childhood traumas, becoming paranoid or generous, learning philosophy, learning to be careful, disadvantaged or advantaged, suffering loss, and sometimes getting lucky. This is the subject who has a past, a psychological profile, and allotted or elected social roles. Biographical subjectivity is the viewpoint of that subject.

More specifically, I have in mind a personal point of view from which an individual's own conceptual framework and evaluative system come into play in her perceptions, judgments, and emotional responses.<sup>6</sup> Because we have such widely varied biographies, we are exposed and disposed to widely varied forms of conceptualizing the world. That means not only differences in the terms through which we think, but also differences in what we notice, remember, and forget, or in what gets conceptually elaborated or dropped. By the same token, the differences in our biographies produce differences in evaluation so that we find significance or triviality, urgency or indifference in different places. Here is one story.

Tess is possessive. She is not selfish, but she takes pride in her own belongings. She is one of those people for whom ownership means a lot. In more relaxed moments, rather than daydreaming or reading a novel, she simply sits and lets her gaze rove over the cozy arrangement of furni-

ture, the worn science fiction novels neatly arrayed on homemade bookshelves, the pictures acquired over the years that hang in carefully selected spots. Often you find Tess refinishing or rearranging furniture, making or buying new touches for her home. She is the kind of person who gets irritable when an ill-hung gate goes unrepaired. She is, you might say, finely tuned to her property.

Tess thinks in terms of decorative schemes, craft knowledge, ownership and care of private property, and homes. Her patterns of perceptual attention embody this style of thought. She notices that ill-hung gate, though others might not. And the objects of perceptual attention fall under descriptions compatible with her conceptual framework. She sees the gate under the description "needing to be repaired."

What is important to note is that even though Tess adopts a personal point of view, there is nothing necessarily epistemically subjective in doing so. Although a limited and limiting framework, so too is any conceptual framework. The scientific and artistic points of view, for example, similarly restrict the range of perceptual salience and intentional descriptions. What the scientist, the artist, and Tess observe differs partly because they find value, importance, and urgency in different places, and partly because they employ different conceptual frameworks. An astronomer notes a star's size, color, and position: "It is a red giant in the Andromeda nebula." An artist, perhaps a Van Gogh, heeds how fiercely the stars whirl over black and lonely trees. And Tess notes with satisfaction how clearly visible they are from her backyard. What inclines us to call Tess's framework personal but the scientist's public is that the scientist's conceptual framework as a whole is shared among scientists, while Tess's is unlikely to be repeated. But the components of Tess's framework are shared. Her style of thinking draws off of and amalgamates several shared conceptual frameworks, particularly those of the homemaker and the capitalist.

To say that an observation, interpretation, judgment, or moral appraisal is biographically subjective is to say both that that individual is disposed to make just that observation and that it expresses or reveals her personal identity. A friend and I, for instance, watch *Gone with the Wind*. I remark that Scarlett O'Hara is a tragic heroine who breaks conventional moral standards, risking social ostracism and loneliness to enable her family to survive in the post-Civil War world. She is admirable. My friend disagrees, saying that she is selfish, scheming, and devoid of moral feelings of sympathy, gratitude, and guilt. Not coincidentally, our different interpretations are psychologically explicable. I identify with Scarlett, seeing myself as a survivor who can do what has to be done. He firmly believes that women should conform to the model of femininity embodied by Melanie. Our interpretations are biographically subjective, revealing who we are; the film, arguably, equally supports both.

This is not to say that all judgments, observations, and interpretations are biographically subjective. We may perform Freudian analyses for fun, note a car wreck because it is tragic, or acquire bits of trivia without having any special disposition to do so. We may also deliberately bracket our own personal point of view in order to take up a generally shared conceptual framework. Nor is it to say that we can be biographically subjective only from our own point of view. The skillful novelist enlivens her characters by imaginatively constructing their biographical subjectivities, and sympathy and compassion often require the capacity to adopt another's personal point of view.

While I have been arguing for the independence of the notion of biographical subjectivity from that of epistemic subjectivity, there is one feature of biographical subjectivity that may sorely tempt us to equate the two. When eighteenth-century British moralists talked about self-interestedness, they had in mind the evaluative bias that often creeps into biographically subjective judgments and perceptions. We cannot duplicate for strangers the concern we feel for our friends' and relatives' welfare. We notice our friends' virtues, but our enemies' flaws. And famine at home seems more important than famine abroad. One might also note that some very odd things become objects of inordinately high personal estimation: for Tess, her worn science fiction books; for others, perhaps a high school ring, a pressed flower, a VW Beetle. In short, the evaluative system structuring a biographically subjective viewpoint often seems misaligned with the impersonal worth of things and people.

What I want to suggest, though briefly, is that such idiosyncratic, personal evaluations can be epistemically objective, but we justify them differently. Think for example of a farmer who has spent years building his tomato business, sweating over mortgage payments, suffering through droughts, working with his wife as a partner, and teaching his children the business. His personal evaluation may be that no price would be high enough for his land. Yet he may also accept its low market value. Two different standards of evaluation are coming into play here. On the one hand, there are impersonal considerations for setting the value of land to people in general: low public demand for small farms, flooding of the market, isolation from urban centers. On the other hand, there are personal considerations having to do with preserving his life's integrity: he has long-range plans to remain in agribusiness and the farm is the locus of memories, projects, and relations with others that are central to his self-identity. These biographical considerations make it reasonable to set the value of the farm to himself very high.

While getting impersonal evaluations right may require bracketing our own lives, the only way to get personal evaluations right is to adopt a biographically subjective viewpoint. We may still tragically misjudge what

really matters to our personal identity. But this will be because we are self-deceptive or insufficiently self-reflective, not because of the subjectivity of the viewpoint.

#### EMOTIONAL SUBJECTIVITY

The point I have been driving toward is this: If it is the connection with personal biography that makes emotions subjective, then the line between emotion and reason will have to be drawn very thinly. I have argued that, in what we call our rational life, there is a fit between style of thinking, pattern of knowledge and ignorance, and motivation to pursue particular fields of knowledge on the one hand and personal biography on the other. Reason, too, bears traces of the personal. The result is multiple connections between cognition and emotion. First, many beliefs are emotionally charged. The belief that women deserve equal pay, for example, may well be more emotionally charged for women than for men. Second, one's personal conceptual framework tailors not only the contours of one's cognitive life, but also one's emotional repertoire and temperament. The centrality of moral notions in a conceptual framework, for example, will produce simultaneously a cognitive disposition to make moral judgments and an emotional disposition to experience guilt and moral indignation. Whatever line can still be drawn between reason and emotion turns out to be an epistemologically uninteresting one. It amounts to no more than this: Some of our beliefs and cognitive processes have either no or only a very weak connection with our personal biographies, and hence are relatively independent of our emotional life; others have a stronger connection and are emotionally charged.

Once we stop looking for emotion's irrationality and begin looking for its biographical meaningfulness, we get a deeper analysis of the nature of emotion. In the following subsections, I want to explore what happens to the analysis of emotion when emotional subjectivity is understood biographically.

#### Emotion-Statements and Belief-Statements

According to cognitive theories of emotion, which I take to be basically right, emotions logically or quasi-logically presuppose beliefs that both define the emotion and, if true, justify it. Love involves thinking highly of the beloved; embarrassment presupposes the belief that one has committed a public faux pas; jealousy requires a belief in one's entitlement to attention from another, and so on. One reason for logically connecting emotion and belief is, as Errol Bedford (1956–57) argued, that emotion statements elliptically express categorical evaluations. "I regret having skipped fourth grade" elliptically expresses something like "Skipping fourth grade was det-

perimental." This expressive function might explain why we pepper our narratives with emotion-statements, but it does not explain why we sometimes choose emotion- over belief-statements. For if emotion-statements are merely ellipses, it would seem reasonable always to eschew emotion-statements in favor of straight talk. Yet talking about our emotions seems to convey something that talk about our beliefs does not. Moreover, using emotion language is a familiar, convenient strategy for avoiding arguments. "I feel . . ." removes evaluations at least partially from the public domain of appraisal and criticism, but "I believe . . ." and "It is true that . . ." normally make them fair game for critical attack. In short, emotion-statements seem to be more than mere ellipses. The "more" comes from emotion's biographical subjectivity. Emotions presuppose, and thus emotion-statements express, evaluations made from a biographically subjective viewpoint. "I resent sexual injustice" locates sexual injustice in the speaker's life as something that matters to her and gets lived against, while "Sexual injustice is undesirable" dislocates it from the speaker's biography. Sometimes choosing "elliptical" emotion-statements over belief-statements, then, makes sense. "I resent," "I regret," and "I feel" linguistically specify the personal point of view and the connection between belief and biography. That using emotion-language partially immunizes evaluations from critical attack also makes sense. We are in a privileged, though not incorrigible, position to know what is good or bad, detrimental or helpful, in relation to the structure of our lives. By contrast, belief-language places evaluations in the domain of public criticism, because it expresses biographically objective (impersonal) evaluations; and we are not in a similarly privileged position to know what is of value or disvalue to the structure of people's lives in general.

## Teleology

Philosophers have often remarked on the teleology of emotion. Descartes claimed that passions "dispose the soul to desire those things which nature tells us are of use" (1989, art. 52). Robert Solomon (1977) argues that emotions are oriented toward maximizing self-esteem. And quite a few philosophers, including Dewey, Ryle, Sartre, and Solomon, have advanced "obstacle theories" of emotion: emotions respond to and emotional behavior and imagery try to sidestep obstacles thwarting our intentions.<sup>7</sup> (Sartre's famous sour grapes story has the fox solving the problem of unreachable grapes by resentfully transforming them into imagined green ones.) Understanding emotional subjectivity biographically would explain emotion's teleology. Emotion-founding beliefs, on this reading, always have a strong connection with personal biography. Thus a belief in the desirability of nuclear disarmament that is only part of a "purely academic" stance could not lead to disappointment over failing arms negotiations, but connected, say, to a

commitment to antinuclear organizations, it could. If emotional beliefs are biographically loaded in this way, then it would be no surprise if emotions typically occur in contexts where our self-esteem is at stake, or where some obstacle impedes our projects, or, in the case of positive emotions (joy, pride, love, euphoria), where the road to carrying through our projects is unusually clear.

### Appropriate Emotions

Appropriate emotions are ones that are epistemically objective; there is some sense in which they fit the facts. Ronald de Sousa (1980) gives an especially useful and sociologically sensitive analysis of what this “fitting the facts” might come to. We learn, identify, and judge the appropriateness of emotions, he argues, by reference to paradigm scenarios. Crystallized in literature, film, adages, and metaphors, paradigm scenarios specify “first a paradigm situation providing the characteristic objects of the emotion . . . , and second, a set of characteristic or ‘normal’ responses to the situation” (142). When successfully enculturated, people intuitively recognize and put paradigm scenarios into play even if they cannot always articulate the content of those scenarios. (Thus emotions may not presuppose particular beliefs so much as overall readings of situations in light of emotional paradigm scenarios.) The loosely defined nature of these paradigm scenarios means, for de Sousa, that there are no rigid criteria for emotional appropriateness. At best, we can roughly assess the degree of fit between an emotional paradigm and a given situation. The paradigm scenario for love in *Romeo and Juliet*, for instance, only marginally fits twentieth-century love relations or love between elderly people. Moreover, a wide variety of emotional paradigm scenarios may fit a given context. The scenarios for anger, humility, and fortitude may equally fit a scene of public reprimand. Where there is this flexibility, de Sousa suggests we choose the most appropriate scenario by looking at both the consequences of adopting one paradigm over another (fortitude may preserve relationships, while anger destroys them) and the cultural rankings of paradigms (fortitude is a virtue, anger is not).

What I want to add to this account is the subjective element in appropriate emotional response. While considering the cultural rankings and the consequences of adopting particular paradigms may partly determine emotional response when more than one paradigm fits our situation, our own emotional temperaments also play a role. Consider this story.

Emilio, who was recently promoted up the corporate ladder, wants to prove that he can handle his new responsibilities. He begins working longer than usual hours, attending more job-related social functions. Although his wife, Sarah, complains, Emilio shrugs off her complaints, saying this is just temporary. But months pass with no letup, and Sarah finds herself devel-

oping a close friendship with her riding instructor, whose company substitutes for Emilio's lost attention. Eventually, their social intimacy leads to sexual intimacy.

Jealousy, anger, self-recrimination, and regret all fit Emilio's situation. He may gauge his emotional response by the consequences for his marriage, or his job, of one emotion versus another. Or he may ask what he ought most to feel in a marriage. But these objective considerations will be tempered and molded by his subjective conceptual framework. That is, given the subjectivity of emotion, Emilio's emotional response will likely reflect his own conceptual framework and evaluative system. If his own role in creating events dominates his thinking, he will focus on his failure to respond to Sarah's needs and his responsibility for her seeking a supportive relationship elsewhere. So he will view his situation through the paradigm scenario for self-recrimination. If instead he is prone to fatalistic thinking, noting inescapable causal determinants, he will focus on how his habitual overestimation of job requirements and Sarah's proclivity for close personal relationships have combined to make this situation unavoidable. The scenario for regret will be most salient. Alternatively, if he operates with a moralistic framework where obligations to himself are central, he will focus on Sarah's infidelity and her insensitivity to his needs, viewing his situation through the scenarios for anger and jealousy. Thus although the paradigm scenarios for jealousy, anger, self-recrimination, and regret may equally fit Emilio's situation, he will not be equally disposed to feel all four emotions.

In short, the beliefs founding appropriate emotions are still biographically subjective. Securing the appropriateness of one's emotions does not require taking an objective, impersonal point of view and impartially surveying all the facts. It simply requires that the interpretive story one is subjectively most likely to tell about one's situation be true.

That a person is disposed to tell just that story also enters into our assessments of emotional appropriateness. In daily life, we sometimes sense that even though an emotion is right for the situation, it is wrong for the person; it is out of character or reveals inconsistency of character. Jealousy, anger, self-recrimination, and regret may all be appropriate for people in general to feel in situations like Emilio's, but they may not all be appropriate for Emilio.

In years past, Emilio has had quiet affairs. He avoided guilt by telling himself they were excusable lapses because the women meant nothing to him. But now, discovering Sarah's affair, he furiously accuses her of betraying his trust. She defends herself, saying, "What are you so mad about? You obviously didn't think there was anything wrong with your playing around!"

Sarah's complaint is that, even if she has done something wrong, Emilio should not be pointing it out. His past behavior and lack of guilt indicate a toleration of outside affairs that conflicts with his present obsession with

betrayed trust. Emilio's anger is inappropriate not because it is unfounded, but because the interpretive story fueling his anger conflicts with the interpretive stories he told in other, similar situations. Emotional appropriateness requires not just fit with the facts, but also fit with a consistent, subjective style of thinking.

### Psychological Explanations

Emotions seem to demand psychological explanations in a way that beliefs do not. "Why do you believe that?" asks for justifying reasons, not a psychological profile. "Why do you feel that way?" may also ask for a justification, but giving a psychological explanation instead, or adding it to the justification, would not be odd. By contrast, it seems unreasonable, or at least oddly irrelevant, to throw in psychological explanations after a belief has been justified. Moreover, when people have unusual, inappropriate emotions, the morally correct tack is more often taken by trying to understand the emotion's psychological genesis than by bluntly criticizing its appropriateness. We are not nearly so circumspect about false beliefs. Even if "Your argument for animal rights is ridiculous" is insensitive, "Your compassion for animals is stupid" seems far worse.

Why should we settle for psychological explanations of emotions but not of beliefs? Perhaps it is because people can be held responsible for their beliefs but not their emotions. Descartes thought so. The will has infinite power to assert, deny, or refrain from judgment, but only limited control over instinctive, externally caused emotions. The biographical subjectivity of emotion and belief suggests a different answer. On the one hand, we may have underestimated the relevance of psychological explanations to understanding beliefs, and with that underestimated the insensitivity of critiquing a person's beliefs. If, as I have argued, what occupies our thinking, what we accept as true, and our patterns of knowledge and ignorance are often tied to personal biography, then what we believe cannot be cleanly separated from who we are. At most, only some of our beliefs will not be biographically subjective. Emilio, for example, may acknowledge Sarah's culpability without being prone to think about it. Only in these cases will understanding the psychological etiology of beliefs be irrelevant and will critiquing their falsity leave us as individuals unscathed. On the other hand, I have argued that emotions, even appropriate ones, are always tied to biographically subjective beliefs. Because tied to beliefs we have a psychological disposition to hold, emotions take psychological explanations. So even if Emilio appropriately feels regret, we can still ask why he is psychologically prone to fatalistic thinking about the unavailability of events rather than prone to moralistic thinking about others' culpability. The subjectivity of emotion also means a tight fit between emotion and self. Compassion for animals grows out of a distinctive way of living in the world

that perhaps has its origin in memories of childhood pets, or a failure to socialize well with people, or an adoption of nurturing roles. To criticize emotion, then, is to criticize the person, her memories, her way of life, her style of thinking, and evaluative systems. Such critiques cannot leave us as individuals unscathed.

#### NOTES

This essay was originally published in 1989 in *The Philosophical Forum*. I have made minor editorial changes to the text.

I thank the National Endowment for the Humanities for financial assistance in writing the first draft of this chapter. Amelie Oksenberg Rorty's "Explaining Emotions," in her anthology *Explaining Emotions* (1980), was the conceptual spark for my account of emotional subjectivity; Jane Tompkins taught me that knowledge can be personal and emotional.

1. See, e.g., Searle 1983's description of emotion.
2. Especially helpful to me, though this is by no means an exhaustive list, were M. Jaggard 1989, Lloyd 1985, and Keller 1985.
3. "All emotions have this in common, that they make the same world appear . . . in which the relationship of things to consciousness is always and exclusively magical" (Sartre 1948, 80).
4. Jaggard 1989 develops this conception of emotions as an epistemic resource.
5. I am indebted to Jane Tompkins for pointing out the gulf between academic theorizing and what matters in our daily life. See Tompkins 1987–88.
6. What I am calling a "conceptual framework" here is virtually the same as the notion of "cognitive set" that I develop in Calhoun 1984. Although that earlier piece clarifies what I have in mind by a conceptual framework, the piece exaggerates the distinction between our emotional life and our rational life by refusing to notice that our beliefs, too, are shaped by our personal style of thinking and by presuming that in emotion-belief conflicts it must be one's emotions, not one's beliefs, that are irrational. I am now inclined to think that Tess should have trusted her resentment and used it as a resource for knowledge, rather than discrediting her emotions.
7. See Dewey 1950, Sartre 1948, and Solomon 1993.

*This page intentionally left blank*

# IV

## Emotions and Rationality

*This page intentionally left blank*

## Emotions, Rationality, and Mind/Body

PATRICIA GREENSPAN

This chapter attempts to connect recent cross-disciplinary treatments of the cognitive or rational significance of emotions with work in contemporary philosophy identifying an evaluative propositional content of emotions. An emphasis on the perspectival nature of emotional evaluations allows for a notion of emotional rationality that does not seem to be available on alternative accounts. Though naturally expressed in mentalistic terms, the perspectival account does not imply dualism. What does the evaluating, or has the propositional content, is a state of positive or negative affect with a physiological basis, including connections with cortical brain states in typical cases of human emotions. What the account adds to this is the element of normative interpretation, as needed to give emotions a serious role in practical reasoning.

In the bookstores I frequent there are now quite a number of popular or semipopular works urging rejection of the old opposition between rationality and emotion. They present evidence or theoretical arguments that favor a reconception of emotions as providing an indispensable basis for practical rationality. Perhaps the most influential is neurologist Antonio Damasio's *Descartes' Error* (1994), which argues from cases of brain lesion and other neurological causes of emotional deficit that some sort of emotional "marking"—of memories of the outcomes of our choices with anxiety—is needed to support learning from experience.

Damasio's work has an interesting connection to such issues as how to understand psychopaths, agents who lack normal feelings of guilt and other moral motives based on empathy.<sup>1</sup> It seems that psychopaths are not like the rational "amoralists" of philosophic lore but rather are unable to

follow through reliably on long-term plans they make in their own interests. A failure of emotional empathy—with one's own future self, in effect—apparently yields elements of practical irrationality.

On the other hand, Damasio wrongly sets up Descartes and mind/body dualism as a philosophic foil for his view.<sup>2</sup> His real target seems to be computationalism and similar views in cognitive science (“the mind as software program”).<sup>3</sup> He even implicitly recognizes at one point toward the end of the book that his announced target, Descartes's *cogito*, does include emotions, or at any rate their mental aspect (“suffering”), and he cites Descartes's detailed account of emotions in *The Passions of the Soul*.<sup>4</sup> But Descartes's explanation of emotions in that work in terms of “animal spirits” (essentially an outdated predecessor of neurological impulses) seems to bridge body and mind (or soul), despite his official dualism. The titles of both books—Damasio's and Descartes's—may be somewhat unfortunate.

More generally, the recent neuroscientific work on emotions seems to take all but neurophilosophy and similar approaches within philosophy as necessarily opposing the project of recognizing the cognitive or rational role of emotion. In a rough-and-ready way, emotions are assumed to fall on the “body” side of the “mind/body” distinction for anyone who would allow that much talk in mentalistic terms.

There are other recent popular works dealing with evolutionary psychology and related subjects that do make use of some philosophic literature for insight into the moral role of emotion. These essentially follow Darwin's attempts to explain the development of the “moral sense” in terms of social emotions in animals.<sup>5</sup> A particular focus is eighteenth-century British moral philosophy, with its attempts to base ethics on human emotional nature. Sometimes the approach is put to conservative political uses by “sociobiologists” and others, and sometimes it is dismissed on just those grounds by political opponents, especially feminists. A current popular book that attempts something less ideological (though still committed to a basis in the mind's innate structure, on a version of the view derived from Chomsky) is Stephen Pinker's *How the Mind Works* (1997).<sup>6</sup>

One thing many of these discussions seem to have in common is an importance assigned to emotions in rational terms specifically for resistance to full rational control. Emotions are treated as cases of “rational irrationality”: they are of use to us rationally, in promoting our long-term ends, in part because they function as barriers to rational deliberation.<sup>7</sup> They protect us from the need or the tendency to reason things out from scratch at every stage or in every respect, often to the detriment of rapid response or reliable follow-through or the ability to form relationships of mutual trust. In social terms, they serve as “commitment devices,” making it demonstrably difficult for us to act as we otherwise would on the basis of narrow self-interest.<sup>8</sup> The extreme case of uncontrolled anger, for instance, communicates a “hell-bent” retaliatory urge in a way analogous to throwing the

steering wheel out the car window in a game of chicken—Schelling’s classic game-theoretic case, with two cars hurtling toward each other, about to crash unless one of them swerves.<sup>9</sup>

Just because emotions are somewhat recalcitrant to reason, then, they are accorded a crucial role in rational design—in creating a human nature (or a range of human natures) that is up to the human task—whether the design in question is evolutionary, cultural/political, or pedagogical. There are other recent works from a psychological or psychiatric perspective marshaling evidence for the role of emotional development in early childhood as a foundation for normal cognitive learning.<sup>10</sup> A further area of application to individual cases is psychotherapeutic redesign: there is a huge collection of psychological self-help and related literature (some of it theoretically respectable) dealing with emotions in rational terms. For that matter, the general line of thought here fits easily with self-developmental approaches within philosophy stressing Aristotelian notions of character building or habituation in virtue.

However, there is still a kind of disconnection from contemporary philosophy (outside cognitive science) in the treatment given in many of these works to understanding the nature of emotions, what emotions are. From within philosophy Paul Griffiths’s recent book *What Emotions Really Are* (1997) exploits and widens this gap to the extent that its evolutionary approach to emotions caricatures and dismisses the alternative, “propositional attitudes” approach and sweeps aside one of the main questions the latter attempts to deal with, essentially a normative version of Damasio’s question of the bearing of emotions on practical reasoning.<sup>11</sup> It is what philosophy has to teach on that subject that needs to be brought out in response to Damasio’s error (and also, I think, Griffiths’s) of confusing ongoing attempts to understand emotions in mentalistic terms with a certain competing research program in cognitive science.

The effect of much contemporary philosophy of emotion has been to identify a rational or potentially rational (rationally assessable) content of emotions, at any rate in typical cases of human response.<sup>12</sup> Emotions can be viewed as having a content expressible in terms of what they “say” about their objects: personal anger, for instance, registers the agent’s perception of a wrong someone presumably has done; pride registers the thought that the agent is somehow praiseworthy.

This approach is generally discussed under the heading of “emotions and judgments,” since it emerged from debate over more extreme versions that simply equated emotions with a subclass of evaluative judgments or beliefs, as the category of propositional attitudes that philosophers were most at home with. But more fundamentally, what is at issue is a view of emotions as registering evaluative information and thus as susceptible to some sort of rational assessment themselves—not automatically to be consigned to the “irrational” category.

There are overlapping theories in psychology that understand emotions in terms of cognitive “appraisals” and similar notions.<sup>13</sup> However, much of the current work in “harder” areas of science (including “cognitive science” areas of philosophy) eschews such talk in favor of a treatment of emotional states as physiological or bodily reactions capable of causal connection with rational thought and action but not themselves capable of rationality.

There may be good heuristic or other practical reasons for adopting this nonmentalistic framework for certain purposes. Within philosophy, the aim is often to lend support to an interdisciplinary scientific research program. Minimizing metaphysical assumptions avoids a lot of potentially divisive dispute. But of course it is the job of philosophy to inquire into assumptions. In this case, the point of doing so need not be particularly metaphysical—to push beyond the categories studiable by science—but rather, as I see it, is more concerned with specifying just how it is that practical reason puts emotions to work. The essential terminology is that of normative assessment rather than mentalistic talk *per se*.

Philosophers have exploited the possibility of representing the evaluative content of emotions in propositional terms since Aristotle, though not always with a distinction between content and causal accompaniments. The Stoics even made out emotions as evaluative judgments. However, unlike Aristotle, they also advocated an asceticism that affects their treatment of emotions. Emotions for the Stoics amount to confused judgments, and their advice is to minimize confusion by cultivating more detached states of mind.

With all their many possibilities of confusion, however, emotions may sometimes embody more accurate perception of the value-laden world than we allow to affect our detached judgments. Our regard for them as quick responses resisting deliberative control is heightened by this assessment: they are not just “quick and dirty” (rationally speaking) but often embody a point of view worth recording even where more reasoned judgments are to hand. There has been a resurgence in recent years of a “judgmentalist” approach with a more positive spin on the value of emotion.<sup>14</sup>

I would modify the approach with an account of emotional rationality that sets it apart from the logic of judgments by allowing for rational options, including conflicting emotional responses by the same person to the same situation—or for that matter, the suppression of emotional response.<sup>15</sup> I refer to this as the “perspectival” account, meaning that rational warrant for an emotional response varies with evaluative perspective in a way not recorded in qualifications to the content of emotion, unlike what is supposed to be the case for judgment.<sup>16</sup>

On the perspectival account, what emotions register, when the mechanism is working properly, is not necessarily the “all things considered” view of things by which we assess our beliefs. To say that an emotion is reasonable, or rationally appropriate, is to say that a certain evaluative

belief that represents the content of the emotion (for anger, for instance, that someone has done me a wrong) would be warranted by a significant subset of the evidence—significant in the sense of “worth holding in mind,” perhaps for moral or other practical purposes.

This is a loose and variable standard, presupposing a further level of normative assessment (*of* the evaluative thought content of emotion as well as *by* it). Rationality in the relevant sense allows for emotional options and even emotional conflict or ambivalence. It does not imply the *irrationality* of an emotion with the opposite content—or of no emotion, emotional suppression or indifference. As a positive evidential assessment of an emotion, “rational” means something like “rationally acceptable,” or adequately grounded in the situational evidence, rather than “rationally required,” or mandated by the evidence, as on the usual standard for assessing belief. What is assessed in the case of emotional evaluation is something more like attention to a *prima facie* belief—holding a certain thought content in mind—as distinct (in some cases) from all-things-considered assent.

Imagine letting yourself get angry about a consumer complaint for the sake of arguing more forcefully with the store.<sup>17</sup> The propositional content of anger here would be something like “The store has dealt with me unfairly.” But I could think this—in the dispositional sense relevant to belief, or even as an object of occurrent attention—without necessarily reacting to it with characteristic phenomenological symptoms of anger. I may have reasons for “letting it go” until I have more time, say, or out of sympathy for an overburdened clerk. On the other side, I have reasons for “letting it happen,” setting up the conditions under which my anger will emerge (for instance, by reviewing the history of my interactions with the store), in order to get some action from the clerk. So I have options here for emotional reaction—appropriate reaction, as assessed in rational terms, relative to the evidence, for a more tolerant analogue of the notion of warranted belief. Either allowing or suppressing the feeling will be appropriate, in the sense of being adequately warranted by the facts of the situation.

To make more detailed sense of the account, we need to distinguish the essentially cognitive notion of emotional rationality as appropriateness—evidential or representational rationality—from strategic or instrumental rationality, the practical notion that I refer to as “adaptiveness.” Adaptiveness would include, say, a straightforward appeal to the usefulness of feelings of anger—in my example, in getting the clerk to yield—whether or not there is a real basis for the reaction. There are two senses of rationality in play in these cases and elsewhere, and they can sometimes come apart. However, I think the strategic notion (adaptiveness) does play a background role in determining the standard of evidence applicable to a given emotion, in contrast to warrant for belief.

That is, how much we demand in the way of evidential backing for an emotion is adjusted to reflect the usual value of its consequences, both for

the individual himself and for people generally. Less evidence is required, for instance, for anger seen as a healthy form of self-assertion with ameliorative effects in the long run than if we interpret it simply as arrogant and destructive. But we can still make a distinction between rational appropriateness as a kind of evidential warrant and social or moral appropriateness, the assessment of a response such as anger simply as fitting or failing to fit social or moral norms in the particular case. One might reject anger on religious grounds, say, and still recognize that there are grounds for it in a particular case as opposed to others. There is a distinction, for instance, between appropriate emotion and emotion that is normal and understandable but not really warranted, on the order of blaming the messenger of bad news.

In general, the perspectival view is able to make sense of the rational validity of conflicting reactive standpoints as well as our ability to shift perspectives in a way that allows for the combination of emotional uncontrol with a degree of strategy.<sup>18</sup> The view appeals to a notion of the propositional thought content of emotions, but my own inclinations in philosophy of mind are basically in the naturalist camp, if the term “naturalist” is understood to allow for serious social influence on emotion. Though it is not set up to record the results of scientific inquiry into emotions, I would hope that the view can accommodate them.

Presumably, on a naturalist account, the full-blown cases of human emotion that my own view takes as paradigmatic for purposes of rational assessment would be made out as involving a complex causal relation between cortical brain states and physiological states and events.<sup>19</sup> By the same token, emotion on the view I have outlined in mentalistic terms involves evaluative thought content but also an element of positive or negative affect that can be said to have that content—to be about what the associated thought is about. This “associated” thought need not be present as a distinct occurrence. Rather, I take the affective element of emotion as a propositional attitude, an attitude with a propositional thought as its content. Affect itself essentially evaluates something as in some respect good or bad—good or bad for the organism (to be sought after or avoided), in the most primitive cases.

I think of this affective element in crude terms as comfort or discomfort—discomfort that some wrong has been done me, in the anger example, say. Discomfort here amounts to a representation in affect of the negative aspect of the emotional evaluation. It (or the various physiological feelings the term covers) can be seen as a “marker,” to use Damasio’s terms in *Descartes’ Error*, of practically significant thoughts—in the sense of propositions it is “about,” not necessarily propositional thoughts held in mind in some independent sense. Discomfort also adds a practical or motivational significance of its own, as a bad or aversive state for the agent to be in, that affords it a role in rational decision making.

This is definitely not meant to say that the only important property of the affective element of emotions is its positive or negative aspect.<sup>20</sup> There obviously is much more to feeling than that (or in some cases, such as surprise, possibly less). Other features of affect such as degree of arousal can enter into the description of an emotion as a felt quality, and for that matter its classification as the particular sort of emotion it is. Early arguments for judgmentalism exploited the inadequacy of affect as a basis for distinguishing different emotion types, but it does not follow that affect adds nothing relevant. My own simple categories are set up for the purpose of rational assessment, not to give a full account of the nature of emotions.

I sometimes speak of affect and evaluation as “components” of emotion, but this is meant in an analytical sense, not implying separable parts. The two components (aspects, elements) are internally connected insofar as emotional affect has an evaluation as its content. The assumption of intentionality at this level of basic feeling can sound mysterious, but in principle it is no *more* so than in more familiar cases involving units of language and thought. In fact, I suspect that the historical or evolutionary account would start with feelings assigned “meanings” by their significance for the organism in a sense that includes their role in behavioral response—meanings in a sense that becomes mental only with later cognitive development.<sup>21</sup> Thought content in this sense, even at later stages of development, need not be a separable mental element; it is the content of a feeling.

Even if there is a more ultimate explanation of emotional intentionality in naturalistic terms, I think we need to speak in terms of propositional content in order to address normative questions of rationality. Consider, for instance, a possible alternative approach based on appeal to the causal histories of emotions. This would involve taking an emotion as rationally appropriate on the basis of its occurrence in a situation that resembles in relevant respects a situation originally associated with it, whether in early childhood, as in Ronald de Sousa’s “paradigm scenarios” account of emotional rationality, or in an earlier evolutionary environment; and whether or not the connection is socially mediated, as on “social constructivist” views of emotion.<sup>22</sup>

Consider jealousy, or particularly the anger component of jealousy. Imagine someone who momentarily feels jealous anger when his wife exchanges glances with another male at a party. To use Aristotle’s definition of anger in the *Rhetoric*, he is reacting to an unjustified slight—or at any rate, what he sees as a slight, or as indicating that a slight is imminent.<sup>23</sup> But this is a first-level normative judgment (an evaluation of the situation) that requires interpretation of past events and their natural and conventional meanings—what a glance means or can mean, what legitimate expectations a relationship confers, that a glance involves or might lead to intimacies that violate those expectations—on a level that is unlikely to correspond in any simple way to connections among brain and physiolog-

ical states interpreted just with reference to a descriptively characterized situational context.

On the level of second-order normative assessment (of the emotion), jealousy might sometimes be assessed as inappropriate to the current situation—if the agent really knows, say, that his wife and the recipient of her glance, a colleague in her area, are reacting to a professional faux pas on the part of someone else at the party—even where situational cues naturally give rise to jealousy because of their resemblance to the paradigm scenario. To explain which cues render an emotion appropriate, rather than merely natural or understandable, given that assessments of practical significance may have changed since the paradigm scenario was established, we seem to need at least implicit reference to the notion of a propositional content, as what the emotion still essentially “claims” about the situation.

Even supposing that a feeling is ultimately explainable in biological terms—meaning that its occurrence is thus explainable—those are not the terms in which we assess it, or could assess it, as rational or irrational in the instrumental as well as the representational sense, for purposes of self-regulation and social life. If we got to the stage where we could treat jealousy reliably with drug therapy, say, someone would have to decide whether it should be treated, and she would have to deliberate on the basis of at least some assessments containing further normative elements.

For a full theoretical understanding of much that goes on in human behavior, moreover, we need to be able to recognize cases where an agent uses emotional response for his own purposes, healthy or not. For instance, we can make sense of someone talking himself into feeling jealous on flimsy or imagined grounds just in order to provoke a kind of interaction with his spouse—to exert control, perhaps, or perhaps just an occasion to express and enhance affection. Though the jealous episode may start out as a “pretense” of sorts, some pretenses are self-fulfilling.

An account without propositional attitudes would seem to be unable, then, to capture all the causal histories and strategic aims that are relevant to assessments of emotional rationality. But the standards of appropriate response come to be internal to a given emotion type—anger, for instance, gets set up as a response to some sort of perceived slight—even if its affective element is first found in infancy as a response to something more basic, such as physical restraint. So propositional attitudes also affect the way we identify emotions.

Let me end with a “sound byte,” summing up the view I have tried to defend: Affect evaluates! Emotional affect or feeling is itself evaluative—and the result can be summed up in a proposition. I think we can have it both ways, that is, about emotions as feelings or judgments. My own view emerged from modification of judgmentalism, but I have concluded that it amounts to a version of the feeling view with enough structure to allow for rational assessment of emotions. It does not make out emotions as

“quasi-judgments” or thoughts with hedonic tone but rather as *feelings* with evaluative thought *content*. On its own this content amounts to a “thought” in something like a Fregean sense—not necessarily an occurrent mental event, at any rate apart from feeling, but rather what a feeling registers or conveys. By isolating it for analysis in the form of a proposition, I have tried to show how we can begin to understand the role of emotions in practical reasoning. However, emotions are and remain feelings.

## NOTES

Earlier versions of this essay were presented at a plenary session on emotions at the XIVth Interamerican Philosophy Conference in Puebla, Mexico, in August 1999 and at a conference on “Rationality and Mental Health” of the Association for the Advancement of Philosophy and Psychiatry in May 2000. Along with members of those audiences, I owe thanks, for comments, to Erich Deise, Scott James, Stephen Leighton, and Kathleen Wallace.

1. See Damasio 1994, 178–79. I take these suggestions further in “Responsible Psychopaths” (unpublished paper).

2. See Damasio 1994, 247ff., for an explanation of what he takes Descartes’s error to be.

3. See Damasio 1994, 250, 248.

4. See Descartes 1970, 331–99.

5. See Darwin 1871, esp. chaps. 3–5.

6. See Pinker 1997, esp. chap. 6. For one of the more emotion-based versions of the sociobiological argument, see Wilson 1993.

7. See Parfit 1984, 14.

8. See Frank 1988 and Greenspan 2000.

9. See Schelling 1980, esp. 22ff.

10. See Greenspan 1997.

11. See Griffiths 1997; compare Greenspan 2004.

12. Note the qualifications; for discussion of some less typical or less developed cases, see Deigh 1994. For my own purposes, it is unimportant whether we use the ordinary term “emotion” for such cases, but for simplicity’s sake I shall often omit the qualification that I spell out here. “Emotion” on the resulting use will be a semitechnical term, somewhat tightened up from its ordinary use, in this case for purposes of rational analysis—without any claim that it represents a “natural kind” as thus construed.

13. See, e.g., Arnold 1969 and Frijda 1986.

14. For the reentry of what I am calling “judgmentalism” into the post-Wittgensteinian Anglo-American literature, see Bedford 1962. Solomon 1976 expands the view and connects it with continental philosophy, especially Nietzsche and Sartre. A more qualified variant of the view with Jamesian physiological elements appears in Lyons 1980.

15. See Greenspan 1980 and my extended account in Greenspan 1988. For a simpler explication of the current point, see Greenspan 1995, chap. 5.

16. I should note that my term “perspectival” is not meant to evoke Nietz-

sche's general perspectival theory of truth (or for that matter, various other uses of the term that apply specifically to desires and emotions). See esp. the perspectival account that de Sousa opposes to Solomon's view of emotions as "subjectivity" in de Sousa 1987, 146–49.

17. See Greenspan 2000.

18. I should note that Griffiths's argument often seems to take for granted (or even to represent as a product of scientific theorizing) an assumption of "passivity," or emotional uncontrol, that essentially erects a barrier against the recognition of emotional strategies—dismissing them as mere "pretenses" of emotion in cases where the strategy is social or cultural and involves cultivating the sense of uncontrol; see Griffiths 1997, 155–597, 233ff., 242ff., and 9, 16, 118, and 120. On the suggestion that pretenses may be self-fulfilling, see Greenspan 2000.

19. I mean this compatibly with the argument in LeDoux 1996 for a subcortical pathway operating in less full-fledged cases of fear, identified as such by a "freezing" response in both humans and lower animals.

20. See Pugmire 1998, esp. 65ff.

21. Griffiths in some ways creates an opening for this kind of account with his defense of an alternative to propositional "content schemata" in terms of ecological significance for the organism (1997, 231). Compare this with the evolutionary "functionalist" or teleological conception of intentionality defended in Millikan 1984.

22. See de Sousa 1987, esp. 181–84.

23. See Aristotle 1991.

## Some Considerations about Intellectual Desire and Emotions

MICHAEL STOCKER

There is a theoretical argument for the existence and importance of intellectual desires and emotions that some theorists may find agreeable and others may find themselves committed to. It is found in or derives from action theory, as usually deployed in regard to physical/bodily action. The argument concerns the need for what is called *desire*. It goes very simply, perhaps invoking one of Mill's methods. A person can have the very same beliefs in the case of both action and lack of action. Thus, the having of those beliefs cannot explain the action, the doing, nor of course can the beliefs explain the inaction. What is needed for the action and the explanation is something else—something that provides, we may say, the energy. This of course is desire. And of course this is true enough to many of our experiences. We see the cake and do nothing until we also have the desire for it, and then we act to get it. If and to the extent that energy—desire, desirous energy—is needed for the movement, we can similarly argue for such a need of energy in the intellectual realm. And, at least in many cases, this also is true to our experience. Facts or hints of facts may be present to mind, but unless we want to proceed—for example, to see where they lead—we may do nothing with them and proceed no further in those investigations. So too, there are claims—made by Kierkegaard, Heidegger, and others—that without certain forms of care and concern, we would not “parse” the world at all, much less as we do; that without such care and concern, nothing would be salient, intellectually or otherwise.

It may be something of an objection that at times it seems that our thoughts have their own momentum, so to speak. They continue no matter what we want or are at least aware of wanting. So, for example, a well-trained radiologist looking at an X ray may be unable to see only blotches

and shadows, as I do, but will, of necessity, see a mass identified as almost certainly a cancer. An adequate study of this involves several lines of thought. One would lead us to investigate how much control and what sorts of control we have in intellectual goings-on. For it might be held that at least in the higher processes there is or can be the same need for desire as in physical action. We often must drive ourselves to continue, to investigate, to check and recheck. We often must take hold of our thinking and direct it. Perhaps in certain daydreamlike states or reveries, the thoughts proceed on their way, continue on their own; or perhaps here, just because of this last, we hold that the thoughts or the thinking processes are infused with desire. A third line would investigate how an earlier need for desire and care bears on whether there now is desire or care: consider again what the radiologist now sees, and the bearing, if any, on this of the desires the person had in learning radiology.

For present purposes we need not pursue these issues. It is enough if it be recognized that for many intellectual pursuits and activities—ranging from the most specific, eventlike or actionlike, to the most general field of investigation or inquiry—care and concern are present and seem important, even necessary, both in learning how to do those activities and also in doing them once one has learned how to do them.

Limitations of time allow me only to suggest, not show, that a large part of education is evoking, shaping, developing relevant care and concern, and that an important feature of good intellectual work is continuing to have and exercise relevant care and concern. Here I would urge, as an example, that a student who is too dismissive of great writings may be encouraged to approach them with more respect, even with loving attention (as Iris Murdoch suggested the mother-in-law should have to understand and appreciate her daughter-in-law). For a correlative example, I would urge that when a teacher is no longer moved by the texts under investigation, no longer reads them with, say, wonder, it may be time for the teacher to take a break, at least from those texts.

These limitations of time also allow me only to suggest, not show, that the objects and forms of care differ from discipline to discipline, or problem to problem—the care needed for a good philosophical investigation of a text differs from that needed for a good examination of its prose style. And finally, these limitations allow me only to suggest, not show, that as these objects and forms differ, so does the care: in short, that different disciplines and problems involve different desires and different emotions.

In this presentation, I want to make some suggestions about the importance and value of emotions and desires for intellectual activity. In particular, I want to show how emotions and desires can be useful and beneficial for intellectual activity. To start on this, I think it important to suggest why we should reject two views, one about emotions and one about intellectual activity. The one about emotions is that they are somewhat special,

somewhat rare events or eventlike things; that they are momentous, describable as mountains and canyons in our everyday life. On my view, emotions can be diffuse, pervasive, and long lasting, forming our background, as well as the tone, the color, the affective taste, the feel of activities, relations, and experiences. In addition to being diffuse, pervasive, and long lasting, emotions and emotionality are the rule, not the exception.

As the psychoanalyst Ernest Schachtel writes in *Metamorphosis*, “There is no action without affect, to be sure not always an intense, dramatic affect as in an action of impulsive rage, but more usually a total, sometimes quite marked, sometimes very subtle and hardly noticeable mood, which nevertheless constitutes an essential background of every action” (1959, 20). An even more widely ranging claim about action and affects is made by another psychoanalyst, Henry Krystal, in *Integration and Self-Healing: Affect, Trauma, and Alexithymia*: “Affects are familiar to everyone. They are part of our experiences, so ordinary and common that they are equated with being human. . . . Yet their very universality and constant presence with us throughout life make them as unidentifiable as the prose in our speech” (1988, xi). A similar view is also offered by the contemporary neurologist Antonio R. Damasio in *Descartes’ Error* (1994).<sup>1</sup> These theorists’ claims contradict the views that emotions are always momentous, that they are rare, and that they are entirely bodily. By telling against the claim of rarity, they also tell against the view that emotions are always or generally disruptive. That view is plausible only if emotions are rare, or only if ordinary life is ordinarily disrupted.

This helps answer a negative claim that emotions are, as such, disruptive. I am interested in answering that claim. But I am also interested in establishing a contrary, positive claim: that emotions are not just features of everyday, healthy, and good life, but also that they help make everyday life—indeed everyday, intellectual life—healthy and good.

There may seem to be, however, something of a problem here—a problem of presentation, not of truth. If, as these theorists argue, emotions are so very common, arguing for their presence might well seem to be an exercise in banality, like arguing that a fish swims in water or that M. Jourdain in Molière’s *Le bourgeois gentilhomme* can speak in prose. Whether or not it is banal to argue for these near omnipresent facts, it is of some considerable interest to explain why they have been missed or even denied. I will be concerned somewhat with this. But there is another area of my concern, which has to do with distinguishing and discussing different emotions, and more particularly the different ways emotions are found in our lives, the different roles they play—and for this volume, the different roles they play in intellectual activity.

Before discussing some of these roles, let us turn to the theme about intellectual activity I earlier said needs to be rejected. This theme comes in different varieties, with the unifying feature of claiming that just as it is

nonsense to think that numbers or number theory are emotional or can be improved by emotions, so too it is nonsense to think this of intellectual activity taken quite generally.

There is, first of all, the claim that action and acts cannot be intellectual—that is, of the intellect. That what we do when we think, reason, plan, investigate is none of it properly called action. Some who make or are committed to this claim may do so inadvertently. Thus, Harry Frankfurt (1979) starts one of his influential articles on action by saying that the problem of identifying and understanding action is the problem of differentiating between those *bodily* doings or goings-on that are actions (i.e., that are intentional) and those *bodily* doings or goings-on that are not actions but are, say, only reflex reactions.<sup>2</sup> In private discussion, Frankfurt said he had no reason to exclude intellectual actions, and indeed he did not realize that he had done this. It was just that in speaking of actions, he, along with so many other philosophers, just naturally thought of *bodily* actions. (I leave it open whether, had he extended his investigations to intellectual actions, he would have had to make any significant changes in what he said.)

Irving Thalberg (1978), among others, explicitly and intentionally argues that the very idea of intellectual action makes no sense. Further, there are arguments, such as those found in Bernard Williams's writings (e.g., Williams 1985), that would seem to exclude emotions from theoretical reasoning, properly so called, even if room for emotions is left in practical reasoning. The nub of the argument precluding emotions from theoretical reasoning is that such reasoning, even if cast in terms of the first person present, in terms of what a given reasoner does or thinks—"I think . . . , therefore . . ."—is really about what is or is not the case (with appropriate probabilities). So, even if a theoretical reasoner reasons "I think . . . , therefore . . . ," this is not really about the reasoner, nor does the reasoner really figure in the reasoning. The reasoning can, and for clarity's sake should, be understood along the lines of, say, "It is probable that . . . , therefore . . ." Room for emotions disappears with the disappearance of the human subject, the reasoner.

My reply to most of these forms of the argument will be Reid-like: I will show the stone and kick it. But I do want to make two more brief theoretical or conceptual points. First, if Thalberg is right that the very idea of intellectual action is nonsense, then we must wonder, even worry, what the chapters in this volume are and what we were doing preparing them, what we are doing in our professional lives that we get paid and esteemed for, what prizes and reputation are for, and so on. Not all labor is manual or even *bodily* labor.

Second, if Williams is right—which I do not think he is—a bright light is focused on the exhaustiveness of the division between theoretical and practical reasoning. For if he is right that the subject disappears in theo-

retical reasoning, we will need a way, a place, to locate so very much other reasoning which, while not theoretical, does not seem practical. Among the outcast or orphan forms of reasoning, we find much if not most creative work, not only in literature and art, but elsewhere, too. So too, we will be hard pressed to find the proper location for much work in the social sciences, in journalism, and even philosophy, itself. This is to speak of serious, often professional, academic endeavors. We will also have problems in locating reveries, daydreams, fantasies, and much casual thought and talk. In all these areas—at least much of the time—the subject does not disappear. The person's way of doing things, the person's construal of and take on the world are exactly to the point and even central. In saying this, I do not mean to cede theoretical reasoning to Williams's personless and thus emotionless point of view. I mean, rather, to suggest that even if he is right about theoretical reasoning, there is much reasoning that is not practical, where the person remains and where there is at least conceptual room for emotions.

But even if it is agreed that there is intellectual activity and that emotions may figure in intellectual activity, we need to specify and discuss different ways this does or does not, can or cannot, happen. On my view—as developed in *Valuing Emotions* (Stocker 1996)—emotions can figure in life in general in any number of ways. Emotions can be, or help make up, the cause of something we do—my anger can explain my pouting or my slapping you. Emotions can help characterize the sort of person who is doing a certain sort of thing—for example, angry gestures. Emotions can help make the doing of something good or bad for the person doing it—for example, interest can make doing research good, or better, for the researcher; boredom can make it bad. Emotions can be what we aim to have or avoid—for example, we go to a movie to be amused. Emotions can figure in the evaluative grounds and criteria not only of the doing, but also of what is done, the product of the doing—for example, a game with one's child is good insofar as it involves certain emotions; a personal relation is an emotional relation. This last holds that without emotions, certain sorts of activities would not be good. Another role would be as a necessary condition for the activity or for people engaging in the activity. So Kierkegaard and Heidegger hold that without care, concern, and interest, nothing would be salient, indeed the world would have no categories.

Having laid out some roles of emotions, we should ask which of these roles they play in intellectual activity. I am going to start with this question. But I want to warn that just as there are all sorts of emotions and all sorts of roles they can play, so too there are all sorts of intellectual activity. And what is true about emotions in and for one intellectual sort may well not be true for others.

This said, let us now turn to some examples of intellectual desire and emotions—desire and emotions in intellectual activities. Cases could be

made for the differing roles of intellectual courage, intellectual generosity, intellectual liberality, intellectual humility, intellectual etcetera. I will focus on intellectual interest and excitement, which I use as instances of both care and desire. For my part, I do not know of any completely accurate characterization of emotions. But one characterization that is often enough useful has it that they are constituted by affectively laden appreciations or construals, often coupled with desire. So we could note that many instances of intellectual interest and excitement depend on evaluations and are about what is good, intellectually or otherwise. We could not understand Watson and Crick and their inquiries into DNA if we did not see this. And clearly, a central part of education and maturation—both in general and also in particular activities and disciplines—involves learning to be interested in what is relevant, important, useful, and beautiful.

This learning involves developing interest and excitement, as well as the ability to mobilize and direct them. It also involves developing an ability to master them and not to be overwhelmed or disrupted by them. We learn both where to direct our interest and excitement, and also how to modulate them in relation to the value, often the intellectual value, of what does or should interest or excite us as students or practitioners. For just as intellectual interest and excitement are important and often essential for good intellectual activity, unmodulated and undirected interest and excitement can impede or harm, even preclude, good intellectual activity.

It is clear, then, that many instances of intellectual interest and excitement are about what is intellectually good. But it is also clear that many instances of them are *not* about what is good. I am here not thinking about what is morally bad, such as sadism, or aesthetically bad, such as kitsch. These can be intellectually interesting, at least as objects of study.

I am thinking, rather, about interest in what has no interest at all. Here we might consider cases where people have “a bee in their bonnet,” an obsession, an *idée fixe*. Sometimes this is said about something that would be interesting, were it true or even plausible or possible: for example, squaring the circle with only a straight edge and a compass. It is also said about matters that are of no interest at all: for example, someone’s “interest” in how many times the letter *e* occurs in the first edition of Hume’s *Treatise*. But even here, we may be able to find a view, albeit a wildly mistaken one, that would make this interesting. Perhaps the person thinks, on numerological grounds, that Hume is to be believed only if that number is prime. If we cannot find even such a view, we might think that instead of showing interest in that issue, that person’s activity is an expression of, say, obsessional disorder. If this is right, these forms of interest and excitement require a connection—whether plausible or implausible, perhaps only a deformed connection or parody of a connection—between them and what is interesting or what is intellectually or otherwise good.

Nonetheless, there are instances of interest and excitement that do not

have even such a weak connection with what is intellectually or otherwise good. Here we might think of the interest or excitement infants have toward moving lights or mobiles. As I see matters, the closest we can connect this with value and evaluation is by holding that it is an age appropriate and developmentally useful activity or protoactivity that deploys and engages what will be so vital later on in life: it is an awakening of consciousness and a beginning of engagement and activity. And these certainly are good.

Some instances of interest and excitement are still more removed from evaluative thoughts. Here we can consider the excitement of infants and toddlers roughhousing with their parents, being tossed up and caught, or being spun around. So too, we can think of the excitement of dogs when they are playing or hunting. We can also consider the excitement of children and adults during roller coaster rides, parachuting, bungee jumping, and the like.

To be sure, in some of these last cases, there is excitement over mastering what is taken as dangerous—and the notion of danger is evaluative. But there are also instances of such interest and excitement that are not based on value or evaluation—again, for example, the excitement of toddlers being tossed up and caught, and children and adults on roller coaster rides.

I have mentioned various examples of excitement and interest for a number of reasons. One is to help us see that what makes an emotion an intellectual emotion may well not be the category of emotion—here, interest and excitement. The second is to borrow sustenance from these other examples: it is clear that a child's interest in and excitement over a game involve emotions. My question is, What is missing from Watson's and Crick's interest and excitement that would lead us to deny that their excitement and interest are emotions? My reply is that nothing is missing; their excitement and interest are also emotions, intellectual emotions.

I am not claiming that the very same excitement and interest can be found in Watson's and Crick's intellectual endeavors as are found in child's play—that, as it were, there is a selfsame excitement or a selfsame interest that occurs wherever there is excitement or interest. The sort of excitement and interest can vary with the object, the field of endeavor. So we might well say—following Aristotle on desire—that just as there is a child's excitement, so there is an adult's; that there is a mutual interaction between the object or endeavor and the specific nature of the emotion.

This leaves conceptual room for those who would argue that in some cases, such as a child's play, excitement is an emotion, but that in the case of Watson and Crick, the excitement is nonemotional. Once again, I would ask, "What is missing from, or present in, these cases of intellectual excitement to preclude their being counted as fully fledged emotions?" And once again, I would answer, "Nothing."

This, however, presents us with the following issue. Warnings against emotions have been part and parcel of our tradition. But our tradition does not warn against intellectual interest and intellectual excitement, and, given its emphasis on the intellect and rationality, it could hardly warn against them. In saying this, I want to insist that we do have some reason to be wary of intellectual interest and excitement. They can be dangerous for moral and intellectual concerns. They can be dangerous in the ways they get us to attend to intellectual matters when we should, instead, be working in nonintellectual areas. So too, they can be dangerous in the ways they get us to work in the wrong intellectual area or too much in one area. Here we need only think of the dangers of overinterest or overexcitement, or interest and excitement about the wrong facts, and the like.

Nonetheless, intellectual emotions are not dangerous as such; nor are they to be warned against as such. Perhaps, in fact, this is why intellectual emotions were not recognized. By not recognizing them, our tradition was able to warn against emotions, taken generally.

Having recognized them, we must modify our tradition as warning against only certain emotions. I will not offer a characterization of the difference between those emotions our tradition finds dangerous and those it does not. But I will offer this speculation, that difference is connected with the view that emotions, or the emotions we are warned against, are bodily, primitive, and essentially arational.

Perhaps then, the commonplace warnings against emotions should be understood as holding that emotions are typically, or too often, dangerous, but that only some instances of cool rationality are. Or perhaps it should be understood as holding that cool belief and cool rationality are not dangerous when, as is possible, the facts and procedures they involve are correct; but that emotions are always likely to be dangerous. Put in terms of people, the claim might be that inquirers who are emotional are likely to go wrong, that their being emotional typically raises questions about their reliability, but that coolly rational inquirers need not go wrong and do not, as such, raise questions about their reliability.

Here we might be reminded of the ways people's personalities and pre-occupations can influence their emotions, and whatever is in turn influenced by those emotions. Much could be said on this topic. For present purposes, it is enough to note the ways angry or self-pitying people are all too likely to be preoccupied by their own grievances and their own status, to be too harsh on others' views, especially views of those they feel have wronged them. These emotions and emotional states are moodlike: they seek out and collect, even create, sustaining or concordant facts (or "facts"), which they then use to justify and sustain that emotion, leading to further seeking, collecting, creating, and coloring.

Put in a related way, emotions can have a life of their own. When people are angry, especially when caught up in and controlled by anger,

they are all too likely to think, feel, and act in certain distinctive ways—that is, angry ways. Such angry people are not just people who, as it happens, are angry but are in all other ways just as they would be were they not angry. They do not just feel, think, and act *with* anger, they feel, think, and act *from* and *because of* anger.

It is, however, tendentious to use these claims to show that emotions are, as such, dangerous for intellectual activity, or are, as such, more dangerous than cool rationality. This claim uses emotions that mislead or overwhelm us, perhaps even emotional mountains or canyons, as their paradigm of emotions. But not all emotions, and certainly not all typical emotions, are like this.

These claims are also tendentious about rationality. To say the least, they express an exceptionally optimistic trust in rationality. Here we would do well to remember that there are many different cognitive personalities, styles, preoccupations, and indeed many different sorts of rational lives or lives of rationality. At least some, perhaps many, of these different cognitive personalities, styles, preoccupations, and lives can be cognitively harmful. To name only a few of these, they include undue credulity or undue skepticism, undue acceptance or undue rejection of tradition and authority, undue acceptance or undue rejection of traditional ways of working with or acquiring data. In addition, there are dangers of coolly misestimating the importance of, or the evidence for or probability of, both particular and general beliefs and theories. There are also dangers of coolly working with coolly held mistaken beliefs and theories—ranging from more ordinary ones, such as phlogiston theory in the 1800s, to extraordinary ones, such as holding in the late 1900s that the universe was created only two hundred years ago. The same also holds for coolly held mistaken theories, techniques, and beliefs about procedures.

So, both emotions and nonemotional thinking can mislead and be intellectually harmful. Here I might note that I see no way even to begin to count the ways emotions or cool rationality can mislead. Nor do I see any way to estimate their relative danger. I do not know how even to begin assessing the claim that emotions are more dangerous than rationality.

Just as emotions and rationality can be intellectually harmful, and just as lack of rationality can be intellectually harmful, so can lack of emotions. To see this, we need only consider lack of interest. This comes in different varieties. Here are some of them. A person can find a problem or a topic “dead,” boring, or not worth attending to: “If I have to do another of these problem sets, run another of these tests, I will scream!” One can be in the combined intellectual-emotional state of “burnout.” One can simply be indifferent to the issue. I find it hard to imagine how any of these forms or states of lack of interest could be thought better for intellectual progress in general or on a particular issue than well modulated, properly directed intellectual interest.

I do want to allow for contrasts between those intellectual approaches which involve emotions that enhance intellectual work and those which do not enhance such work, and perhaps detract from it. But in allowing for such contrasts, I also want to hold that emotions can be involved in the best approaches to intellectual work. And this shows not only that, but also why, there can be no useful contrast between approaches to intellectual work that involve emotions and the best approaches to intellectual work: the best approaches involve emotions.

We might take the claim that emotions are not useful for intellectual work, and are perhaps even dangerous for it, as a way of affirming allegiance to an ideal of the *primacy of rationality*. There may be a way to continue to affirm that ideal even while accepting many of the claims made here. One could hold that emotions are epistemologically and cognitively useful, but only when they are *controlled* by rationality, and that the best rationality requires the *cooperation* of emotions, but only of those that are controlled by reason.

However, it is unclear how this supposed primacy works, when the relations involved also require cooperation. Put in terms of people, rather than forms of mental activity, to the extent that I must cooperate with you and must secure your cooperation, I do not seem in control of you; indeed you could be seen as being in control of me or at least of our joint enterprises. And if we must both cooperate with each other, neither seems to control the other.

Perhaps, then, the claim about the relations between reason and emotion would be further modified to hold that for good intellectual work reason must play the *leading* role. Whether or not this is right, I see no need to pursue the matter. My concern was to show that emotions are not always dangerous for good, or even the best, intellectual work, and more strongly that they are often important for such work. And I have shown that.

These various points can be put in a closely related way, in terms of the neurotic defense of *intellectualization*: briefly, retaining the thoughts, considered as proposition-like, while repressing or dissociating from the affect of emotions and emotionally charged situations. I would point out that the very term *intellectualization* gives important information about what it describes. Intellectualization involves dividing emotions, and other affectively laden elements, into two parts, one with affect and the other just with proposition-like content. It then involves repressing, dissociating from, or otherwise ignoring those affective elements, while keeping accessible what is "of the intellect," the nonaffective content, such as thoughts as propositional content and the neutral, nonaffective holding of this content. It is, we might say, an attempt to avoid the emotional in favor of the intellectual.

One problem for intellectualization is that there are important emotional "aspects" of the intellect, such as intellectual interest and excitement.

There are also the intellectual emotions, or emotions about intellectual matters. This is a problem because, as it shows, to think that one can avoid affect and emotion by "living in the intellect" requires an idealization of the intellect. Like other idealizations, this one idealizes and distorts its object, intellect. Correlatively, it involves a negative idealization, a demonization and distortion of its split-off "negative object," emotions and affectivity.

Intellectualization thus involves an endorsement of a distorted intellect, thinking of intellect only in terms of intellectual, proposition-like content and not in terms of how the content is held and pursued. To put this in terms of my claim about the term *intellectualization*, this distortion is a motivated and indeed a principled distortion, giving expression to the fears, hopes, and goals that lead to and characterize intellectualization.

This requires only a slight modification to make a parallel point about philosophy. By somehow not seeing how important intellectual emotions are for good intellectual work, including good work in philosophy, and indeed by not seeing that there are intellectual emotions, philosophers make it possible for themselves to hold and espouse an idealized and distorted view of reason and of philosophy. Correlatively, to hold that view of reason, it is essential not to see that there are intellectual emotions, much less how important they are for good philosophical and other intellectual work.

Now, my claim is not that if one is emotionally distant, or has defective or distorted emotions, or at an extreme if one is affectless, one must make errors about value. Nor is my claim nearly so strong as David Rapaport's, made at the end of "On the Psychoanalytic Theory of Affects": "Affects . . . are just as indispensable a means of reality testing as thoughts. Indeed, they are more indispensable for reality testing in all except successfully intellectualizing and obsessional characters. Reality testing without the contribution of affect . . . readily changes into obsessional or paranoid magic" (1967, 508). My claim is only that there are deep and systematic connections between emotional, evaluative, and intellectual defects and errors, and also between emotional, evaluative, and intellectual strengths and abilities.

Much, indeed most, remains to be investigated. To suggest some areas for further investigation, I want to raise several issues. First, it remains to be seen how various emotions, even intellectual interest and excitement, help intellectual activity, or how they differentially help, perhaps even differentially harm, various intellectual activities. Here are some ways and areas they can help: they can be among the conditions for being a good intellectual worker, for being in good intellectual shape and having the energy to persevere. So too, they can make the life of an intellectual worker better for their presence. It is unclear in the case of Watson and Crick and DNA that they entered into the intellectual product itself and made it better or worse for their presence. For a contrast, we can see how a poem or play, perhaps even a piece of philosophy, such as one of Plato's dialogues, is

made better by the presence of such emotions. Certainly, some such pieces of work are criticized as dead for lacking such emotions.

Second, how do various intellectual emotions figure differently in different sorts of intellectual doings? It seems that some, but not all, intellectual work pretty much requires or is made much easier by the presence of certain emotions. Here we can think of the work of a biographer: contrast the difficulties faced by someone in emotional attunement with the subject and another biographer who is unable to be in attunement. This also holds for the intellectual, practical tasks of figuring out what would be a good way to play with a child or children of a certain sort. And, of course, it also holds for intelligent play.

Certain forms of intellectual work—of understanding—are impeded by certain emotions and emotional constellations. Here we could look at people who are *enmeshed* with each other. Boundaries between them are too porous or nonexistent, each is too caught up in the life of the other, too involved and overly concerned with that person. That people are enmeshed, and how enmeshment is constituted by particular sorts of moral concern, attention, and understanding, explains how such people are typically epistemologically ill placed or ill equipped to make good evaluative judgments of their own situations and of the needs and interests of others.

Further, it is argued in any number of ways that people who are deficient in their own emotionality and emotional repertoire are all too likely to be unable to understand other people accurately. Here we can think of people who are schizoid, psychopathic, or alexithymic. So the psychoanalyst Harry Guntrip writes in *Schizoid Phenomena, Object Relations, and the Self*, “the ego of the schizoid person in consciousness and in the outer world is delibidinized and feels no interest in objects,” and “As a result of this lack of feeling, schizoid people can be cynical, callous, and cruel, having no sensitive appreciation of the way they hurt other people” (1969, 30). In *Theaters of the Mind: Illusion and Truth on the Psychoanalytic Stage*, psychoanalyst Joyce McDougall writes of people who are alexithymic: “Instead of mentally elaborating their emotional states, they tended to discharge their feelings . . . often in inappropriate ways: through disputes, ill considered decisions, or a series of accidents. . . . It is evident . . . that an inability to capture and become aware of one’s own emotional experience must be accompanied by an equally great difficulty in understanding other people’s emotional states and wishes” (1991, 155, 160).

Here, clearly, we are in the realm of the importance or otherwise of countertransference for understanding others and oneself. I am not thinking just of how a therapist needs skill in reading the countertransference, but also how teachers and many others need this, too. For example, it will serve a teacher well to know what to make of finding a student or class annoying or hostile. A teacher who is unable to understand the source or

import of these feelings will lack an important tool for understanding and teaching. So too, I would argue that, when reading philosophy, we are well served by knowing why certain lines of thought are so annoying or boring, and we are handicapped by not understanding why we are reacting those ways.

To turn, now, to another point, a lack of interest in a mathematical problem set, say, can have one sort of effect in one sort of person, but not another. A very disciplined student might start her evening work with these problems and with great assiduity—to get them over and done with and to keep up a perfect record. Another student might put them off or rush through them. How the work gets done seems more related to the rest of the person's character than the mere presence or absence of interest. But in other cases, such as understanding other people, a lack of certain emotions might be a real handicap. And their lack or presence might bear in important ways on accuracy and completeness.

I have not mentioned—but I do not mean to exclude—the many ways emotionality or various emotions are held to be important for ethics, for acting well, for knowing what acting well is, and for other forms of ethical/practical knowledge.

There is a problem of the role of traits, emotions, and the like in accounts of what we do. It is possible for a person who is not interested in a subject to persevere in it, even to take it up as a part of life's work—for example, out of a sense of duty. So too it is similarly possible for a person who is not brave, who lacks the virtue of bravery, to do a brave act. We are thus faced with the very general problem of describing how only the emotions we have considered figure in intellectual work, and in other doings, too.

Taking up a closely related issue, we are confronted by the difficulties in understanding the role of emotions in many conditions that are named or otherwise characterized in terms of emotions. As just noted, it is possible for a person who is not brave to do a brave act. This may be understood as saying that one does not need the virtue of bravery to do what is brave. It can also be understood or extended to say that one does not need any brave emotion, emotionally dealing well with what is fearful, to do what is brave. Nonetheless, some of us—especially those of us influenced by Aristotle—will hold that for an act to be counted as a brave act, it must stand in certain relations to brave emotions. More strongly, it will be held that there is no way—or no natural way—to identify that act as brave independently of brave emotions. So, for example, even those of us who agree with Aristotle that a mercenary might be a more successful soldier than a brave citizen might also hold (and agree, I think, with Aristotle) that the arena of such success is brave acts and what requires them, and that this arena cannot be characterized except in terms of brave people and brave emo-

tions. These last claims that there is a need for those emotions for the identification or characterization of those acts as brave are, of course, controversial.

Let us now turn to respect and loving attention—two (sorts of) emotions that I earlier suggested are important for good intellectual work. There is, first of all, the issue mentioned above: it is possible to do a mathematics problem set perfectly even when one has no respect or loving attention for the problem or even for mathematics. Nonetheless, we might hold that respect and loving attention are in various other ways useful and important, perhaps even necessary, for good mathematical work. Even those who agree that respect and loving attention can be useful and important might wonder whether those features or sorts of attention that are useful and important can be characterized or even identified independently of those emotions.

To be sure, intellectual emotions such as respect for texts or loving attention to them can be important for intellectual work even if they are not necessary for them. And we have, I think, absolutely convincing evidence that they can be and often are important for them. In saying this last, I may seem to be claiming only that these emotions are instrumentally useful for intellectual activity. But I would contest the “only.” Even if they are instrumentally useful in this way, this is hardly insignificant. Further, even if these emotions are thus shown instrumentally useful intellectual products and work, they are shown to be internal to another “phase” or aspect of our intellect and our intellectual work and life. I am thinking of attention itself. Showing that attention is best for intellectual work when it is made up by, or infused by, those intellectual emotions thus goes well beyond showing that intellectual emotions are instrumentally useful—much less *only* instrumentally useful—for intellectual work and life.

#### NOTES

This essay is drawn from Stocker 1996 and has benefited from discussion at a conference at Santa Barbara City College, spring 2001.

1. See under “background feelings,” e.g., 150–51.

2. However, in Frankfurt 1976, Frankfurt recognizes and usefully discusses intellectual action.

V

Emotions, Action, and Freedom

*This page intentionally left blank*

## Emotion and Action

JON ELSTER

### 1. INTRODUCTION

In studying emotions we can take them as explananda or as explanantia. The two approaches are complementary. If we appeal to emotions to explain actions or other mental phenomena, we are naturally led to inquire into the origin of the emotions themselves. Conversely, the task of explaining the origin of the emotions would be of limited interest if they had no consequences beyond themselves.<sup>1</sup>

The explanans may be, as indicated, another mental state or an action.<sup>2</sup> The mental state in question may be a belief, a desire, or another emotion. Again, the task of explaining such mental states would be of limited interest if they had no consequences beyond themselves. Ultimately, therefore, the study of the emotions is largely motivated by their impact on action.

I shall proceed as follows. In section 2 I consider the impact of emotion on action that is mediated by what Nico Frijda (1986) refers to as *action tendencies*.<sup>3</sup> In section 3 I discuss the role of emotion in sustaining social norms that in turn are capable of generating, or sometimes blocking, action. In section 4 I turn to the indirect causal links between emotion and action that are mediated by cognition. Section 5 offers a brief conclusion.

### 2. ACTION TENDENCIES

It seems to be a fairly robust fact about the emotions that each of them goes together with one or several spontaneous action tendencies that, if unchecked, will result in action.<sup>4</sup> (I return to the nature of the checks in the next section.) Like other putative features of the emotions, this is not

a fully universal one. Hume asserted that “pride and humility are pure emotions in the soul, unattended with any desire, and not immediately exciting us to action” (1960, 367). One might counter that pride tends to induce boasting, which is kept in check only by social norms of modesty, and question whether humility (as distinct from humiliation) has enough of the other features of emotion to count as one at all. More radical counterexamples are provided by the aesthetic emotions. I shall not pursue this matter, however.

Some important examples of emotions with their characteristic action tendencies are given in table 10.1.

As indicated, a given emotion may be accompanied by several action tendencies. In some cases we can easily understand and perhaps even predict which will be produced. The urge to help the person one loves will arise only if he or she is in distress. Whether shame induces an urge to run away or to kill oneself depends on the intensity of the emotion. In other cases, the issue is more opaque. A study of fear in animals asserts that “rather than thinking in terms of two systems for reaction to different classes of punishment, it makes better sense to imagine a single fight/flight mechanism which receives information about all punishments and then issues commands *either* for fight *or* for flight depending on the total stimulus context in which punishment is received” (Hume 1960, 255). In other words, we don’t know. The different action tendencies of guilt may partly depend on the context in intelligible ways. Thus if the person one has harmed is dead, one cannot confess or make amends. Since one cannot undo the harm one has caused, the only way to restore equilibrium is by imposing an equivalent harm on oneself. But people may also respond to guilt by harming themselves even when they could make amends. A person who feels guilty about cheating on his taxes might prefer burning some of his money to mailing an anonymous check to the IRS. Psychoanalysts would say that such behavior is typical of neurotics or “moral masochists,” but I believe it could happen to anyone.

TABLE 10.1

| Emotion  | Action Tendency   |
|----------|---|
| Anger    | Cause the object of anger to suffer (revenge)                           |
| Hatred   | Cause the object of hatred to cease to exist                            |
| Contempt | Ostracism, avoidance  |
| Shame    | “Sink through the floor”; run away; suicide                             |
| Guilt    | Confess; make repairs; hurt oneself                                     |
| Envy     | Destroy the envied object or its possessor                              |
| Fear     | Flight; fight   |
| Love     | To approach and touch the other; to help the other; to please the other |

Action tendencies typically go together with *desires*—that is, desires to act. Typically the aim of a desire is to bring about some state of affairs. Even fear, when inducing a tendency to *flee from* danger rather than *flee toward* safety, aims at bringing about the state of affairs in which one is at some distance from the feared object. In some cases, however, emotions induce *wishes* rather than desires. Wishes, as I define them, do not have causal efficacy. One may wish for a state of affairs to obtain and yet have no desire to act to bring it about. Regret causes a wish that one had acted differently but no desire to change the past. In this case, there is nothing one *could* do to realize the wished-for state of affairs. In other cases, action is possible but not desired. Certain pseudoemotions fall in this category. Some people enjoy very much the feeling of righteous indignation, with regard to the plight of the poor or some other deplorable state. These emotions offer them an occasion for self-congratulation, but not a spur to reach for their wallet. Their emotion is *sentimental*, defined by the feature that “the existence or continuation of the emotion is motivated by the satisfaction experienced in feeling the emotion.”<sup>5</sup>

In hatred, what matters is that the hated person or group disappear from the face of the earth. In envy, what matters is that the envied person lose his or her possessions. In neither case is there an additional emotional satisfaction derived by the state of affairs being realized through *my* agency. In hatred, I may take action if I am well placed to do so, but I might be just as happy if someone else does it for me. In envy, I may even have a positive preference for the other’s ruin not coming about through my agency.<sup>6</sup> Some people who would be happy to see their neighbor’s house burn down would not themselves set fire to it even were there no risk of detection. (They might, however, deliberately abstain from calling the fire brigade.) By contrast, anger, contempt, love, guilt, and shame activate desires where it matters that they are realized through *my* agency. My desire for revenge is not slaked if the object is injured in a car accident. I may even wish for the object to know that he was harmed by *my* action. My desire for atonement is not alleviated if the person I have harmed wins the jackpot in a lottery. My feeling of contempt requires *me* to isolate myself from the object, not merely that the person be ostracized by others. I want to please the person I love, not merely that he or she be pleased.

Emotional action tendencies do not merely induce a desire to act. They also induce a desire *to act sooner rather than later*. To put this idea in context, let me distinguish between *impatience* and *urgency*. I define impatience as a preference for early reward over later reward—that is, a positive rate of time discounting—and urgency as a preference for early action over later action. The distinction is illustrated in table 10.2.

In each case, the agent can take one and only one of two actions, A or B. In case 1, these options are available at the same time, in cases 2 and

TABLE 10.2

|                                   | t <sub>1</sub> | t <sub>2</sub> | t <sub>3</sub> | t <sub>4</sub> |
|-----------------------------------|----------------|----------------|----------------|----------------|
| Case 1: Impatience                | A              | 3              |                |                |
|                                   | B              |                | 5              |                |
| Case 2: Urgency                   | A              |                | 3              |                |
|                                   |                | B              | 4              |                |
| Case 3: Impatience and/or urgency | A              |                | 3              |                |
|                                   |                | B              |                | 6              |

3 at successive times. In case 2, the rewards (whose magnitude is indicated by the numbers) occur at the same later time, in cases 1 and 3 at successive later times. Suppose that in an unemotional state, the agent chooses B in all cases, but that in an emotional state he chooses A. In case 1, the choice of A is due to emotionally induced impatience. In case 2, it is due to emotionally induced urgency. In case 3, it could be due to either or to the interaction of the two.

The classical moralists emphasized that the ability to wait was a hallmark of reason, whereas passions such as anger induce urgency:

How else did Fabius ["the Hesitater"] restore the broken forces of the state but by knowing how to loiter, to put off, and to wait—things of which angry men know nothing? The state, which was standing then in the utmost extremity, had surely perished if Fabius had ventured to do all that anger prompted. But he took into consideration the well-being of the state, and, estimating its strength, of which nothing now could be lost without the loss of all, he buried all thought of resentment and revenge and was concerned only with expediency and the fitting opportunity; he conquered anger before he conquered Hannibal.<sup>7</sup>

This is not to say that urgency is always irrational or counterproductive. If the emotion arises in a situation where waiting could be disastrous, it is clearly adaptive. In the face of acute physical danger, it is often rational to take immediate flight, which is not to say that we flee *because* it is rational. The action is triggered directly, rather than mediated by desires and beliefs as in the standard model of rational behavior. Although the machinery of rationality is idling, the outcome is the same as if it had been operating.

In other cases, however, urgent emotions arise even when nothing would be lost and something could be gained by waiting. The proverb "Marry in haste, repent at leisure" reflects this possibility. The desire for revenge can be so strong that the agent exposes himself to needless danger by acting immediately. An adulterous person who is overwhelmed by guilt may seek immediate relief by confessing to his or her spouse, without pausing to think whether it might not be better for all parties to break off the

adulterous relationship *without* revealing it.<sup>8</sup> Japanese kamikaze pilots, who were presumably in a highly emotional state, had to be trained to abstain from hitting the first target that presented itself and to wait instead until they could do maximum damage.<sup>9</sup>

My impression is that modern writers on the emotions tend instead to argue that emotions increase the *impatience* of the emotional agent. Whereas reason is capable of taking a long-term perspective, passion is myopic. I confess to being frustrated in my search for either explicit assertions or explicit evidence to this effect.<sup>10</sup> There is evidence that when people face a choice between immediate small rewards and larger but delayed rewards, emotional distress causes people to shift toward the former,<sup>11</sup> but this falls short of asserting a link between myopia and emotional arousal more generally. Some emotions, such as intense happiness, might well *enhance* the capacity to defer gratification. Be this as it may, it seems at least plausible that emotions such as anger, fear, envy, and jealousy tend to shorten the agent's time perspective, although in practice it may be difficult to sort out the effects of impatience from those of urgency.

In addition to their capacity for inducing urgency and impatience, a third feature of emotions is relevant for their impact on behavior: they tend to have a short half-life. With some important exceptions, anger and romantic love are not emotions that can be sustained indefinitely at a high level of intensity. This is one reason for believing that the purges after World War II in countries that had been occupied by Germany were motivated, at least partly, by emotion rather than by a desire for justice. The same crimes when judged in 1948 received much milder sentences than when judged in 1945. In explaining time-inconsistent behavior, this mechanism may offer an alternative to hyperbolic time discounting.<sup>12</sup> At time 1, when the agent is in the grip of emotion, he announces that he will make a great effort or a great sacrifice at time 2. When time 2 arrives and the passion has cooled off, he fails to carry out his plan. A person may swear off drinking because he feels guilty toward his family, but when the emotion wears off his resolve may disappear.

### 3. EMOTIONS AND SOCIAL NORMS

Whereas the analysis of section 2 applies to nonsocial emotions, such as fear of an avalanche, as well as to social ones, this section addresses a subset of social emotions: those implicated in the operation of social norms. These norms exercise a strong influence on action, while being in turned sustained by emotions. By this indirect mechanism, emotions affect action.

The idea of social norms can be understood in many ways. Economists and philosophers often view social norms as equilibria in games, more ac-

curately as *convention equilibria*. As all equilibria, these have the property that no one has an incentive to deviate unilaterally from the equilibrium behavior. In addition, they have the feature that nobody wants anyone else to deviate unilaterally either. (The latter feature is not, for instance, part of the bad equilibrium in the Prisoners' Dilemma.) If a telephone conversation is interrupted, there is a need for a convention to stipulate who is going to call again. It's obviously most efficient to have the person who made the call in the first place call again, since he or she is more likely to know which number to call, but the opposite convention is also better than having no convention at all. Suppose that the first convention is widely adopted in a community, but that one of the callers is unaware of it. Even though he did not make the first call, he tries to call up after the conversation was interrupted, but gets only the busy signal because the other person, following the convention, is also calling. When they finally manage to communicate, we can easily imagine the other person expressing disapproval because the first failed to abide by the convention. On later occasions, the memory of this emotional reaction might sustain the behavior dictated by the convention.

This is how I think of social norms: they are rules of behavior sustained by the emotional reactions of other people toward those who violate the rules.<sup>13</sup> As in the case I just mentioned, the rule may also be sustained by self-interest. Once I know about the convention, I have no incentive to deviate from it. And if we take the rule about driving on the right side of the road, it is sustained from three different sources. First, those who drive on the left side might be fined if caught by the police. Second, they might incur the angry reactions of others whose lives they endanger. And third, unilateral deviation is not in their self-interest.

Social norms, however, are sustained *only* by the emotional reactions of others toward norm violators. Let me first be precise about the kind of rules I have in mind. They are non-outcome-oriented injunctions to act or not act in specific ways. Consider two injunctions: "Always wear black in strong sunshine" and "Always wear black at a funeral." The former is instrumentally rational advice, as air circulates more quickly under dark clothes. The latter has no instrumental aspect at all. Moreover, the norm of wearing black at funerals cannot be seen as a convention equilibrium. Although people might be penalized for not wearing black, the penalties are not independent of the existence of the norm. By contrast, the harm you suffer when you drive on the left side of the road and hit another car is independent of the existence of the norm.

Among those (mainly sociologists and economists) who do *not* identify social norms with conventions, there is a tendency to limit social norms to rules whose effect is to reduce negative externalities, such as norms against littering or smoking in public.<sup>14</sup> I believe this definition, too, is inadequate to capture what we intuitively think of as social norms. In fact, in order to

define a homogenous category I shall *exclude* norms that target externalities. The reason has to do with the very different emotions induced by violation of these norms, compared with violation of the other norms I shall consider. If somebody smokes in my presence, he imposes a harm or a risk on me. As a result, I may get angry. When he realizes that he has imposed a harm on me, he may feel guilty, even if I do not show my anger. Note the structure of the causal chain. The perception of a harm is the cause of both my anger and his guilt. We may contrast this case with the violation of a code of dress, such as wearing last year's fashion at a high school prom. The unlucky girl who makes this mistake will be met not with the anger of her classmates, but with their contempt. She will feel not guilt, but shame. Moreover, the structure of the causal chain is different. The shame is caused by the contempt, not by the norm violation. These are properly called *social* norms, whereas the rules against harmful behavior are more naturally thought of as *moral* norms.

At the conceptual level, the most important difference between anger and contempt is that the first is triggered by the *action* of the norm violator, the second by his or her *character*. Similarly, guilt is triggered by the belief that one has done a bad action, shame by the feeling that one is a bad person. The guilty person can hope to rid himself of guilt by making amends, but the person in the grip of shame can hope only to escape the contemptuous look of others. I do not want to make too much of the distinction. Littering in public may elicit anger in some observers and contempt in others. There are emotions that involve anger and contempt at the same time, as when a social inferior violates a moral norm. What matters for my purposes is that there are clear-cut cases on both sides of the large borderline area.

Violations of social norms are met with contempt, triggering avoidance or ostracism. The correlative emotion in the norm violator is the devastating feeling of shame. When other people refuse to have dealings with a norm violator, he will also suffer materially in a number of ways. He will not be able to find anyone to marry him, to employ him, or to enter into a partnership with him. At the same time, those who refuse to deal with him may also suffer by virtue of giving up the opportunity for mutually profitable transactions. I want to state two propositions about these various effects of norm violations. First, the feeling of shame is a much worse consequence of contempt than the material deprivations imposed on the norm violator. Second, the shame is shaped more powerfully by what it costs others to ostracize him than by the material effects of ostracism on him. The depth and intensity of contempt are best expressed by the willingness of others to forego material benefits. There seems to be considerable consensus on the first proposition.<sup>15</sup> The second proposition receives less attention but seems to follow quite naturally from the first.

Social norms may reinforce spontaneous presocial action tendencies

or, on the contrary, block their expression. In some societies, the spontaneous tendency to take revenge—as when I stumble over a stone and retaliate by kicking it—is reinforced by strong social norms. In other societies, the tendency is kept in check, at least partially, by the norm of turning the other cheek. In some societies, strong social norms against behavioral expressions of envy tend to mute the action tendencies of that emotion. In other societies, the norm against sticking one's neck out tends to license aggressive behavior toward those who stand out. In some societies, prudential fear is accepted, even applauded, and visceral fear at least condoned.<sup>16</sup> In others, both prudential and visceral fear are stigmatized as dishonorable. As these examples show, *emotions of shame and contempt regulate the behavioral expressions of other emotions*. In fact, these emotions may even regulate themselves. In many societies, third parties express disapproval of those who express their disapproval too strongly. In Tahiti, there is both control by shame (*ha'ama*) and control of shame: "Although gossip is an important part of 'shame control,' the words designating gossip have a pejorative tone, and gossiping is said to be a bad thing to do. Ideally, the behavior which would produce shame on becoming visible has to spontaneously force its way into visibility; people are not supposed to search out shameful acts. Such a searching out is itself a *ha'ama* thing" (Levy 1973, 340).

#### 4. COGNITION-MEDIATED IMPACT OF EMOTION ON ACTION

The relation between emotion and cognition is threefold. First, emotions usually have cognitive antecedents. Second, emotions can be the target of cognition (and thus trigger metaemotions). Third, emotion can shape or distort cognition (and thus change the induced emotion). These three causal mechanisms can interact in intricate ways, as I shall try to illustrate by the examples of envy and anger. They all involve the cognition-mediated transmutation of emotion as a result of a threat to one's self-image.

Envy is triggered by the cognition that another has something I want, in both senses of the verb. I may not be aware, however, of the fact that I am envious. Although my envy may be obvious to others, I do not myself draw the correct inference from my denigrating remarks and obstructionist behavior.<sup>17</sup> One day, however, the fact that I am envious dawns upon me. Living as I do in a society in which envy and its expressions are strongly disapproved of, I feel acute shame (or maybe guilt; it does not matter for the present purposes). This metaemotion can set up a strong cognitive pressure to rewrite the script. Even with the most worthy rival it may be possible to come up with a story according to which he or she acquired the coveted good immorally, illegally, and perhaps at my expense. Once that new set of

beliefs takes hold, the horrible emotion of envy is transmuted into the wonderful emotions of righteous indignation or righteous anger.<sup>18</sup> Anti-Semitism probably owes a great deal to this mechanism. Its main result is to *remove the inhibition on action* caused by the social stigma on expressions of envy.

Seneca writes that “men whose spirit has grown arrogant from the great favor of fortune have this most serious fault—those whom they have injured they also hate (*quos laeserunt et oderunt*)” (II.xxxiii). For the rich and powerful, assuming guilt for the result of their reckless behavior is inconsistent with their pridefulness. Instead, they try to justify their action by rewriting the script to put the blame on the other party, thus providing a reason for harming him even more. Whereas the action tendency of guilt is to undo unjustified harm done to others, this mechanism makes one compound it. We may understand one of La Rochefoucauld’s maxims along similar lines: “We often forgive those who bore us, but we cannot forgive those who find us boring” (M 304).

Emotions can also affect cognition and hence action in ways that have nothing to do with a threat to one’s self-image. As Seneca noted, the urgency of emotion detracts from the efficacy of belief formation: “Reason grants a hearing to both sides, then seeks to postpone action, even its own, in order that it may gain time to sift out the truth; but anger is precipitate” (I.xvii). Here, emotion causes suboptimal investment in information. In addition, emotion may prevent us from drawing correct inferences from the information that we do possess. A Swedish proverb says, “We believe easily what we hope and what we fear.” The person in the grip of romantic love may interpret the most insignificant actions or utterances by the other person as a sign that the emotion is requited. Even refusal is seen as a sign of an emotion that is so strong that it has to be suppressed. Conversely, the person in the grip of jealousy may, like Othello, interpret the most innocent behavior as evidence that his fears are justified. In such cases, the beliefs are shaped directly by the emotion, rather than indirectly by emotionally induced urgency.

Emotion and other “visceral factors” may also prevent our ability to anticipate future emotional states.<sup>19</sup> Because of a “cold-hot empathy gap,” it is difficult for a person in a calm state to imagine with sufficient vividness what he might feel (and be led by his feelings to do) in an excited state. The six people who killed themselves in France in June 1997 after being revealed as consumers of child pornography probably did not anticipate just how horrible the shame would feel. Conversely, a “hot-cold empathy gap” makes it difficult for someone in an excited state to imagine that it will not last forever. Once the shame hit them, the same individuals were probably unable to imagine that it would ever get less painful.

## 5. CONCLUSION

The discussion in sections 2–4 indicates only some of the ways in which emotion can have a causal impact on action. Elsewhere I have briefly sketched twenty mechanisms through which this impact can occur,<sup>20</sup> only some of which have been considered here, albeit in more detail. The most glaring omission in the present chapter might seem to be the neglect of future emotional experience as the motivation for action—emotion as the terminus ad quem of action rather than its terminus a quo.

Emotional experiences, to be sure, can be immensely satisfying, at least as valuable (and usually much more so) than hedonic satisfaction. Given that economics has built an immensely sophisticated and powerful apparatus for explaining action by consumers in terms of hedonic pleasure seeking, it would seem appropriate to try to do the same for action motivated by the search for emotional satisfaction. There is, however, a paradox attached to emotional satisfaction that does not arise in the case of hedonic experiences. Let me introduce it by way of an example.

In 1998, I was in Paris when France won the World Cup soccer championship. I remember my French daughter-in-law calling me from Oslo and screaming with happiness, “On a gagné, on a gagné!” Afterward, I went for a walk in Paris and observed the remarkable collective ecstasy generated by the victory. Now, it is fair to say that this victory was a low-probability event. In fact, and this is the key point, the ecstasy occurred *because* it was a low-probability event. *Surprise* has a multiplier effect on emotional satisfaction that it does not have on hedonic satisfaction.<sup>21</sup> This effect also explains the intense emotional pain caused by the surprise elimination of France in the 2002 competition.

In other words, because the *actual* utility of the event depends on its ex ante probability, that probability enters twice into its *expected* utility. This fact blocks the use of a standard tool of economic theory, the cardinal von Neumann–Morgenstern utility function. Cardinal utility is linear in probabilities: the utility of an uncertain prospect that might yield A with probability  $p$  and B with probability  $(1-p)$  is  $p$  times the utility of A plus  $(1-p)$  times the utility of B. If the utility of A exceeds that of B, the utility of the prospect will go up as  $p$  goes up. With emotionally based utility, however, the relation might not even be monotonic: the expected utility of an outcome might go up as its probability goes down.

I am not an economist, so I cannot tell whether this fact is a serious obstacle to the systematic study of emotion. What seems clear, however, is that many of the greatest emotional experiences in life—good or bad—come as a surprise. We may not even have attached *any* ex ante probability to them at all. Those who insist on planning their emotional life might not have much of one.

## NOTES

1. I am not saying the task would be of *no* interest. The question why music is capable of generating emotion is an intrinsically interesting issue in the philosophy of mind that is largely unrelated to the philosophy of action.

2. In addition, emotions may explain nonintentional behavior such as a trembling hand, or physiological events such as a heart attack following a great rage.

3. My debt to Frijda is so large that I shall make only this blanket acknowledgment to his work.

4. These tendencies are not mere dispositions, which need to be triggered to bring about action. They are already triggered by the emotion. We may think of them as incipient actions or “virtual actions,” as Thomas Aquinas said with respect to the destructive urge in envy (II.36, art. 3).

5. Budd 1985, 96. See also Tanner 1976–77.

6. Seneca, III.v. Anger “exceeds spite and envy; for they desire a man to be unhappy, while anger tries to make him so.” As I argue below, the passive character of envy is not intrinsic to the emotion but a result of the disapproval envy-inspired behavior meets in others.

7. Seneca, I.xi.

8. The seemingly unselfish propensities induced by guilt and love may, therefore, have a selfish origin.

9. Hill 2002. A related fact is that the Palestinian organizers of suicide attacks try to eliminate candidates with a suicidal disposition (Drori 2002).

10. Frijda 1986, 120, comes close to saying that emotion induces myopia but does not explicitly say so. Some writers seem to blur the distinction between emotions having a short time horizon and their being short-lived. Thus when in *The Federalist* no. 6 Hamilton refers to “momentary passions and immediate interests,” he seems to distinguish between a *short-lived* motivation and a *short-sighted* one. He nevertheless assimilates the two in no. 15, when he opposes “general considerations of peace and justice to the impulse of any immediate interest or passion.” In the related case of addiction, several writers have discussed the idea of endogenous changes in discounting rates (Becker 1996, 120; Orphanides and Zervos 1998; O’Donoghue and Rabin 1999).

11. See Tice, Braslavsky, and Baumeister 2001.

12. For recent summaries of the relation between preference reversal and hyperbolic discounting, see the contributions by Skog and by O’Donoghue and Rabin to Elster 1999.

13. For a fuller account see Elster 1999, chap. 3.

14. This is, for instance, the main line of argument in Coleman 1990; see also Arrow 1971.

15. Thus Lovejoy 1961 quotes Voltaire: “To be an object of contempt to those with whom one lives is a thing that none has ever been, or ever will be, able to endure. It is perhaps the greatest check which nature has placed upon men’s injustice”; Adam Smith: “Compared with the contempt of mankind, all other evils are easily supported”; and John Adams: “The desire of esteem is as real a want of nature as hunger; and the neglect and contempt of the world as severe a pain as gout and stone” (181, 191, 199).

16. For this distinction, see Gordon 1986.

17. For a fuller discussion of various ways of being unaware of one's emotions, see Elster 1999, 255–60.

18. Following Aristotle, I view anger as caused by someone harming my self and indignation as caused by the sight of someone enjoying undeserved fortune.

19. See Loewenstein 1996 1998.

20. See Elster 1999, 328–31.

21. See Mellers, Schwartz, and Ritov 1999.

## Emotions and Freedom

JEROME NEU

Once one recognizes the centrality of thought in constituting emotions, it becomes easier to appreciate the role of social, political, and cultural factors in shaping our emotional lives. If one thinks of emotions as un-analyzable feelings fixed in our biology, different social contexts might trigger those fixed feelings and different societies might value them differently, but one might readily miss the ways in which society shapes thought and so perhaps more directly shapes our emotional world.

It does not follow that we can simply choose our emotions by choosing our beliefs and thoughts, as some—such as Sartre—have supposed, if only because we don't in general have such direct willful control over our thoughts and beliefs (any more than we have direct willful control over the nature of the society in which we find ourselves and our emotional possibilities being shaped). It does, however, leave room for the hope, the Spinozist hope, that understanding can make us free, can transform—to some degree at least—our emotional lives. Understanding the role of thought in emotion can help us better understand both the possibilities and the limits here. It is partly for this reason that I consider understanding the conceptual structure of particular emotions valuable. Take jealousy.

### JEALOUSY

Jealousy is interestingly complex. It is compounded of fear and anger and is, at least in its erotic forms, importantly tied to love. In understanding jealousy, I would emphasize the constitutive role of thoughts (as with other emotions), in particular thoughts involving fear of loss; more specifically, fear of alienation of affections (that is, loss of those affections to a third

party, a rival), and, more generally and more deeply, fear of annihilation. I believe that a focus on fearful thoughts enables us to see many important connections (for example, with love) and distinctions (say from envy). The constituent thoughts and beliefs provide a way into understanding the psychological and social conditions and implications of the emotion—including issues about psychogenesis and about eliminability. To what extent are the emotions we experience natural and inevitable? What else might we have to give up if we wanted to eliminate jealousy or envy or other painful emotions?

It is important that jealousy and envy are not the same. Othello is jealous, Iago is envious. Jealousy is typically over what one possesses and fears to lose, while envy may be over something one has never possessed and may never hope to possess (think of envy over someone else's beauty or intelligence). Indeed, the focus of envy is typically the other person, the possessor, rather than the particular thing or quality one is envious over (a thing that may not in itself even be desirable to the envier, whatever its perceived value to the present possessor). In jealousy there is always a rival, believed or imagined, but the focus of concern is the valued object (centrally a person, a desiring as well as a desired object). For jealousy, but not envy, the other must be seen as a genuine rival for the object: their gain is one's loss. The evil eye, on the other hand, can be directed at anyone who prospers; it needn't be at the envier's expense. Similarly, *schadenfreude* (joy at another's suffering), the inverse of envy (pain at another's success), may be impartial in that the other's loss need not involve a material advantage to the person who takes pleasure at it. This (apparent) aloofness of envy may make it more intractable.

Of course, our usages are very flexible in this area, but important connections may be more readily perceived if we preserve certain distinctions that ordinary language does not always insist upon. If we restrict jealousy to relations with people, the place of the desire to be desired and for affection comes into sharper focus. And once the central relation in envy is seen to be between the envier and the person envied (even if the envy is *over* a thing or quality), the alternatives of malicious and admiring envy become clear. In the case of malicious envy, one wants to lower the other (to one's own level or below); in the case of admiring envy, one wishes to raise oneself (to become like the other). I would argue that these two types of envy have different instinctual sources and developmental paths, and that, as a result, malicious envy, unlike admiring envy and unlike jealousy, may not have appropriate objects—that is, the explanation for its occurrence may always involve pathology.

Going with these differences, and looking to a psychogenetic account of the origin and place of jealousy, there are grounds to distinguish between jealousy and envy in relation to the hopes for emotional transformation

connected with two types of ideals: communitarian and socialist. (The personal and the political are sometimes mistakenly assimilated. The ideal of the loving community is not the same as the ideal of the just community, though one might wish for, and work for, both.) Despite the hopes of social reformers, the possibility of jealousy cannot be eliminated. It is wrong to think that jealousy is always necessarily misdirected, that it cannot have appropriate objects—on the contrary, it can. The presence and persistence of jealousy have more to do with the development of self-identity than with the possession (understood as ownership) of others; and while the underlying fears may make us prone to pathological forms of jealousy, it is also the case that jealousy is tied to certain forms of love—so the elimination of the possibility of jealousy might involve the loss of much else. Once one appreciates the thoughts that underlie jealousy, their relative independence from claims of right and their attachment to basic fears, one can see that jealousy holds a place in human life as fixed as human vulnerability and the need for certain types of love. On the other hand, the same difficulties do not, it seems to me, stand in the way of the hopes of social reformers in relation to envy. Which is not to suggest that envy can be readily dislodged from its place in human life, nor is it to say that the harmful consequences of jealousy cannot be ameliorated. Some of the relevant arguments depend on conceptual clarification, some draw on the resources of literary study and the social sciences. The desirability or undesirability of certain emotional reactions, like the desirability and undesirability of certain social structures (which may themselves sometimes be established partly to deal with those reactions—e.g. certain institutions of marriage and of punishment), must take into account an understanding of their nature and sources and the related conditions of change.

It should be recognized that, on the social level, jealousy may do important institutional (indeed, even biological) work. Jealousy is not a merely bourgeois passion; it is not confined to societies with capitalistic or monogamous social arrangements. (That it is not confined to a particular class within such societies goes without saying.) Every society that prefers and sanctions certain social arrangements over others, which is to say every *society*, will have room for jealousy: it serves to reinforce and protect the preferred arrangements (in particular, the preferred distributions of sexual affection). While social arrangements may vary, whatever the social arrangements, jealousy may serve to reinforce them.

Perhaps more important, on the individual and psychological level, jealousy need not be simply a pathological form of possessiveness, misdirected toward people as though they were things. It has deep developmental roots that connect the fears involved with essential processes of identification, with who we are and who we may become through our relations with others. A proper account of the thoughts and feelings in jealousy needs to

be psychogenetic. The place of identification in making jealousy both more and other than simple possessiveness (in the sense of ownership) can, I think, be understood through Freud's account of the Oedipus complex (that triangular set of ambivalences concerning love and hate, supporting figures and rivals, that we all must negotiate) and by utilizing some of Winnicott's insights about development: insights having to do with transitional objects (those teddy bears and security blankets of our early years that manage to be both part of us and yet independent, our first possessions), with the mother's face as a mirror in which the infant can recognize himself (it is through such early acknowledgement and appreciation that we develop a sense of a continuous real self, and the need for such sustaining reciprocity persists in later life), with the importance of being alone in the presence of another in relation to developing independence and the capacity to be truly alone (and giving psychological depth to later notions of someone "being there" for us even when they are not physically present), and finally with hatred's constructive role in identity formation and the ability to love (testing our boundaries and our dependencies). The result of our complex developmental histories is that jealousy is not necessarily pathological, though it certainly can be. The mere fact that jealousy may be connected with the development of self-identity rather than the possession of others (as though they were things) is not itself enough to ensure that either one's identity or one's jealousies will not be pathological. The result is also that our forms of attachment, and the risks of loss that they entail, would have to be very different if the possibility of jealousy were to be eliminated. We don't get to simply choose the psychological mechanisms by which we are formed or the intricate emotional complexes through which we must negotiate our lives. These limits, which might be thought of as limits on our freedom, should, however, also be understood as providing some of the rich possibilities for attachment and interdependence that make life meaningful and valuable.

While I don't think it worthwhile to seek an exceptionless set of necessary and sufficient conditions for jealousy (or any emotion, for that matter)—our usages are simply too loose to expect a stable set of such conditions and (as Wittgenstein taught us) it is often a mistake to assume ordinary language concepts have a clear essence—valuable connections and distinctions can emerge if we seek patterns in our attributions of emotion. And these may have political significance. For example, as I've suggested, if we maintain certain distinctions that ordinary language would indeed permit us to ignore, we may achieve an understanding of the instinctual and developmental roots of jealousy and envy that might otherwise elude us. And this may help us better understand the prospects for individual and social transformation. Something similar may be said for pride.

## PRIDE

How are we to understand the emergence in our time of pride, one of the traditional seven deadly sins, as the banner under which marginalized groups assert their claims, as the theme of identity politics? As with jealousy, a consideration of pride may help clarify how a structured account of the conceptual and other conditions of a particular emotion can reveal features of our identity and their dependence on social beliefs and political institutions. Christian theology still condemns the sin of pride, as it has done for centuries. Indeed, in many accounts, pride represents the worst of all sins, for in pride a person is thought to turn his or her back on God. In extreme cases, it is charged, pride may lead people to imagine that they are themselves God, self-sufficient rulers of their own fate. Humility is urged instead. It needs to be recognized, however, that (proper) pride may elevate and even redeem people who are suffering and oppressed. To see this, one need not insist that God is dead. The blanket and unthinking condemnation of pride has costs. In our time, the invocation of pride has in many instances served to suffuse the downtrodden (and ultimately the wider society) with the attitudes necessary for transforming lives of desperate poverty or self-loathing, and to overcome socially imposed disadvantages and disabilities. This is not to deny the risks of socially isolating and divisive group identities, of the “narcissism of minor differences” that sometimes erupts into racial and ethnic hatred and even civil war. (These risks, however, should not be confused with “sin.”)

To see the appropriate personal and political place of pride, one must properly understand the differing roles of responsibility and value in the constitution of pride. In particular, responsibility for a characteristic is *not* a conceptual condition for pride in that characteristic (whether the pride be in one’s country, one’s family, one’s race, or even one’s chosen sports team), but—by contrast—positive valuation *is* a conceptual condition of pride, and that feature allows room for the transvaluation of values when previously denigrated characteristics claim recognition and acceptance, as in Black Pride, Gay Pride, Deaf Pride, and so on. While one may not be able to directly will social change, one can take steps to transform social attitudes, and one—especially when one becomes many, becomes a social movement—may sometimes succeed.

One might think responsibility should be a condition of pride—that, for example, pride should be restricted to virtues and achievements rather than be extended to include natural endowments and gifts. Responsibility in turn might be seen as conditioned on causal role or individual choice. But despite the many possible senses of “responsibility,” responsibility is *not* a condition of pride. While there are conceptual constraints of other sorts on pride, there is no *conceptual* error in claiming to be proud where one cannot claim responsibility (whether one is proud of a sports team, one’s

cultural heritage, one's parents, or one's height). If responsibility were a condition of pride, a politics of pride in group identity, where the characteristic defining group identity (whether skin color or sexual preference, ethnic or national origin) was not itself something deliberately chosen, would make no sense. The point of claiming such pride is different.

The political value of pride in identity politics derives (at least partly) from the internal place of values within pride. (When O. J. Simpson allowed as how he was "not proud" of his wife abuse, he was using "pride" to mark his choice of values, in this case to show his acceptance of community values.) On all accounts, the source of pride must be seen as an achievement or an advantage; pride involves positive valuation. That is a conceptual condition. The point of pride as a member of a group, the pride of belonging, depends on some distinctive virtue of the group, on its perceived value. Claiming group membership is a way of claiming the associated value for oneself. This reflects the conceptual dependence of pride on positive valuation. (On Hume's excessively mechanical account, lacking the belief in value, one would lack the double association needed to produce pride. Rather, I would say, lacking the needed belief, whatever was produced would not be considered pride.) That is, group pride, the pride of membership or belonging, like the pride of ownership, depends on value—the subject, like the owned object, is seen as valuable. The twist in recent identity politics is in the seeing of value.

Identity politics involves transvaluation, a reversal of received values: a previously despised property comes to be seen as valuable: "Black is beautiful." Earlier majority values or norms are rejected as mistaken, biased, blind. A previous source of shame becomes a source of pride. The point is *not* that one should not be ashamed of one's skin color (for example) because one cannot help it, did not choose it, and so is not responsible. The point rather is that one should not be ashamed of one's skin color because there is nothing wrong with it in the first place.

Of course there are problems with traditional identity politics, some stemming from the grayness of many categories. Some aspects of our identity are fixed independently of what we think or would like to think; some depend on choices we make and allegiances we adopt. But in either case, criteria for identity can be (and very often are) contested. What and who is *in* a particular category? Even a category such as race, which might appear straightforwardly biological, can be problematic: skin color may provide no sure index of anything and we may all in the end be multiracial.

Looking to another disputed category, gay behavior, desires, inclinations, and attitudes can all vary in more ways than marked even by Kinsey's classifications (exclusive, occasional, etc.), and that before account is taken of the unconscious. Who are "we" for purposes of group membership? If we think of the gay-identified as excluding the repressed or closeted homosexual, we may be focusing too much on the voluntaristic aspects of

identification, where identification is self-identification (as with sports fans). But where the political problem arises from identification, and stigmatization, by others, perhaps a politically relevant notion of identification must be broader (even if it risks objectification of individuals and reification of the categories of the others—after all, the struggle is with or against those very others). Even when one is not asked and does not tell, one may be discriminated against, one's life restricted. It may, for example, be less possible to become who one would be happiest being.

Some go beyond issues of grayness. So far as political radicals who deny the very existence of the relevant categories aim at denying privileged valuations of *either* side of dichotomies, the message may ultimately be the same as "Black is beautiful" or "Gay is good" or "Deaf power." For the point, typically, is not to say black is better than white, or gay is better than straight, or deaf is better than hearing, but simply to deny the denigration of the minority position. The point is to demand political equality, equal concern, and respect.

And the notion of "respect" at stake repays scrutiny. The pleasure that Hume discerns in pride is ultimately a form of self-approval. But self-approval is ambiguous in a way that may help explain the dual attitudes, sin to be avoided and virtue to be sought, toward pride itself (whether regarded as a character trait or a passion). We can understand the ambiguity in terms of a contrast between self-esteem and self-respect. Self-respect, having to do with one's rights and dignity as a person, may be noncomparative. Self-esteem, having to do with one's merits and self-valuation, may depend on the standards of value in one's society and how one compares oneself with other members of that society. Thus understood, of self-respect one cannot have too much, of self-esteem one obviously can.

So far as pride is a matter of self-respect and the respect that is owed one as a person, quite independent of special individual (or group) merits, everyone is entitled to it. It is a condition of moral identity. This leaves self-esteem as quite another matter. One of the errors of certain recently popular self-help psychologies is to suppose that increasing self-esteem is simply a matter of changing one's attitude rather than the more strenuous activity of changing one's life. So far as esteem depends on merit, a pride that simply depends on deciding one is "OK" (whatever one does) becomes like the sinful individual pride of old: one falls into unjustifiable self-satisfaction. Group credit too, or "bragging rights," does little to advance claims based on merit unless responsibility (as well as "nearness") can somehow be claimed. So far as group pride gives self-respect and asks for respect from others based on one's common humanity and equal moral rights, there need be no sin, no error.

Greek *hubris* (thinking oneself superior to the gods), like Christian pride (thinking oneself independent of God, self-sufficient), involves placing oneself above one's station. This is one of the features of pride that makes it

peculiarly appropriate as the banner for political movements that seek to change the station of those in them—that seek a transvaluation of values. Both identity politics and the sort of radical politics that denies the very categories used in discrimination, whatever their views on whether God has died, deny that the social valuations and positions that denigrate certain groups and privilege others are ordained by God. Times have changed. The death of God would leave the concept of sin with little conceptual foothold. But even in a world where God is still believed to preside, an attack on social hierarchy need not be regarded as sin, for it is not an attack on God: social hierarchy is not a matter of natural law, is not God-given. These political movements are challenging positions in the political world rather than a God-given order. And, as just noted, on an individual level, the self-approval that is characteristic of pride may be ambiguous, and the different significances may be understood in terms of a contrast between self-esteem (which can be excessive and unjustified) and self-respect (which does not depend on invidious comparison and may be essential to human dignity). A politics of self-respect, where the self has a social identity, may not be so ungodly after all.

#### FORGIVENESS

Self-respect may also place certain constraints on forgiveness, that is, on the forswearing of resentment for injuries and the disrespectful attitudes that may underlie them. Feeling resentment and anger (and variants such as indignation) when appropriate may be a condition of self-respect, and so failure to feel appropriate anger may be a sign of insufficient concern for one's rights and dignity, insufficient self-respect. What is involved in forgiveness is an interplay of attitudes. What is resented is an attitude and what changes when one forgives is one's attitude toward the person whose attitude originally caused resentment.

Resentment focuses on the intention or lack of due care and respect that an injury may convey, and forgiveness involves a change of heart toward the wrongdoer. Understanding forgiveness as forswearing resentment enables one to see systematic connections among the sorts of reasons that may serve as appropriate grounds for forgiveness. For example, both "repentance" and "old time's sake" enable one to distinguish the attitude manifested in an agent's act and the current fundamental attitude of the agent, and so make sense of St. Augustine's somewhat mysterious counsel to "hate the sin, but love the sinner." (The mystery arises because people are usually taken to be identified, to some degree, by and with their acts.) There is a disparity in messages communicated. The divorce between act and agent, or between the attitude manifested in an act and the attitude of the agent, helps us see what shifts when understanding leads us to move from resentment to forgiveness.

In order to preserve the inner goodness of the wrongdoer, perhaps to make it easier to go on loving the sinner while condemning the sin, people sometimes distinguish an inner (real and true) self and an outer (false and determined self). But the separation is as false as the Cartesian split between mind and body that it mimics—both approaches treating the real or essential self as though it were a disembodied mind. It is the schizoid vision of the self popularized by R. D. Laing in the 1960s. The perhaps comforting vision of a well-meaning (and, in the full schizoid version, all-talented and omnipotent) self should be resisted. There are several protections against the metaphysical and moral temptation to regard one's inner or mental life as somehow "true" and one's bodily life (with its overt, observable actions) as external and somehow "false." The first is to consider carefully what "false" might mean here. In most senses (except where it is equated from the start with things bodily and visible) it can apply equally to things mental and physical. That is, emotions and thoughts may be as "false" as social roles in the sense of being, for example, undesired, unchosen, or disliked. Properly understood, the true/false distinction cuts across the mind/body distinction, rather than running parallel to it. Mind (mental states) can be false as well as true. Bodily states can be true as well as false. This connects with a second major protection against the schizoid delusion: the recognition that not all social roles are false. We build our identity partly through others' perception and recognition of us. Some of the social roles that make us who we are we in fact desire and choose. Being a parent, friend, student, lover need not be "false" just because each is a social role involving an embodied interacting life. And a third remedy to a schizoid split of mind and body is to consider what constitutes a "mental state." Philosophers such as Gilbert Ryle have emphasized the behavioral aspects of intelligence, knowing how, vanity, and so on. As Wittgenstein put it, "The human body is the best picture of the human soul." The self inevitably becomes empty if it is regarded as disembodied because the attribution and existence of many psychological states depends on their bodily expression. The moral comfort of a retreat to a well-meaning inner self can be bought only at the cost of gross distortion of just what it is that makes us who and what we are. Sinners cannot shed their sins by a simple metaphysical shift in identity.

Of course, actions that cause harm, even actions that intentionally cause harm, need not reflect attitudes that express disrespect. Think of the innocent victims of Allied bombing in Europe in World War II. Their suffering may be seen as the price of defeat of the Nazis (or more accurately as part of the price—the bombers too ran risks and made sacrifices). They (the victims who survived) might well say, "I forgive you my injuries, I would have done it too" or "It was necessary." It might make as much sense to say "there is nothing to forgive" as "I forgive you," but the second formulation can be a way of acknowledging a shift in attitude based on a full understanding or appreciation of the situation. It acknowledges just-

fication for a harm. To say simply “there is nothing to forgive, you did the right thing” might fail to recognize and properly note the existence of tragedy. There are situations in which the best that one can do is still wrong, has a moral cost that remains to be regretted by the agent and (perhaps) forgiven by the victim. Or, to put it slightly differently, the lesser of two evils may still leave the chosen evil a wrong (an undeserved harm), even while ultimately the right thing to do. Therapy can sometimes be an effort to get someone to see that not every injury is an affront, that harm may be done (even intentionally) without disrespect.

One needs to be careful about the role of necessity in such matters. We don't typically forgive people because of some general belief in determinism, because of some notion that no one is ever responsible because actions can always be traced back to causes ultimately outside their control. Determinism as such is not the issue. When we excuse or forgive someone, it is not because of some general belief in causal order in the universe; it is not because we believe that every event has a cause. Resentment is forestalled or inhibited in particular cases for particular reasons, broadly classifiable in terms of the voluntariness of the particular (otherwise) offensive or injurious act (where certain cognitive conditions, such as nonculpable ignorance, or certain control conditions, such as being pushed or the absence of viable alternatives, prevail) or in terms of the competence or capacities of the agent (where at the time of action or always there are special pressures or the agent is psychologically abnormal, or is simply a child) or in terms of the character of the relation between the injurer and the injured (as when we forgive someone “for old time's sake”).

Our commitment to ordinary participant reactive attitudes toward each other (attitudes that include resentment, and gratitude, and love) does not depend on a denial of determinism, and an acceptance of determinism need not undermine those attitudes. We can add that an acceptance of determinism would not underwrite universal forgiveness—however desirable or problematic such a universal response. While Christianity and some forms of therapy might encourage unbridled forgiveness for the sake of communion, community, and calm, an appropriate resentment may lead to valuable restraint in others and be necessary to justice and (as noted) to self-respect.

What shifts in forgiveness are attitudes, and so blanket encouragements to forgiveness may face special difficulties. If someone seeks to forgive simply to ease his or her own mind, for the sake of self-therapy, whatever kind of closure is achieved might not amount to forgiveness. Indeed, in such a case, one might suspect that the relevant attitude does not really get shifted, it is just the expression of anger that gets suppressed. Since the attitude of the wrongdoer presumably still stands (the therapeutic interests of the aggrieved give no ground to separate agent from act), a response to the affront is always liable to be provoked anew. It is not clear that forgiveness to make ourselves feel better, to free us to move on, is “forgiveness,” that is, a gen-

uine change in attitude toward the offender or the offense. More broadly, the notion that forgiveness is the only way to achieve closure so one can move on is of course mistaken. The notion that one must achieve closure before one can move on may also be mistaken. And the notion that understanding inevitably leads to forgiveness and thus to closure is perhaps least plausible of all. The notion of “closure” is itself problematic when we are dealing with an interplay of attitudes, which by their nature, especially in ongoing relationships, are always in flux.

Attitudes, however, are complex. If our resentments are not simple matters of choice, can forgiveness be? And even if we can, somehow, shift our inner attitude, is such a shift by itself enough to constitute forgiveness? In all circumstances? Pardoning or showing mercy certainly requires a shift in outward behavior; might forgiveness (at least sometimes) require as much in order for the supposed shift in inner attitude to be taken seriously?

Just as it may be difficult to separate offending wrongdoers from their acts, it may be difficult to separate would-be forgivers from theirs: a change of heart in the would-be forgiver without a change in behavior and treatment may not be enough to constitute genuine forgiveness. While attitudes certainly matter, it is not always clear that an attitude can be taken to have changed if one nonetheless demands one’s pound of flesh, insists first (or after) on extracting the full punishment.

Attitudes are not typically under the direct control of our will. So, even if one were persuaded one would be better off if one forgave someone who had trespassed against him or her, one might find oneself unable to forgive. That is not necessarily something (a further something) to blame oneself for: “I am an unforgiving person.” Perhaps the forgiveness is undeserved. Perhaps the offense is in a sense unforgivable (due to its seriousness, its egregiousness, or the depth of betrayal involved). Perhaps the incapacity to forgive is specific to this offense and this offender—not a perpetual unyielding and self-righteous disposition. For example, in the case of a psychopath insensitive to moral rights and obligations, one may not forgive him (where that involves restoring him to full human relations) because it seems more appropriate to dismiss him (regard him as not a moral agent at all). Perhaps there is no ground for separating the agent from the act. (Understanding is not by itself sufficient for forgiveness. Why forgive the unrepentant wrongdoer?) And perhaps the wound itself is of a kind that renders the victim incapable of forgiveness. The interplay of attitudes needs to recognize a third kind of injury. Apart from whatever grievous harm might have been done and whatever morally offensive message might have been sent along with and through it, there is always a risk of moral injury: the person who is the victim of injustice may become capable of injustice in turn (and withal incapable of forgiveness). It is that sort of moral damage that deeply concerned Socrates, and it is the fear of it that sometimes makes abusers and oppressors relentless: they may fear the justified resentment and the

unrestrained revenge of their victims. And all may fear other consequences of such moral damage. Think of the molested child who becomes a child molester. Think of the victims of genocide whose fear of genocide leads them to commit the same crime.

If we cannot simply and directly will our anger and resentment away, steps can be taken, and perhaps sometimes ought to be taken. Therapy depends upon the hope that attitudes can be changed—if not by a direct act of will, by a variety of techniques (which give varying place to reason, thought, and argument). Spinoza's therapy for anger, and for passive emotions generally, involves seeking wider understanding, ultimately *sub specie aeternitatis* (under the aspect of eternity). Modern therapeutic movements are picking up on a philosophical, as well as a Christian, theme.

Spinoza counsels that we avoid as far as possible passive and painful emotions, such as hatred and anger, and points out, among other things, that if we appreciate “that men, like other things, act from the necessity of nature, then the wrong, or the Hate usually arising from it will occupy a very small part of the imagination, and will easily be overcome” (part 5, prop. 10, scol.). In effect, he is suggesting a revision of belief about the operation of causes, so that the object of anger will be seen as just an element of a necessary structure—a change that would inevitably alter the character of the emotion. And the intellectual activity, the search for and consideration of broader causes, is itself a pleasure and so alleviating. Along similar lines, he points out that if we become aware of the multiplicity and complexity of causes, an emotion will have many objects and we will be less affected toward each than if we had regarded one alone as the cause (part 5, prop. 9).

Spinoza's advice, especially the urging to seek wider understanding, contains good sense. Nonetheless, there is a risk of mistaking the sort of necessity Spinoza speaks of as some sort of justifying inevitability. We do not, however, need to reject determinism outright in order to leave room for judgment. We make finer discriminations based on the character of particular causal stories. Every action has an explanation. Only some explanations excuse. Belief in determinism need not provide excuses for the victimizers, nor (despite Spinoza's optimism on this) need it be a comfort to the aggrieved. Of course, when we manage to take a wider perspective on the travails of our life, even if we don't come to regard them as inevitable, we may come to regard them as trivial. Certainly from a God's-eye view, our concerns may seem absurd. But it is not obvious that we always can assume such a perspective, or even that we should if we could.

Sometimes we are able to direct our attention (though there are limits even on this, as many experiences of unrelieved boredom confirm), and choosing a perspective and so perhaps shifting attitude may sometimes be like that. But why should we take God's point of view or think that the perspective of eternity and the universe is somehow more correct than a

more limited perspective? The mere possibility of such an alternative is not enough to make our concerns unjustified—at least not once we recognize that justification must always come to an end. Recognition of alternative views need not leave us with an ironic view of the seriousness with which we take ourselves, when we properly, by our own standards, do take ourselves seriously. After all, what we are looking out onto are our individual human concerns. Such concerns might disappear within some vastly larger picture, but why should a point of view that makes them invisible be thought to make their position (in relation to us) clearer? The concerns remain real for us, and the issue is what is the correct perspective *for us*. (The notion of a “correct perspective” itself determined from no point of view seems unintelligible.) Even if we somehow thought the God’s-eye view the correct one, it seems clear that we could not sustain it. (Aristotle recognized that we are neither simply gods nor animals, though our natures may participate in characteristics of both.) And again, even if we could sustain it, that would not show that what matters to us does not really matter to us or should not matter to us. We love and (yes) we hate, and the reasons of our hearts cannot be simply dismissed just because we can imagine a perspective from which our reasons might no longer move.

The God’s-eye view, like the perspective of determinism, is not really ours. It is not what our attitudes toward others and ourselves depend upon. Perhaps we can look from such a perspective in rare philosophical moments (like looking from the point of view of the stars and seeing the earth as an insignificant little planet), but there is no reason that we should seek to shift from the perspective through which we must inevitably live our lives or give higher priority to an ultimately impossible standpoint.

While enlarged understanding may always offer some benefits, I myself would hesitate to attempt to move permanently into a wholly expansive view, not only because I don’t think one could permanently succeed, but because more particular perspectives seem to me often appropriate. That something might not matter from a God’s-eye view does not mean it does not matter.

#### ACTIVITY AND PASSIVITY IN EMOTIONAL LIFE

As always, one must be careful in thinking about necessity. One often-neglected aspect of freedom depends on a recognition of those constraints that are given—that is, an acknowledgement of the necessities of our nature. An individual human being is not unfree because he or she cannot (unaided by technology) fly. A bird that has had its wings clipped *is* unfree. (Human susceptibility to jealousy is not, in itself, a disabling shackle. It goes with other features of our condition.) Of course, when we mistake social impositions for the necessities of nature, we turn contingencies into constraints. We can be shackled by our own failures to appreciate the differ-

ence. (Hence the value of the sort of questioning of values engaged in by various political pride movements.) One of the great values of the imagination, as of historical and cross-cultural study, is to open us to possibilities that may not be obvious to us in our immediate circumstances. Literary and personal explorations—as well as philosophical ones—can help us to discover real limitations and to overcome illusory ones. Of course, one of the ways to discover limitations is to test them, to try to overcome them.

While we may be subject to attitudes and emotions we experience as given, attitudes and emotions are also subject to change. While we may sometimes intentionally express our emotions, they may sometimes manifest themselves in ways seemingly not subject to our will. It must be recognized that there are degrees of activity and passivity among emotions and within particular emotions—which should not surprise us if we recognize that freedom is in general a matter of degree (despite the insistence of some that it is all or nothing, that we are either existentially free always and everywhere or determinedly unfree always and everywhere, independent of our individual intentions and efforts). Getting our actions to match our intentions, having our desires match our values, experiencing emotions that we can identify with, that flow from what we regard as our genuine natures and attitudes, are achievements—often difficult achievements, and always only partial achievements.

Control over emotions (despite the perhaps wishful thought of Sartre and others who treat all emotion as action), like control over beliefs, is limited. Belief, which aims at truth, is constrained by the evidence we acknowledge. (I think Spinoza, who refused to distinguish a separate faculty of willing in relation to belief, was closer to the truth about the relation of belief and will than Descartes, who insisted error was due to the extension of our will beyond our understanding.) Our responsibility for our beliefs does not end, however, with the limits on our will. There is always the question of whether to act on the beliefs we happen to have and the even more crucial question of what efforts and attitude to take toward gathering evidence in the formation and maintenance of beliefs. (Self-deception has more to do with these matters than with simply and directly choosing to believe that which we would prefer to be true.) All of these complications in relation to belief, given the centrality of belief and thought in emotions, carry over to the realm of emotions, judgments, and attitudes. If forgiveness is forswearing resentment, the question arises of whether (and if so, how) we can choose to forgive. Can we choose not to be angry? At best it seems a process, sometimes involving steps over which we have only limited control. Not that forgiveness is simply a matter of anger management—the interplay of morally appropriate (or inappropriate) attitudes is at stake. A part of the process may involve, as Spinoza suggested, correcting our understanding. A part may involve examining our own desires.

Are desires simply given or are they criticizable and changeable? Certainly desires are modifiable and manipulable—much advertising is based on this fact. It aims to create and to shape desires. Desire can also be made to go away. People fight even addictions—there are a variety of conditioning and other techniques. The question is whether we can modify our own desires, not by self-alienated manipulation, but by reasoning about what is desirable and by the intelligent appreciation of experience. Education depends on that hope, as do self-education and self-development. Asking what is desirable and how we might come to desire it may be the best way to achieve certain sorts of freedom.

The thoughts involved in desires and emotions need not be conscious. Our ordinary concept of thought allows room for a distinction between phenomenological (typically self-conscious and articulate) and explanatory (typically not conscious and not explicit) senses of the term. That distinction does not depend on belief in the psychoanalytic unconscious, though it importantly leaves room for it.

Many thoughts that we ascribe both to ourselves and to others are less than fully explicit without being withheld from consciousness by dynamic forces of repression. Thus, I might explain to the police officer who stops me for jumping a light that “I thought the light had turned green.” In such a case (assuming I am being as honest as I can), I am not claiming that I explicitly thought “the light has turned green and now I can go forward,” any more than when I change gears I (as an experienced driver) have to think explicitly “I am in neutral and must now shift into first.” Such actions are intentional (I don’t, usually at least, shift gears by accident), they are done knowingly, but they do not require conscious explicit spelling out of their guiding thoughts. Indeed, we reach for the thought as an explanation for our behavior only when things go somehow awry in our usually semi-automatic behavior (and when we can rule out alternative explanations, such as a mechanical failure in the car causing it to lurch forward before the light has changed). In sum, we use the concept of “thought” in both explanatory and phenomenological senses. Sometimes we ascribe a thought on the basis of being explicitly aware of it. (That is the phenomenological sense.) Sometimes we ascribe it on the basis of its filling an explanatory need. (Wittgenstein is full of examples of “thought” in this second, explanatory, sense.) And all of this is part of the perfectly ordinary understanding and functioning of the concept of “thought.” The explanatory use of the concept of thought and thinking is essential to self-understanding and our understanding of others.

Spinoza distinguishes between active and passive thought in terms of the explanation of the thought’s occurrence. Imaginative thought tends to be passive: it mirrors physiology and it is dominated by memory and association. It is only explicit thought with a normative and argumentative

structure that is on the side of active emotion. If we are to be free and have control over our emotional lives, we must, according to Spinoza, seek to replace passive thought with active (rational and directed) thought. But the desire to displace imagination and passive emotion risks the elimination of aesthetic and sexual response—for in these areas memory, association, and imagination are essential. By contrast, for the poet William Blake the difference between active and passive does not have these implications. For him imagination involves energy and activity.

Blake understood that, as he put it, “A tear is an intellectual thing.” So, in a sense, are all expressions of emotion. So, in a related sense, are all emotions. Because of this fact—the fact that emotions are discriminated from one another on the basis of and in part constituted by thoughts, beliefs, judgments, and the like—changing one’s beliefs can be a way of transforming one’s emotions. Not that one can simply and directly choose one’s beliefs (that is part of the puzzlement of self-deception); but how one conceives, perceives, and understands the world will in large measure determine how one experiences it. And how one understands oneself will affect who one is. While it is not the case that thinking simply makes it so, in the realm of the mental at least, knowledge affects the thing known. This great power of reflexive knowledge is, as Spinoza understood, what makes room for human freedom.

We generally regard ourselves as having control over what actions we take on the basis of the feelings we are given. But feelings do often seem given. Loving and forgiving are not things that, in their usual forms, we can simply choose to do. They are often thought of as gifts given to us; we experience them and in turn pass them on to others, the objects of our love and forgiveness. But hatred and resentment too typically come unbidden. We may be as passively subject to them as we are to the more positive emotions. And then we may wish to change, and happily, sometimes we may succeed. The fact is that we can both inhibit and cultivate emotion—we can change our minds. Most remarkably, so far as mental states are constituted by our beliefs about them, in the realm of the mental, reflexive knowledge may have the power to transform the thing known. This is rather different from the realm of the physical. A person’s height and other physical attributes remain whatever they are independent of the person’s beliefs about them. But thoughts play a rather different role in relation to emotions. They crucially are what differentiate among, within, and between emotional states. That is, the differences between regret, remorse, guilt, and the like are largely conceptual and not a matter of feelings in the narrow sense of sensations. While bare regret may centrally involve the thought “it would have been better otherwise,” remorse adds thoughts about personal responsibility and about fault—that is, about one’s causal role in the regrettable situation and about the mistake involved being a moral one. As far as sensations go, regret and remorse may feel just the same. Bodily

feelings may mark the difference between emotions and nonemotions, but they are not (in general) what marks an experience as the distinctive emotion it is. And the object of a particular emotion may shift with our beliefs about its cause. Thus anger directed at someone we believe to have harmed us must shift (perhaps to anger directed at ourselves, perhaps to something else) once we are persuaded that we were mistaken in our attribution. Even where irrational states persist (we remain frightened of flying even after being persuaded that flying is less dangerous than forms of transport we do not fear), our bifocal vision must turn what was a naive fear into something else (perhaps now an acknowledged neurotic symptom) because the crucial belief no longer remains simply what it was. And wider understanding—of ourselves, of causes in the world, of what matters—can have wider repercussions.

The situation for expressions of emotion is in some ways similar to the situation in regard to emotions themselves. Expressions of emotion can include obviously voluntary actions (bodily gestures and conventionally symbolic expressions). If we narrow our focus to the apparently universal and apparently involuntary (including some facial expressions), issues of choice and control become more intricate. Is crying an action? This question has a number of different dimensions. Is crying a matter of choice? Is it something we can control? Is it something that just happens to us? These questions are not the same. We can sometimes control involuntary bodily activities that we cannot initiate. In those cases, we do not so much choose them as actions as choose not to stop them once they have started. And sometimes involuntary bodily responses can be actively induced. How they are induced at will is itself an interesting question, sometimes revealing about the normal mechanisms. Actors cultivate various techniques in relation to crying. (Many actors cry by turning their thoughts in sad directions.) Children in general quickly learn the instrumental and manipulative uses of crying. (Do they use a technique to make themselves cry?)

Some bodily states are voluntary, and so especially suitable for the communication of feeling as gestures or facial expressions. Such gestures and expressions can be given culturally variable significance, but because of certain uniformities in our inclinations to respond to standard situations there is some uniformity across cultures. Some bodily states are nonvoluntary, and so while less suitable for the deliberate expression of feeling, they may nonetheless effectively manifest feelings. Indeed, that a certain state cannot be readily called up at will may help it to serve to mark sincerity of feeling. But even nonvoluntary states can often be inhibited at will and sometimes called up at will. Many states are thus neither simply voluntary nor nonvoluntary. Crying is such a state, smiling is another. We can successfully inhibit a smile, or sometimes we may smile despite ourselves, or, more important for present problems, we may call up a smile for a purpose. The purpose may be personal and social, as in a polite smile at a friend's

joke, or even commercial, as in the professional smile of a flight attendant. What does it take to call up a smile or shed a voluntary tear? In particular, while I won't pursue the inquiry here, does the production of an expression of emotion require or involve the production of the feeling or emotion normally (naturally, nonvoluntarily) expressed?

My central interest is moral identity. Whether writing about jealousy, or pride, or incest, or self-deception, I have sought to bring out how particular emotions and particular psychological mechanisms, in light of their conceptual conditions and taking them in their social, cultural, political, instinctual, and developmental contexts, either sustain or threaten our sense of who we are and our hopes for who we might become. Emotions are not simply given, they have their conditions, conceptual and social, and distinguishing among the various conditions and understanding their nature and limits is a step toward controlling our lives. My work is informed and motivated by the Spinozist hope that, because of the peculiarities of reflexive knowledge, understanding our lives can help change them, can help make us more free.

#### NOTE

The discussions and arguments concerning particular emotions such as jealousy and pride, as well as the considerations concerning the expression of emotions, referred to in this paper are elaborated (with full citations) in essays included in Neu 2000. The general theory of the emotions that informs these discussions and its relation to the views of Hume, Spinoza, and Freud are elaborated in Neu 1977. Finally, the discussion of forgiveness here is more fully developed in "To Understand All Is to Forgive All—Or Is It?" in Lamb and Murphy 2002.

VI

Emotion and Value

*This page intentionally left blank*

## Emotions as Judgments of Value and Importance

MARTHA NUSSBAUM

Nun will die Sonn' so hell aufgeh'n  
Als sei kein Unglück die Nacht gescheh'n.  
Das Unglück geschah nur mir allein.  
Die Sonne, sie scheineth allgemein.

Du musst nicht die Nacht in dir verschränken,  
Musst sie ins ew'ge Licht versenken.  
Ein Lämplein verlosch in meinem Zelt,  
Heil sie dem Freudenlicht der Welt.

[Now the sun is going to rise, as bright  
as if no misfortune had happened during the night.  
The misfortune happened only to me.  
The sun sends light out neutrally.

You must not fold the night into yourself.  
You must drown in eternal light.  
In my tent a small lamp went out.  
Greetings to the joyful light of the world.]

Friedrick Rückert  
(text of the first of Mahler's *Kindertotenlieder*)

It is almost impossible to understand the extent to which this disturbance agitated, and by that very fact had temporarily enriched, the mind of M. de Charlus. Love in this way produces real geological upheavals of thought. In the mind of M. de Charlus, which only several days before resembled a plane so flat that even from a good vantage point one could not have discerned an idea sticking up above the ground, a mountain range had abruptly thrust itself into view, hard as rock—but mountains scul[p]ted as if an artist, instead of taking the marble away, had worked it on the spot, and where there twisted about the another, in giant and swollen groupings, Rage, Jeal-

ously, Curiosity, Envy, Hate, Suffering, Pride, Astonishment, and Love.

Marcel Proust, *A la recherche du temps perdu*

The story of an emotion, I shall argue, is the story of judgments about important things, judgments in which we acknowledge our neediness and incompleteness before those elements that we do not fully control. I therefore begin with such a story, a story of fear, and hope, and grief, and anger, and love.

I.

Last April I was lecturing at Trinity College, Dublin. As my mother was in the hospital convalescing after a serious but routine operation, I phoned at regular intervals to get reports on her progress. One of these phone calls brought the news that she had had a serious complication during the night, a rupture of the surgical incision between her esophagus and her stomach. She had developed a massive internal infection and fever, and, though she was receiving the best care in a fine hospital, her life was in jeopardy. This news felt like a nail suddenly driven into my stomach. With the help of my hosts I arranged to return on the next flight, which was not until the following day. That evening I delivered my scheduled lecture, on the subject of emotions. I was not then the same exuberant self-sufficient philosopher delivering a lecture, but rather a person barely able to restrain tears. That night in my room in Trinity College, I had a dream in which my mother appeared emaciated and curled into a fetal position in her hospital bed. I looked at her with a surge of tremendous love and said, "Beautiful Mommy." Suddenly she stood up, looking as young and beautiful as in the photographs of the time when I was two or three years old. She smiled at me with her characteristic wit and said that others might call her wonderful, but she preferred to be called beautiful. I woke up and wept, knowing that things were not so.

During the transatlantic flight the next day, I saw, with hope, that image of health before me. But I also saw, and more frequently, the image of her death, and my body wanted to interpose itself before that image, to negate it. My blood wanted to move faster than the plane. With shaking hands I typed out paragraphs of a lecture on mercy, and the narrative understanding of criminal offenders. And I felt, all the while, a vague and powerful anger—at the doctors, for allowing this crisis to occur, at the flight attendants, for smiling as if everything were normal, and above all, at myself for not having been able to stop this event from happening, or for not having been there with her when it did. On arriving in Philadelphia I called

the hospital's intensive care unit and was told by the nurse that my mother had died twenty minutes before. My sister, who lived there, had been with her and had told her that I was on my way. The nurse asked me to come and see her laid out. I ran through the littered downtown streets as if something could be done. At the end of a maze of corridors, beyond the cafeteria where hospital workers were laughing and talking, I found the surgical intensive care unit. There, behind a curtain, I saw my mother in bed, lying on her back, as I had so often seen her lying asleep at home. She was dressed in her best robe, the one with the lace collar. Her make up was impeccable. (The nurses, who had been very fond of her, told me that they knew how important it had been to her to always have her lipstick on right.) A barely visible tube went into her nose, but it was no longer hooked up to anything. Her hands were yellow. She was looking intensely beautiful. My body felt as if pierced by so many slivers of glass, fragmented, as if it had exploded and scattered in pieces round the room. I wept uncontrollably. An hour later I was on my way to my hotel, carrying my mother's red overnight bag with her clothes and the books I had given her to read in the hospital—strange relics that seemed to me not to belong to this world any more, as if they should have vanished with her life.

## II.

This story embodies several features of the emotions which it is my endeavor to explain here: their urgency and heat; their tendency to take over the personality and move one to action with overwhelming force; their connection with important attachments, in terms of which one defines one's life; one's sense of passivity before them; their apparently adversarial relation to "rationality" in terms of cool calculation or cost-benefit analysis, or their occasionally adversarial relation to reasoning of any sort; their close connections with one another, as hope alternates uneasily with fear, as a single event transforms hope into grief, as grief, looking for a cause, expresses itself as anger, as all of these can be the vehicles of an underlying love.

In the light of all these features, it might seem very strange to suggest that emotions are forms of judgment. And yet it is this thesis that I shall defend. I shall argue that all these features are not only not incompatible with, but are actually best explained by, a version of the ancient Greek Stoic view, according to which emotions are forms of evaluative judgment that ascribe great importance to things and persons outside one's control. Emotions are thus, in effect, acknowledgements of neediness and lack of self-sufficiency.<sup>1</sup> The aim is to examine this view and the arguments that support it, showing how the original Stoic picture needs to be modified in order to be philosophically adequate. In this way I hope to restore to the

philosophical and political discussion of emotion a dimension that has too frequently been overlooked in debates about whether emotions are “rational” or “irrational.”<sup>2</sup>

My focus will be on developing an adequate philosophical account. But since any adequate account in this area must respond not only to the data of one’s own experience and to stories of the experience of others, but also to the work done to systematize and account for emotional experience in the disciplines of psychology and anthropology, I draw on those disciplines as well. Neo-Stoic views have recently been gaining ascendancy in cognitive psychology, in work on helplessness and control,<sup>3</sup> and on emotion as “appraisal” of that which pertains to a creature’s “thriving”;<sup>4</sup> and in anthropology, in work on emotion as an evaluative “social construction.”<sup>5</sup> Since the Stoic view needs to be connected to a plausible developmental account of the genesis of emotion in infancy, I also draw on pertinent material from the object-relations school of psychoanalysis,<sup>6</sup> which converges with the findings of cognitive psychology and enriches the account of the complexity of human history.<sup>7</sup>

Throughout, the explananda will be the genus of which grief, fear, love, joy, hope, anger, gratitude, hatred, envy, jealousy, pity, guilt, and other relatives are the species. The members of this family are distinct, both from bodily appetites such as hunger and thirst as well as from objectless moods such as irritation or endogenous depression. Through there are numerous internal distinctions among the members of the family, they have enough in common to be analyzed together; and a long tradition in philosophy, beginning from Aristotle, has so grouped them.<sup>8</sup>

### III.

The Stoic view of emotion has an adversary: the view that emotions are “nonreasoning movements,” unthinking energies that simply push the person around and do not relate to conscious perceptions. Like gusts of wind or the currents of the sea, they move, and move the person, but obtusely, without vision of an object or beliefs about it. In this sense they are “pushes” rather than “pulls.” This view is connected with the idea that emotions derive from the “animal” part of our nature, rather than from a specifically human part—usually by thinkers who do not have a high regard for animal intelligence. Sometimes, too, the adversary’s view is connected with the idea that emotions are “bodily” rather than “mental,” as if this were sufficient to make them unintelligent rather than intelligent.<sup>9</sup>

The adversary’s view is grossly inadequate and, in that sense, it might seem to be a waste of time to consider it. The fact, however, that it has until recently been very influential, both in empiricist-derived philosophy and in cognitive psychology,<sup>10</sup> and through both of these in fields such as law and public policy,<sup>11</sup> gives reason to reflect on it.<sup>12</sup> A stronger reason for

reflecting upon this view lies in the fact that the view, though inadequate, does capture some important aspects of emotional experience, aspects that need to figure in any adequate account. If we first understand why this view has the power that it undeniably does, and then see why and how further reflection moves us away from it, it will lead to an understanding of what we must not ignore or efface in so moving away.

Turning back to my account of my mother's death, we now find that the "unthinking movements" view does appear to capture at least some of what went on: my feeling of a terrible tumultuousness, of being at the mercy of currents that swept over me without my consent or complete understanding; the feeling of being buffered between hope and fear, as if between two warring winds; the feeling that very powerful forces were pulling my self apart, or tearing it limb from limb; in short, the terrible power or urgency of the emotions, their problematic relationship with one's sense of self, the sense of one's passivity and powerlessness before them. It comes as no surprise that even philosophers who argue for a cognitive view of emotion should speak of them this way: Seneca, for example, is fond of comparing emotions to fire, to the currents of the sea, to fierce gales, to intruding forces that hurl the self about, cause it to explode, cut it up, tear it limb from limb.<sup>13</sup> It seems easy for the adversary's view to explain these phenomenal for if emotions are just unthinking forces that have no connection with our thoughts, evaluations, or plans, then they really are just like the invading currents of some ocean. And they really are, in a sense, non-self; and we really are passive before them. It seems easy, furthermore, for the adversary to explain their urgency for once we imagine these forces as extremely strong.

By contrast, the neo-Stoic view appears to be in trouble in all these points. For if emotions are a kind of judgment or thought, it would be difficult to account for their urgency and heat; thoughts are usually imagined as detached and calm. Also, it is difficult to find in them the passivity that we undoubtedly experience: for judgments are actively made, not just suffered. Their ability to dismember the self is also overlooked: for thoughts are paradigmatic, as it were, of what we control, and of the most securely managed parts of our identity. Let us now see what would cause us to move away from the adversary's view and how the neo-Stoic view responds to our worries.

What, then, makes the emotions in my example unlike the thoughtless natural energies I have described? First of all, they are *about* something; they have an object. My fear, my hope, my ultimate grief, all are about my mother and directed at her and her life. A wind may hit against something, a current may pound against something, but these are not *about* the things they strike in their way. My fear's very identity as fear depends on its having an object: take that away and it becomes a mere trembling or heart-leaping. In the same way, the identity of the wind as wind does not depend on the particular object against which it may pound.

Second, the object is an *intentional* object: that is, it figures in the emotion as it is seen or interpreted by the person whose emotion it is. Emotions are not *about* their objects merely in the sense of being pointed at them and let go, the way an arrow is let go against its target. Their aboutness is more internal and embodies a way of seeing. My fear perceived my mother both as tremendously important and as threatened; my grief saw her as valuable and as irrevocably cut off from me. (Both, we might add—beginning to approach the adversary's point about the self—contain a corresponding perception of myself and my life, as threatened in the one case, as bereft in the other.) This aboutness comes from my active way of seeing and interpreting: it is not like being given a snapshot of the object, but requires looking at it, so to speak, through one's own window. This perception might contain an accurate view of the object or it might not. (And, indeed, it might take as its target a real and present object, or be directed at an object that is no longer in existence, or that never existed at all. In this way too, intentionality is distinct from a more mechanical directedness.) It is to be stressed that this aboutness is part of the identity of the emotions. What distinguishes fear from hope, fear from grief, love from hate—is not so much the identity of the object, which might not change, but the way the object is perceived: in fear, as a threat, but with some chance for escape; in hope, as in some uncertainty, but with a chance for a good outcome;<sup>14</sup> in grief as lost; in love as invested with a special sort of radiance. Again, the adversary's view is unable to account for the ways in which we actually identify and individuate emotions, and for a prominent feature of our experience of them.

Third, these emotions embody not simply ways of seeing an object, but beliefs—often very complex—about the object.<sup>15</sup> It is not always easy, or even desirable, to distinguish between an instance of *seeing x as y*, such as I have described above, from the belief that *x is y*. In order to have fear—as Aristotle already saw it<sup>16</sup>—I must believe that bad events are impending; that they are not trivially, but seriously bad; that I am not in a position to ward them off; that, on the other hand, my doom is not sealed, but there is still some uncertainty about what may befall.<sup>17</sup> In order to have anger, I must have an even more complex set of beliefs: that there has been some damage to me or to something or someone close to me;<sup>18</sup> that the damage is not trivial but significant; that it was done by someone; that it was done willingly; that it would be right for the perpetrator of the damage to be punished.<sup>19</sup> It is plausible to assume that each element of this set of beliefs is necessary in order for anger to be present: if I should discover that not *x* but *y* had done the damage, or that it was not done willingly, or that it was not serious, we could expect my anger to modify itself accordingly or recede.<sup>20</sup> My anger at the smiling flight attendants was quickly dissipated by the thought that they had done so without any thought of disturbing me or giving me offense.<sup>21</sup> Similarly, my fear would have turned to relief—as

fear so often does—had the medical news changed, or proven to be mistaken. Again, these beliefs are essential to the identity of the emotion: the feeling of agitation by itself will not reveal to me whether what I am feeling is fear or grief or pity. Only an inspection of the thoughts will help discriminate. Here again, then, the adversary's view is too simplistic: severing emotion from belief, it severs emotion from what is not only a necessary condition of itself, but a part of its very identity.

Finally, there is something marked in the intentional perceptions and the beliefs characteristic of the emotions: they are all concerned with *value*, they see their object as invested with value. Suppose that I did not love my mother or consider her a person of great importance; suppose I consider her about as important as the branch on a tree near my house. Then (unless I had invested the branch itself with an unusual degree of value) I would not fear her death, or hope so passionately for her recovery. My experience records this in many ways—not least in my dream, in which I saw her as beautiful and wonderful and, seeing her that way, wished her restored to health and wit. And of course in the grief itself there was the same perception—of enormous significance, permanently lost. This indeed is why the sight of the dead body of someone one loves is so painful: because the same sight that is a reminder of value is also an evidence of irrevocable loss.

The value perceived in an object appears to be of a particular sort—although here I must be more tentative since I am approaching an issue that is my central preoccupation. The object of the emotion is seen as *important* for some role it plays in the person's own life. I do not fear just any and every catastrophe anywhere in the world, nor (so it seems) any and every catastrophe that I know to be bad in important ways. What inspires fear is the thought of the impending damage that threatens my cherished relationships and projects. What inspires grief is the death of a beloved, someone who has been an important part of one's life. This does not mean that the emotions view these objects simply as tools or instruments of the agent's own satisfactions: they may be invested with intrinsic worth or value, as indeed my mother had been. They may be loved for their own sake, and their good sought for its own sake. But what makes the emotion center around her, from among all the many wonderful people and mothers in the world, is that she is *my* mother, a part of my life. The emotions are in this sense localized: as in the Rückert poem in the epigraph, they take up their stand "in my tent," and focus on the "small lamp" that goes out there, rather than on the general distribution of light and darkness in the universe as a whole.

Another way of putting this point is that the emotions appear to be eudaimonistic—that is, concerned with the agent's flourishing. And thinking about ancient Greek eudaimonistic moral theories will help us to start thinking about the geography of the emotional life. In a eudaimonistic ethical theory, the central question asked by a person is "How should I

live?" The answer lies in the person's conception of *eudaimonia*, or human flourishing. The conception of *eudaimonia* includes all that to which the agent ascribes intrinsic value; for instance, if one can show that there is something missing without which one's life would not be complete, then that is sufficient argument for its inclusion.<sup>22</sup> The important point is this: in a eudaimonistic theory, the actions, relations, and persons that are included in the conception are not all valued simply on account of some instrumental relation they bear to the agent's satisfaction. This is a mistake commonly made about such theories under the influence of utilitarianism and the misleading use of "happiness" as a translation for *eudaimonia*.<sup>23</sup> Not just actions but also mutual relations of civic or personal *philia*, in which the object is loved and benefited for his or her own sake, can qualify as constituent parts of *eudaimonia*.<sup>24</sup> On the other hand, they are valued as constituents of a life that is my life and not someone else's, as my actions, as people who have some relation with me.<sup>25</sup> This, it seems, is what emotions are like, and this is why, in negative cases, they are felt as tearing the self apart: because they have to do with<sup>26</sup> damage to me and to my own, to my plans and goals, to what is most urgent in my conception of what it is for me to live well.

We have now gone a long way toward answering the adversary, for it has been established that his view, while picking out certain features of emotional life that are real and important, has omitted others of equal and greater importance, central to the identity of an emotion and to discriminating between one emotion and another: their aboutness, their intentionality, their basis in beliefs, their connection with evaluation. All this makes them look very much like thoughts after all, and we have even begun to see how a cognitive view might itself explain some of the phenomena the adversary claimed on his side—the intimate relationship to self-hood, the urgency. But this is far removed from the neo-Stoic view, according to which emotions are just a certain type of evaluative judgment. For the considerations we have brought forward might be satisfied by a weaker or more hybrid view, according to which beliefs and perceptions play a large role in emotions, but are not identical with them.

We can imagine, in fact, three such weaker views, each with its historical antecedents:<sup>27</sup>

1. The relevant beliefs and perceptions are *necessary conditions* for the emotion.
2. They are *constituent parts* of the emotion (which has non-belief parts as well).
3. They are *sufficient conditions* for the emotion, which are not identical with it.

The logical relations among these options are complex and need scrutiny. (1) does not imply but is compatible with (3). (3) does not imply but is

compatible with (1). (1) is compatible with (2)—the beliefs may be necessary as constituent elements in the emotion; but we might also hold (1) in an external-cause form, in which the beliefs are necessary conditions for a very different sort of thing that is not itself a belief. The same can be said for (3): a sufficient cause may be external or internal. (2) is compatible with (3), since even if the belief is just a part of the emotion, and not the whole, it may be a part whose presence guarantees the presence of the other parts.

We have gone far enough, I think, to rule out the external-cause form of (1) and of (3), for we have argued that the cognitive elements are an essential part of the emotion's identity, and of what differentiates it from other emotions. So we are left, it appears, with (2)—whether in a form in which the belief part suffices for the presence of the other parts, or in a form in which it is merely necessary for their presence. What are those other parts? The adversary is ready with a fall-back answer: non-thinking movements of some sort, or perhaps (shifting over to the point of view of experience) objectless feelings of pain and/or pleasure. A number of questions immediately come to mind about these feelings: What are they like if they are not *about* anything? What is the pleasure *in*, or the pain *at*? How are they connected with the beliefs, if they do not themselves contain any thought or cognition?<sup>28</sup> These questions will shortly be reviewed.

#### IV

I must begin a fuller elaboration and defense of the neo-Stoic view by saying something about judgment. To understand the case for the view that emotions are judgments, one needs to understand exactly what a Stoic means when he or she says that; I think we will find the picture intuitively appealing, and a valuable basis (ultimately) for a critique of the familiar belief-desire framework for explaining action.<sup>29</sup> According to the Stoics, then, a judgment is an assent to an appearance.<sup>30</sup> In other words, it is a process that has two stages. First, it occurs to me or strikes me that such and such is the case. (Stoic appearances are usually propositional, although I shall later argue that this aspect of their view needs some modification.) It looks to me that way, I see things that way<sup>31</sup>—but so far I haven't really accepted it. Now there are three possibilities. I can accept or embrace the appearance, take it into me as the way things are: in this case it has become my judgment, and that act of acceptance is what judging is. I can repudiate it as not the way things are: in that case I am judging the contradictory. Or I can let it be there without committing myself to it one way or another. In that case I have no belief or judgment about the matter one way or the other.<sup>32</sup> Consider a simple perceptual case introduced by Aristotle.<sup>33</sup> The sun strikes me as being about a foot wide. (That's the way it looks to me, that is what I see it *as*.) Now I might embrace this appearance and talk and act accordingly; most children do so. If I am confused about astronomy,

I may refuse to make any cognitive commitment on the matter. But if I hold a confident belief that the sun is in fact tremendously large, and that its appearance is deceptive, I will repudiate the appearance and embrace a contradictory appearance. There seems nothing odd here about saying both that the way of seeing the world is the work of my cognitive faculties and that its acceptance or rejection is the activity of those faculties. Assenting to or embracing a way of seeing the world, acknowledging it as true, *requires* the discriminating power of cognition. Cognition need not be imagined as inert. In this case, it is reason itself that reaches out and accepts that appearance, saying, so to speak, “Yes, that’s the one I’ll have. That’s the way things are.” We might even say that this is a good way of thinking about what reason *is*: an ability by virtue of which we commit ourselves to viewing things the way they are.

Let us now return to my central example. My mother has died. It strikes me, it appears to me, that a person of enormous value, who was central to my life, is no longer there. It feels as if a nail has entered my insides; as if life has suddenly a large rip or tear in it, a gaping hole. I see, as well, her wonderful face—both as tremendously loved and as forever lost to me. The appearance, in however many ways we picture it, is propositional: it combines the thought of importance with the thought of loss, its content is that this importance is lost. And, as I have said, it is evaluative: it does not just assert, “Betty Craven is dead.” Central to the propositional content is my mother’s enormous importance, both to herself as well as to me as an element in my life.

So far we are still at the stage of appearing—and notice that I was in this stage throughout the night before her death, throughout the long transatlantic plane ride, haunted by that value-laden picture, but powerless to accept or reject it, for it was sitting in the hands of the world. I might have had reason to reject it if, for example, I had awakened and found that the whole experience of getting the bad news and planning my return trip home had been just a nightmare. Or, I might have rejected it if the outcome had been good and she was no longer threatened. I did accept that she was endangered—so I did have fear. But whether or not she was or would be *lost*, I could not say. But now I am in the hospital room with her body before me. I embrace the appearance as the way things are. Can I assent to the idea that someone tremendously beloved is forever lost to me, and yet preserve emotional equanimity? The neo-Stoic claims that I cannot. Not if what I am recognizing is that very set of propositions, with all their evaluative elements. Suppose I had said to the nurses, “Yes, I see that a person I love deeply is dead and that I’ll never see her again. But I am fine: I am not disturbed at all.” If we put aside considerations about reticence before strangers and take the utterance to be non-deceptive, we will have to say, I think, that this person is in a state of denial. She is not really assenting to *that* proposition. She may be saying those words, but there is

something that she is withholding. Or, if she is assenting, it is not to that same proposition but perhaps to the proposition "Betty Craven is dead." Or even (if we suppose that "my mother" could possibly lack eudaimonistic evaluative content) to the proposition "My mother is dead." What I could not be fully acknowledging or realizing is the thought "A person whom I deeply love, who is central to my life, had died," for to recognize this is to be deeply disturbed.

It is of crucial importance to be clear about what proposition or propositions we have in mind. For, if we were to make the salient proposition one with no evaluative content, say, "Betty Craven is dead,"<sup>34</sup> we would be right in thinking that the acceptance of that proposition could be at most a cause of grief, not identical with grief itself. The neo-Stoic claims that grief is identical with the acceptance of a proposition that is both evaluative and eudaimonistic, that is, concerned with one or more of the person's most important goals and ends. The case for equating this (or these) proposition(s) with emotion has not yet been fully made, but so far it appears far more plausible that such a judgment could in itself be an upheaval. Another element must now be added. The judgments that the neo-Stoic identifies with emotions all have a common subject matter: all are concerned with vulnerable externalities: those that can be affected by events beyond one's control, those that are unexpected, those that can be destroyed or removed even when one does not wish it. This implies that the acceptance of such propositions reveals something about the person: that she allows herself and her good to depend upon things beyond her control, that she acknowledges a certain passivity before the world. This emerges in the complex combination of circumstantial and evaluative considerations that must be present in the relevant propositions.

At this point, it can be concluded not only that the judgments described are necessary constituent elements of the emotion, but that they are sufficient as well. It has been argued that if there is no upheaval the emotion itself is not fully or really present. The previous arguments suggest that this sufficiency should be viewed internally: as that of a constituent part itself causes whatever other parts there may be. I have spoken of the way in which the relevant judgments are a part of the identity conditions of the emotion; however, there is need for further analysis, since it may still appear counter intuitive to make the emotion itself a function of reason, rather than a nonrational, cognitive movement.

Well, what element in me *is* it that experiences the terrible shock of grief? I think of my mother; I embrace in my mind the fact that she will never be with me again—and I am shaken. But how and where? Does one imagine the thought as causing a trembling in my hands, or a fluttering in my stomach? And if so, does one really want to say that this fluttering or trembling *is* my grief about my mother's death? The movement seems to lack the aboutness and the capacity for recognition that must be part of

an emotion. Internal to the grief must be the perception of the beloved object and of her importance; the grief itself must quantify the richness of the love between us, its centrality to my life. It must contain the thought of her irrevocable deadness. Of course, one could now say that there is a separate emotional part of the soul that has all these abilities. But, having seemingly lost one's grip on the reason for housing grief in a separate non cognitive part, reason looks like just the place to house it.

The adversary might now object that this is not yet clear. Even if one concedes that the seat of emotion of must be capable of many cognitive operations, there also seems to be a kinetic and affective aspect to emotion that does not look like a judgment or any part of it. There are rapid movements, feelings of pain and tumult: are we really to equate these with some part of judging that such and such is the case? Why should we not make the judgment a cause of emotion, but identify emotion itself with these movements? Or, we might even grant that judgment is a constituent element in the emotion, and, as a constituent element, a sufficient cause of the other elements as well, and yet insist that there are other elements, feelings, and movements, that are not parts of the judgment. I have begun to respond to this point by stressing the fact that we are conceiving of judging as dynamic, not static. Reason here moves, embraces, refuses; it moves rapidly or slowly, surely or hesitantly. I have imagined it entertaining the appearance of my mother's death and then, so to speak, rushing toward it, opening itself to absorb it. So why would such a dynamic faculty be unable to house, as well, the disorderly motions of grief? And this is not just an illusion: I am not infusing into thought kinetic properties that properly belong to the arms and legs, or imagining reason as accidentally colored by kinetic properties of the bloodstream. The movement toward my mother was a movement of my thought about what is most important in the world; that is all that needs to be said about it. If anything, the movement of my arms and legs, as I ran to University Hospital, was a vain mimesis of the movement of my thought toward her. It was my thought that was receiving, and being shaken by, the knowledge of her death. I think that if anything else is said it will sever the close connection between the recognition and the being-shaken of that experience. The recognizing and the upheaval belong to one and the same part of me, the part with which I make sense of the world.

Moreover, it appears that the adversary is wrong in thinking of the judgment as an event that temporally precedes the grieving—as some of the causal language suggests. When I grieve, I do not first of all coolly embrace the proposition “My wonderful mother is dead” and then set about grieving. No, the real, complete, recognition of that terrible event (as many times as I recognize it) *is* the upheaval. It is as I described it: like driving a nail into the stomach. The thought that she is dead sits there (as it sat before me during my plane ride) asking me what I am going to do about

it. Perhaps, if I am still uncertain, the image of her restored to health sits there too. If I embrace the death image, if I take it into myself as the way things are, it is at that very moment, in that cognitive act itself, that I am putting the world's nail into my own insides. That is not preparation for upheaval, that is upheaval itself. That very act of assent is itself a tearing of my self-sufficient condition. Knowing can be violent, given the truths that are there to be known.

Are there other constituent parts to the grief that are not themselves parts of the judgment? In any particular instance of grieving there is so much going on that it is very difficult to answer this question if one remains at the level of token identities between instances of grieving and instances of judging. We have a more powerful argument—and also a deeper understanding of the phenomena—if we inquire instead about the general identity conditions for grief, and whether there are elements necessary for grief in general that are not elements of judgment. In other words, would we withdraw our ascription of grief if these elements were missing? I believe that the answer is that there are no such elements. There usually will be bodily sensations and changes involved in grieving, but if we discovered that my blood pressure was quite low during this whole episode, or that my pulse rate never went above sixty, there would not, I think, be the slightest reason to conclude that I was not grieving. If my hands and feet were cold or warm, sweaty or dry, again this would be of no criterial value. Although psychologists have developed sophisticated measures based on brain activity, it is perhaps intuitively wrong to use these as definitive indicators of emotional states. We do not withdraw emotion-ascriptions otherwise grounded if we discover that the subject is not in a certain brain-state. (Indeed, the only way the brain-state assumed apparent importance was through a putative correlation with instances of emotion identified on other grounds.)

More plausible, perhaps, would be certain feelings characteristically associated with emotion. But here we should distinguish “feelings” of two sorts. On the one hand, there are feelings with a rich intentional content—feelings of the emptiness of one's life without a certain person, feelings of unrequited love for that person, and so on. Such feelings may enter the identity conditions for some emotion; but the word *feeling* now does not contrast with the cognitive words *perception* and *judgment*, it is merely a terminological variant for them. As already mentioned, the judgment itself possesses many of the kinetic properties that the “feeling” is presumably intended to explain. On the other hand, there are feelings without rich intentionality or cognitive content—for instance, feelings of fatigue, of extra energy. As with bodily states, they may accompany emotion or they may not—but they are not necessary for it. (In my own case, feelings of crushing fatigue alternated in a bewildering way with periods when I felt preternaturally wide awake and active; but it seemed wrong to say that

either of these was a necessary condition of my grief.) So there appear to be type-identities between emotions and judgments; emotions can be defined in terms of judgment alone.

## NOTES

This article is based on the first of my Gifford Lectures delivered at the University of Edinburgh, spring 1993. The subsequent lectures not only offer further arguments for the theory and extend it to the analysis of other emotions but also argue that the theory as stated here needs to be modified in certain ways in order to yield an adequate account of the development of emotion and of the emotions of non human animals. I address various normative questions about the place of emotions, so defined, in an account of public and private rationality. I cannot hope here to provide more than a sketch of those further developments, and, hope that the reader will understand that some questions that may arise about this theory are questions that are addressed later. Despite these drawbacks, I did want to put forward this particular essay as my attempt to honor the memory of Bimal Matilal, not only for its subject matter, but because it is at the core of my work, rather than a peripheral addendum. Matilal was a scholar of profound insight and intellectual courage, whose contribution to philosophy is sui generis, a paradigm of cross-cultural historical and philosophical inquiry. I also knew him as a person possessing great warmth, grace, and wit, whose particularity these abstract terms do not go very far toward conveying.

1. I discuss the Stoic view historically in Nussbaum 1994, chap. 10. Some parts of the argument of this lecture, especially in sec. IV, are closely related to that argument; but I have added new distinctions and refinements at every point in the argument, and, in secs V and VII, have substantially modified my position. Further modifications occur subsequent to the material of this article.

2. Some elements of a related philosophical position are in Lyons 1980, Solomon 1993, Gordon 1987, and de Sousa MIT 1987. None the emotions' cognitive content.

3. See esp. Seligman 1975.

4. See esp. Lazarus 1991; Ortony, Clore, and Collins Press, 1988; and Oatley 1992. For a related view, with greater emphasis on the social aspects of emotion, see Averill 1982.

5. See, Lutz e.g., 1988. See also Briggs 1970.

6. Above all, see Fairbairn 1952, Bollas 1987, and Chodorow 1980; with much reservation and criticism, Klein 1984, and 1985. Experimental psychology, anthropology, and psychoanalysis are brought together in an illuminating way in Bowlby 1982, 1973, 1980.

7. Most of the detailed discussion of all this material is in parts of the project subsequent to this paper; I include the references to convey an idea of my larger design.

8. The word I shall use for the explananda is *emotions*. The Stoic view used the term *pathe*—previously a general word for “affect”—in order to demarcate this

class and to isolate it from the class of bodily appetites. For this reason, the philosophical tradition influenced by Stoicism has tended to use the word *passions* and its Latin and French cognates. To contemporary ears, this word denotes a particular intensity, especially erotic intensity, as the more inclusive Greek term *pathē* did not. I therefore use *emotion* as the best translation and the best generic term—although I shall comment both on the kinetic element that led to the original introduction of that word and also on the element of passivity that is stressed in the Greek term.

9. I believe, and argue subsequently, that emotions, like other mental processes, are bodily, but that this does not give us reason to reduce their intentional/cognitive components to non intentional bodily movements. For my general position on mind/body reduction, see Nussbaum and Putnam 1992.

10. See the illuminating criticisms of both in Kenny 1963, which shows that there is a close kinship between Humean philosophy and behaviorist psychology.

11. We see such views, for example, in the behaviorist psychology of Richard Posner (1990, 1992). Even many defenses of emotion in the law begin by conceding some such view of them—for documentation of this point, see Nussbaum 1993.

12. The Stoics had similar reasons: the adversary's view was represented, for them, by some parts of Plato, or at least some ancient interpretations of Plato.

13. See my Nussbaum 1994.

14. This difference of probabilities is not the whole story about the difference between fear and hope. In my case, where there was both a serious danger and a robust chance of escape, both were possible, and the shift from the one to the other depended on whether one focused on the possible good outcome or on the impending danger.

15. Subsequently, I argue that in the case of animal emotions, and in the case of some human emotions as well, the presence of a certain kind of *seeing as*, which will always involve some sort of a combination or predication, is sufficient for emotion.

16. Aristotle 1991, 11.5.

17. One might argue with this one, thinking of the way in which one fears death even when one knows not only that it will occur but when it will occur. There is much to be said here: does even the man on death row ever know for sure that he will not get a reprieve? Does anyone ever know for sure what death consists in?

18. Aristotle insists that the damage must take the form of a "slight" suggesting that what is wrong with wrongdoing is that it shows a lack of respect. This is a valuable and, I think, ultimately very plausible position, but I am not going to defend it here.

19. See *Rhesorid* II. 2–3.

20. In my case, however, one can see that the very magnitude of accidental grief sometimes prompts a search for someone to blame, even in the absence of any compelling evidence that there is an agent involved. One reason for our society's focus on anger associated with medical malpractice may be that there is no way of proving that medical malpractice did not occur—so it becomes a useful target for those unwilling to blame hostile deities, or the cosmos.

21. Anger at self is a more intractable phenomenon, since it is rarely only about the events at hand; I discuss this elsewhere in my project.
22. On this, see Aristotle, *Nicomachean Ethics*, 1; and for a particular case, IX.9, on the value of *philia*.
23. For the misreading, and a brilliant correction, see Prichard 1935 and Austin, 1961.
24. For a good account of this, where *philia* is concerned, see Cooper 1980.
25. The contrast between such eudaimonistic and more impartialist views is brought out and distinguished from the contrast between egoism and altruism in Williams 1973.
26. As we shall see, “have to do with” should not be construed as implying that the emotions take the conception of *eudaimonia* as their *object*. If that were so, they would be in error only if they were wrong about what conception of value I actually hold. On the neo-Stoic view they are about the *world*, in both its evaluative and its circumstantial aspect. If I grieve because I falsely ascribe to a thing or person outside myself a value he or she does not really possess (Stoics think of all grief as such), I am still really grieving, and it is true to say of me that I am grieving, but the grief is false in the sense that it involves the acceptance of propositions that are false.
27. See Nussbaum 1994, chap. 10.
28. By “cognitive” processes I mean processes that deliver information (whether reliable or not) about the world; thus, I include not only thinking, but also perception and certain sorts of imagination.
29. I discuss this issue in a subsequent chapter of my project.
30. See Nussbaum 1994, chap. 10, with references to texts and literature.
31. It should be stressed that despite the usage of the terms *taking in* and *acknowledging*, this notion of appearing is not committed to internal representations, and it is fully compatible with a philosophy of mind that eschews appeal to internal representations. It seems that neither Aristotle nor the Stoics had an internal/representationalist picture of the mind; nor do I. What is at issue is *seeing x as y*: the world strikes the animal a certain way, it sees it *as such-and-such*. Thus the object of the creature’s activity is the world, not something in its head (or heart). In this essay I proceed as if all these ways of seeing can be formulated in linguistically expressible propositions. Subsequently I argue that this is too narrow a view to accommodate the emotional life of children and other animals, as well as many of the emotions of human adults. And it neglects the fact that other forms of symbolism—music, for example—are not simply reducible to language but have expressive power in their own right.
32. Aristotle points out that such an unaccepted “appearance” may still have some motivating power, but only in a limited way: as when a sudden sight causes one to be startled (but not yet really afraid), see *De Anima* III.9, *De Motu Animalium* II. Seneca makes a similar point concerning the so-called pre-emotions or *pro-patheiai*: see *De Ira* II.3; it is remarkable that Richard Lazarus reinvents, apparently independently, the very same term, *pre-emotions*, to describe the same phenomenon in the animals he observes (1991). The Greek sceptics suggest that one might live one’s entire life motivated by appearances alone, without any beliefs—pointing to

the alleged fact that animals are so moved. But their case is dubious, since, for one thing, it seems to misdescribe the cognitive equipment of animals.

33. *De Anima* III.3.

34. Of course the moment we insert the name of a human being, there is some evaluative content and some moral theories would urge that this is all the value there should properly be, in any response to any death.

## Feelings That Matter

ANNETTE BAIER

### EMOTIONS AND THE IMPORTANT

We all accept the idea that emotions are reactions to matters of apparent importance to us: fear to danger, surprise to the unexpected, outrage to insult, disgust to what will make us sick, envy of the more favored, gratitude for benefactors, hate for enemies, love for friends, and so on. And sometimes the felt emotion can precede knowledge of precisely what the danger, the insult, the nauseating substance, and so on is. Emotion then plays the role of alerting us to something important to us—a danger, or an insult. As I write this chapter, a young man on trial for stabbing his mother to death in the family home (just down the road from where I live), whose defense is insanity, claims memory loss for the time of the murder but says he knows he must have done it, since, quite apart from the overwhelming physical evidence, he has “the guilty sort of feeling, like I have done something” (*Otago Daily Times*, Aug. 29, 2001). This is a rare and doubtless pathological case, but emotions can on occasion play the role of showing us that something important has occurred before we clearly understand what exactly it is.

In such cases emotions alert us to important matters, good or ill. And the emotion itself may at least help constitute the good or the ill. Descartes says all the good or ill of this life depends on the passions. Hume and many other writers about human passions have divided them into the pleasant and the unpleasant, on the one hand those that respond to, alert us to, or constitute goods; on the other hand, ills. There are some purely unpleasant emotions, such as boredom, grief, and guilt, and some purely pleasant ones, such as relief and joy. But as Hume (and Kant) knew, gratitude, although

occasioned by what is a good to us, may be itself unpleasant for a proud person to have to feel, and anger, response to a perceived injury, can be invigorating and releasing, not altogether unenjoyable. Hume would explain such cases by saying that the pleasure of receiving help is mixed with the pain of humility, of needing the help, the pain of being injured with the satisfaction of incipient aggression to the injurer. There surely can be mixed feelings evoked by one event or situation. But some individual emotions, or at any rate states for which we have a single name, while they have a distinctive phenomenological feel, seem to have an essentially mixed hedonic tone—nostalgia, for example. And some, I shall suggest, are neutral in hedonic tone, neither pleasant nor unpleasant. Surprise and interest seem of this hedonically neutral sort, unless boredom is the worst evil. I want to direct attention on an emotion very close to interest, perhaps a variant of it.

Consider this case: a person receives a long distance phone call from a close relative. When she answers the phone the first words her caller says, after greeting her, are “Are you sitting down?” At once she knows that the message to come is of importance, and she feels an appropriate emotional disturbance. As she finds a chair and seats herself, she may reply, “Why? Has someone died?” But she may not jump to that conclusion, and the news may be momentous but good, say that a son listed missing in action has after many years been found safe and well. She certainly feels strongly while awaiting the news that is about to be given her. She will go on, once the news is broken to her, to feel joy or sorrow, but the first feeling seems neither joy nor distress. Interest, concern, anticipation, and nervousness, yes, but more than that, some sort of shock, and intent seriousness. For what she now anticipates is no ordinary good or bad news, unlikely to cause her to need support. Nor is uncertainty alone enough to explain the emotion she feels even before the big news is given her. But what name has this emotion, felt as the important, simply as such? *Interest* seems not quite right, since one can be interested in quite trivial news, or relayed gossip, which one could with no danger receive while on one’s feet. *Concern* in its older sense of “what regards one” would be close, but in its contemporary English sense it is too close to anxiety for the hedonically neutral emotion I am after.

In the case I have sketched, the opening question creates drama, and until the momentous news is given there will be uncertainty. Hume (1973, 441–42) noted that uncertainty itself intensifies an emotion, as does mixture of contrary emotions from simultaneous different causes. His example of mixture, the man who gets, at one time, news both of the loss of a lawsuit and of the birth of his son, resulting in an alternation of extreme joy with extreme distress, can be adapted for my purpose. Suppose this man is waiting for news of both his lawsuit and the delivery of his child. On Hume’s view, the uncertainty will make both fear of losing at court and

hope of a safe delivery especially violent. Suppose a messenger appears, so he at once knows that one of the uncertainties is about to be ended, but not which or how. He will feel this so far nameless emotion, no doubt along with his fear and his hope, and considerable impatience. But in this case he will likely tell from his messenger's face whether that person is the bearer of good or bad tidings, so it is unlike my telephone-call case, in which no fear or hope precedes the call. My adaptation of Hume's example will not be a pure case of an emotion reserved for the important, as such, as distinct from the important threat, loss, insult, enmity, or for the important joy, victory, honor, friendship. And pure cases of a feeling reserved for the important may be quite rare.

#### EXPRESSION OF EMOTION

What made my adaptation of Hume's case impure, as an example of an emotion reserved for the important as such, was both the expectation of getting news and the inevitable bodily expression of sympathetic emotion in the messenger that indicates whether the news is good, bad, or mixed. As Hume emphasized, we do tend to sympathize with each other's emotions, and this is facilitated by the point that Darwin made that we have evolved to share information about what emotions we are feeling by our involuntary bodily expression of them. For, quite apart from sympathy, we need to know if our companions are angry with us, and whether they hate us or not. Do we need also to know what others find important? Well, we will know that, up to a point, by seeing and hearing any of their emotional displays, since all emotions are felt as something taken to be of some importance, something that affects us. But is there a special bodily expression that shows our feeling that we are encountering or considering something of definite importance? What would be the face and posture of the one seating herself to hear news so momentous that she should not receive it while standing unsupported?

She might go pale. Her face would be attentive, that is to say (if we accept Carroll Izard's analysis of the bodily expression of attention and interest [1971, 242]), her lower eyelids may be slightly raised as if to focus better, her lips slightly opened, her chin dropped. The plates of psychologists who, like Izard and Paul Ekman, have catalogued the bodily expressions of different emotions tend to cut their subjects off at the neck, as though the face is all that counts, but, as Darwin knew, stance and movement of arms and hands are also revealing. One thing our nervous receiver of important news will not be doing is shrugging her shoulders, expressing nonchalance. She might assume a crouched protective posture. Even should she later dismiss the news as not as important as her caller deemed it, and shrug off the honor she has received, say, or the blow, as long as she treats what is coming as important, her shoulders will, like her chin, be lowered, to

take on board what she is told and its significance for her. Harry Frankfurt, who has written about the important in our lives, says that we are the beings to whom things matter (1988a, 80), and that seems undeniable. By the same token we can say that we are the mammals with shoulders we can shrug, or lower, to dismiss as unimportant or to accept as important.

It might be agreed that the shoulders have, among their many expressive uses, the particular expressive function I have claimed, namely the acknowledgement of or refusal to acknowledge importance, without agreement that any special emotion is thereby expressed. Our bodies can express our wills' determination, as well as our emotional state. And Rodin could make posture express thinking. So why have I claimed that there is an emotion that is reserved for the putatively important, as fear is for the putatively dangerous? Had there been such an emotion, surely Aristotle, the Stoics, Descartes, Spinoza, or Hume would have included it in their lists. Darwin, who does discuss the shoulder shrug (after quizzing missionaries to confirm that it was universal) takes it to express the antithesis of aggression, to express helplessness, at least when accompanied by suitably disempowering arm and hand movements—that is, elbows in, hands opened outward (1998, chap. 11). Darwin was, in this discussion, interested mainly in the expression of the emotions we share with other animals, and so it is not surprising that he should not have discussed an emotion such as the one I am postulating if it depends on peculiarly human capacities, and that he omitted to consider the shrug's purely general dismissive function, as distinct from the aggression-dismissing function of raised shoulders along with demobilized arms and hands.

Did any writer about emotion recognize this emotion I am discovering or inventing? Aristotle has the *spoudaios*, the person who is serious about things, but this is an ongoing attitude, for the Stoics a virtue, rather than an emotion. Descartes thinks forms of wonder and awe are of great importance, but does not, as far as I know, mention a feeling for the important as such, as distinct from the admirable or the despicable, the providential and the catastrophic. If the feeling I am postulating is the antithesis of nonchalance, we could call it "chalance." (Or, if you prefer, "souciance," the antithesis of insouciance, but that sounds too close to *souci*, worry or concern. The French *soin* may be better.) As nonchalance is temporary lightness of being, chalance or seriousness may be granted to be a temporary state of being bowed down with some weighty matter. (The German *wichtig*, meaning "important," is related to *gewichtig*, meaning "heavy.") But it might be deemed a mood or attitude, even a spell of thoughtfulness. Why call it an emotion?

To answer that, we need to have some general account of what emotions are, and what distinguishes them from pleasures and pains, wants, attitudes, moods, resolves, beliefs. I assume that emotions are felt occurrent mental states with intentional objects, and that, while not themselves be-

liefs, they involve beliefs, or sometimes merely suspicions or wishful thinkings. While not themselves wants or resolves, they tend to lead on to them. Fear, for example, involves the belief that one is in danger from what one fears, and usually the desire to escape it. Emotions are felt, and they are episodic, lasting minutes rather than days. Moods, like them in many respects, are longer lasting and have very vague and general intentional objects, or none at all. Attitudes, like emotions and moods, affect motivation but need not be felt by the one who has them, who may be completely unaware of her attitude. In this last respect attitudes are like beliefs. Emotions, unlike any of these other mental states (except perhaps moods and attitudes), tend to have not just typical physiological accompaniments outside the brain—ones that might, like butterflies in the stomach, be unobservable to an onlooker—but also stereotyped involuntary cross-cultural bodily expression.

Paul Ekman makes this last a necessary condition of calling anything an emotion; to be exact, he writes that nothing counts as an emotion unless there is “a distinctive universal facial expression associated with that state” (1984, my emphasis). Ekman, with this requirement in mind, finds there to be only six emotions: surprise, anger, fear, disgust, delight, and distress. Philosophers’ lists tend to be longer, to include wonder, jealousy, envy, guilt, and shame. I do not think that my thumbnail sketch of what counts as an emotion is very controversial (it is, for example, pretty much in agreement with Helm 2001), but there is some disagreement on how thought-mediated a state can be and still count as an emotion. Most agree with Hume that there must be some “idea” component in an emotion—thus surprise, but not startle (i.e., being startled), counts an emotion. But some theorists, such as Paul Griffiths (1997), refuse to count any beyond Ekman’s basic six as emotions, deeming states like jealousy and guilt too thought-mediated, too brainy, to so count. Emotions proper, he thinks, all involve distinctive physiological changes outside the brain, in blood pressure, muscle tension, and so on, which go with their involuntary bodily expression. But jealousy or guilt or resentment we may keep to ourselves, secret, as it were, in our brains, not secreted in sweat, or other give-away bodily signs.

In postulating an emotional state of chalance, or gravity, directed on what one takes to be of some importance, I am perhaps stretching the admittedly ragbag philosophical category of a passion or emotion, inherited from Descartes, Spinoza, and Hume, but I am keeping the requirement that there be a distinctive feel to an emotion, as well as a distinctive thought content, and some motivational potential. I also assume there will always be some physiological change, something like Descartes’s animal spirits agitating themselves in distinctive emotion-specific ways, leading to some distinctive, involuntary, facial or postural expression that others can read. So, in discovering or inventing chalance, I am much encouraged by the fact that there does in this case seem to be a universal bodily expression, at

least of the admittedly faint negative emotion of finding something of no importance, namely the eloquent “so what?” shoulder shrug. But if you ask me if *chalance*, *gravity*, is not more of an ongoing attitude to what we take to matter, or a resolve to give it due attention, or a tendency to think about it, than any sort of affect, all I can do is ask you if you have not experienced the special feeling, neither especially pleasant nor, like *souci*, unpleasant, a sort of inner settling feeling, with which we encounter, re-encounter, consider, or remember something that matters much to us. (The example I began with involved important news, and so some sense of shock, but *chalance* is more usually felt at unchanging matters of importance to us, not reserved for changes in the landscape of the important.)

Among the things that may evoke such a feeling are moods, attitudes, beliefs, and also other emotions, whose significance may perhaps belatedly strike us. Suppose we hear from a friend with whom we have not been in touch for years. We are pleased, feel delight. Then later, when someone who hears us unaccustomedly singing asks us why we are so cheerful all of a sudden, we realize the significance of our joy, how much that friend, and our feeling for him, matters to us. We accept the importance of the friendship in our life, as we might not have done if not prompted to reflect on it, and react to it. In such a case the felt emotion of *chalance* will be a metaemotion, whose object is affection for the friend with whom we have resumed communication. Emotions and friendships, enmities and angers, can be felt as important, and usually, but not always, their degree of importance to us will correlate with the importance to us of the friend, the enemy, the one we are angry at. Should, however, our anger make us ill, even when we no longer care much about the person who made us so angry, nor want any revenge, the anger may continue to matter much, while its object has come to matter less. We may have to be given drugs to quiet and subdue the crippling anger. Then we will have occasion to feel *chalance* at our anger, take it seriously, while no longer finding the object of the anger so important in our life. But normally, that is to say in non-pathological cases, the emotion will matter only as long as its object does.

#### EMOTION AND THE WILL

Of course when something matters to us we will usually act accordingly. Our plans and goals will usually show what we find important. Recent philosophers who have written about this elusive topic, what matters, have shown rationalist and voluntarist tendencies. Harry Frankfurt speaks of what we will to will, of “volitional necessities,” and of our “investing ourselves” in what we “cannot bring ourselves” not to care about. But he also cites Trollope’s character Lord Fawn, in *The Eustace Diamonds*, every feeling in whose nature revolts against a decision he thought he had taken, and preventing him going ahead with it (Frankfurt 1988b, 183). Feeling may

be what prevents us from disregarding what really matters to us. Frankfurt takes such a case, where feeling revolts against a decision, as still a case of the will, of a (perhaps feeling-prompted) change of mind. He writes, in a later essay, "To care about something is not merely to be attracted to it, or to experience certain feelings. No one can properly be said to care about something unless, at least to some degree, he guides his conduct in accordance with the implications of his interest in it" (1999, 87). For him, Luther proclaiming, "Here I stand. I can no other" is a paradigm of "volitional necessity." Luther's certainly was a case of resolve and action. Is his implacable face the face of the one recognizing what matters? His stance was likely pugnacious as he spoke those famous words, and he likely did feel chalance, feel that the occasion was momentous. He certainly was not shrugging his shoulders.

But not all emotions lead to resolve and motivated action in the way outrage can lead to defiance. Hume thought pride was "completed within" itself (1973, 367). (It leads at most to strutting.) And grief often leads only to helpless laments. There may be nothing to decide, when what is important to one is the loss of a loved one. Once any decision about a memorial is taken, the grief that matters to us may have no outlet in intentional action. When it returns with special intensity on anniversaries of the death, the most one may be able to do, and not always even that, is take fresh flowers to a grave. The importance to us of the loss, and the person lost, will show more in strength of feeling than in any acts of the will. Our will may get involved in coping with the initial grief but is powerless to change the fact of the loss. Depth of emotion, not resoluteness of will, is what will show how much we care.

Frankfurt is not alone in taking our future-directed intentions to be what best shows what we care about. Charles Taylor endorses Alastair MacIntyre's talk of "quests" as showing what we take to be worth caring about, and Michael Bratman (2001) has written that it is a "deep fact about us" that our agency is temporally extended, so that our lives can be structured by long-term plans. It is an equally deep fact about us that our emotions re-echo over time, that grief at loss, guilt at neglect, recur long after there is anything we can do about the lost one or the neglected one, reminding us of what mattered and matters to us, giving us, I suggest, occasion for feeling chalance.

Sometimes feeling may contradict what even acted-on will purports to reveal about what matters to a person. Suppose a person in her sixties, after a good life, decides to risk it by giving a kidney to be transplanted into a younger stranger who will die unless a suitable kidney is made available. This decision reflects, the donor thinks, the fact that her own life expectancy is of no great concern to her, that she is content with the life she has already had. But the night before the surgery she cannot sleep and realizes that her continued life does matter more to her than she thought. Her

emotions show her the truth about how much she cares. Her decision had misrepresented that. She may, if especially strong-willed or noble, go ahead with the risky surgery, but that will not show that her life's continuance is of little concern to her, as she had thought and claimed earlier. Thought and decision, even acted-on decision, can lie about what matters to us, or how much it matters. The emotions the sleepless would-be kidney donor is subject to likely will include anxiety, perhaps regret, or puzzlement at herself, but as she thinks how this may be her last period of conscious thought and looks back on her life and forward to her death, she will, if I am right, also feel weighted and grave. She will feel chalance.

The person who really does not care about her life ending could spend the night sleeping soundly (as King Charles I of England supposedly did before his execution), or, if wakeful, reading an amusing book, or joking and clowning around with friends. Or she might calmly update her will. She need not shrug off the risk she is taking, in the sense of denying it to be real, but she might, if really content to die, shrug her shoulders when others call her act self-sacrificial. Such a nonchalant attitude to one's own end is not likely to be often found. (We might admire it if we found it. Hume, when he wrote in his *History of England* about Charles's death, clearly admired the royal calm. When his own death approached, he too was calm, almost lighthearted. I do not want, by focusing on seriousness about something, to agree with the Stoics that the serious person is morally better than the one who has cultivated nonchalance on matters most people find weighty. Nor were either Charles I or his later historian, Hume, without some due seriousness in preparing appropriately for their own deaths. Charles's last word was "Remember!", spoken to Bishop Juxon, whom he had instructed to implore the absent crown prince to forgive his father's killers. We admire both the serious preparations Charles made and his sound sleep. We admire both Hume's serious attention on his deathbed to the posthumous publication of his dialogues, themselves not without sly satire, and his ability to joke with his visitors almost to the last.)

#### WHAT REALLY MATTERS

I have said that it is our emotions, or lack of them, that will speak the truth about if and how much something matters to us. It may be objected that emotion too can surely be wrong about that. Fears can be exaggerated, even sometimes self-fulfilling, anger crippling, envy unbased, pride vain and silly. (Buddhists supposedly recognise 84,000 dysfunctional emotions, and as many antidotes.) In cases like these, the danger, the insult, the cause of envy, the honor or accomplishment in which foolish pride is taken, is not important enough to justify the person's felt emotion. Or we may later find our earlier mild reactions too muted. Thought and reflection, perhaps after discussion with others, may correct what the initial emotion got wrong. If

there is a special feeling, gravity or chalance, which is directed specifically on what matters to a person, can it not also be wrong, and need correction?

This postulated feeling that something is of great importance to us, and its antithesis, will usually be among the most thought-mediated and reflective of our emotions. Indeed it will typically come into play when other earlier emotions are self-criticized. One may later shrug off the accomplishment in which one earlier felt exaggerated pride. Or one may feel, on reflection, that a past insult should have angered one more than it did at the time. How much trust should we put in our feelings about what matters? The relative who phones to tell someone of the safe return of a lost son, or the finding of his corpse, will have no doubt that the news matters to her hearer and is surely right about that. Many beliefs, memories of earlier communicated anguish, and sympathy all feed into her request that her hearer seat herself. And the hearer infers from that request that something momentous is about to be revealed. The emotion she feels as she awaits the news is inference-based and imbued with trust concerning how well her caller can judge what will matter to her. Then when she gets the good (or bad) news, let us suppose that her son is found and safe, its impact will be mediated by all her past anxiety.

What matters to a person stays in the mind, and memory preserves what relates to that with particular tenacity, as experts on improving one's memory are well aware. What is of little concern to us we tend to forget quickly. (How many of your past shoulder shrugs can you recall?) What stays in the memory and keeps resurfacing to the forefront of attention is what mattered and matters. But it is said that memory can lie, and so, it might be suggested, can our feelings about what is and is not of importance, which affects what memory retains. Of course any emotion based on a false belief or unsound inference can be in that derivative sense false. If the caller who asks her hearer to be seated before she continues goes on to tell her a joke so funny that she might have fallen over laughing, the feeling of chalance will have been misplaced. Jokes, however good, are not occasions for that. But if there is no mistaken factual belief, nor faulty inference, can the feeling itself mislead us? Can we not attribute, on the strength of it, too much importance to something, exaggerate or underrate how much it matters?

Were there such a thing as objective mattering, God's-eye or, rather, God's-shoulder mattering, and were that reliably communicated to us, then our personal findings of importance could be said to be correct or incorrect in comparison with the divine standard. But for nontheists, the most we can expect is that criticism of personal findings of relative importance may come from later such feelings, and from spokespersons for cultural priorities. We may grant the adolescent that he does not care about tidiness but try to get him to care. In our rhetoric with him we may well say things like, "You are wrong to think tidiness does not matter—it matters to us

who live with you.” In my childhood there was a nasty little song that ran, “Don’t care was made to care. Don’t care was hung. Don’t care was put in a pot and boiled till he was done.” This indicates how we try to change what matters to a person, when we do. We work on what already matters to them—in this case whether or not one gets to be hanged and boiled. We do manipulate, as best we can, other people’s feelings about what matters. But that does not really establish that the changed feelings are more correct, in any other sense than more politically correct, than what they superseded. We can change our minds, or, if I am right, our hearts, about how much something mattered, but that is what it will be, a change. We update our priorities, but the later ones, even when better informed, need not be any wiser than the earlier. Wisdom is a good sense of what matters more than what, but it takes it to discern it. Our criterion of relative wisdom will keep up with our changes of mind about what matters most. We will disapprove of too frequent reversals or fluctuations in our evaluations of what matters. Vanessa Bell is reported to have refused to go to social occasions at which formal dress was expected, since such grand parties “changed one’s values” in unwanted ways. There is a kind of integrity in not having one’s version of what matters to one change too easily with change of scene, or of company.

To appreciate the sense in which a person’s feelings about what matters to him at a particular time are the final word on that, consider a person who faces a driving test on his birthday and refuses any celebration until the test is behind him. He may seem, to those close to him, to be taking the matter unduly seriously. They assure him that he is well prepared for the test, that he is a good driver, so has no need to worry. He may reply, “I admit I am a little nervous. Passing this test is very important to me.” Should there be thought to be any real chance that he might fail, we might tell him that it was not the end of the world if this happened. But it might be the end of his world, the world he wants to continue in. For if the birthday he refuses yet to celebrate is his eighty-fourth, and if failure to pass this particular test would mean the end of his driving life, we might have to agree with him that the test was a serious matter. His mobility and independence would be at stake. We would still try to point out that reduced versions of these undisputed goods might still be available to him, but we would not be correcting or challenging his feeling that such goods matter very much. Aging may bring new things into a serious light—renewing one’s driving license in one’s eighties cannot be taken as lightheartedly as it might have been earlier, for the experienced driver. But the older person’s sense that mobility and independence matter is continuous with the toddler’s and the adolescent’s valuing of them. It does not take the wisdom of age to discern their value, merely to realize more vividly how temporary our hold on them may be. Should our man fail the test, his life will be seriously the worse, however stoically he adapts to his reduced style of life.

To say that our feelings about what matters have the final word is at the same time to say that feelings decide value. As Hume said of some passions, feelings of chalance “properly speaking, produce good and evil, and proceed not from them” (1973, 439). Values by definition matter, and how much more one thing matters than another determines what comes before what in our long-term plans, what stands out in our retrospective surveys of our lives, what images return, and what emotions resound. What matters is what we mind about, have minded about, will mind about. Charles Taylor (1989, 23) has criticized as “naturalist illusion” any Humean account of moral or other value that takes it to be simply the projection of our own passions, however reflective the passion. For Taylor, there must be “hyper-goods,” discerned in “strong evaluation.” The values thus discerned, he says, are “not rendered valid by our own desires, inclinations, or choices, but rather stand independent of these and offer standards by which they can be judged” (4). Now offering standards by which desires, inclinations, and choices can be judged is one important role of emotions such as regret, remorse, shame, sorrow. A feeling can be what prevents Frankfurt’s decision taker from bringing himself to carry it out. But Taylor wants higher than human standards to validate human evaluations. He wants “ontological frameworks” within which to find “spiritual values.”

There is a persistent tendency in philosophy to make a mystery out of value, and the word itself, by this point in its history, may encourage this, at any rate more than “mattering” does. Even G. E. Moore, guru of the Bloomsbury group, for whom the adjective “important” was a favorite term of appraisal (at least according to some sour critics, such as Ethel Smyth, to whose musical compositions the “Bloomsbury word” was apparently not often enough applied)—even sensible Moore, who, as Woolf ironically says, “made us all wise and good,” finds something nonnatural about goodness. He took value to be discerned by “intuition,” whatever that is. (Woolf, persuaded by her male Cambridge friends in 1908 to read *Principia Ethica*, reported in a letter to Clive Bell that it caused in her brain “a feeble disturbance, hardly to be called thought.”<sup>1</sup> The moral epistemology of Moore’s book, I have to confess, had a similarly faint effect on me, and certainly did not make me wise and good.)

I am offering the feeling of chalance as a naturalist alternative to Moorean intuition, Frankfurtian metawillings, and Taylorian strong evaluation. But I find myself in belated agreement with Taylor in thinking that it may take upright posture to communicate the sort of evaluations of what does and does not matter, that we are familiar with, in ourselves. Having earlier scoffed at Taylor’s stress on the way we walk, I am now finding the shoulders of we relatively broad- and mobile-shouldered mammals the means by which we communicate our findings of what does and does not matter—

and communication matters, if anything does. (I do not, of course, want to make the implausible claim that only to we who can shoulder things can anything matter. For all expressed emotions show something about what matters to us, and to other animals. My cat's tiny shoulders are too incipient for shrugging or lowering without lowering her whole front body with them, but she leaves one in little doubt what does and does not matter to her. She may not feel chalance, but her walk can seem to be nonchalant, she can turn a cold shoulder, and she can dismiss things by turning her back on them.)

Thomas Nagel (1979, 23) writes that "if there is reason to believe that nothing matters, then that does not matter either." Is he shrugging his shoulders as he communicates this very logical conclusion? He says that the one who thinks there is reason to think that nothing matters will live his life with irony, rather than despair. But irony is not the same as the dismissal of importance that I have taken the shrug to signify. The ironist will wear a faint world-weary smile. What we smile and laugh at is usually of some importance to us. Laughter can be used to mock others' priorities and solemnities, but as Freud knew, the objects of our amusement are of some importance to us. We may laugh when a pompous man slips on a banana skin, but we would mind if we were the ones who slipped, and it is because we know what people mind that we find the spectacle funny. Our jokes can reveal our deepest concerns. In any case, few could honestly say that nothing matters to them, that they take nothing seriously.

A slightly different variant on Nagel's nihilist's question is whether mattering itself much matters. Should one trust any of one's feelings or judgments about what does or does not matter? One seems doomed to trust at least one of them, even if it is the judgment that mattering does not matter, only fun does. If on reflection one finds all one's previous findings of importance exaggerated, perhaps shrugging off one's old concerns, or even swearing off any use of the Bloomsbury word, this is usually because some new concern, say the danger of a world war, makes the old fade into insignificance, or the applicability of some new term of appraisal, such as "cool," comes to matter more. Mattering is the ontology of minding, and what we mind does often change over time, even when fickle fluctuations are not evident. What we mind may shrink in scope as we age, but total apathy will be a rare, and usually a pathological condition. What is more likely is that, over time, while some things cease any longer to matter, new, but not altogether new, things come to matter. (It has taken me seventy-two years to find the shoulder shrug important.) But changed values and priorities will usually show their genealogical links with earlier concerns, as well as show their cultural inheritance. I inherit an interest in the expression of emotion from the authors I have read, from Descartes, through Hume, to Darwin, Izard, Ekman, and Eibl Eibesveldt. And my interest in arms and shoulders could be traced back to my mockery in my APA pres-

idential address of Charles Taylor's emphasis on our two-leggedness, our upright, armed, dignity-affording walk. However spiritual, or ironic, our transvalued values, they will pick up on our earlier values, either by refining them, or by vehement denial of them, or by humor at their expense. Birth, death, birthdays, anniversaries of deaths usually continue to matter, even to revolutionaries, terrorists, and subversives. Black humor, defacing gravestones and disinterring graves, is a backhanded agreement with the conventional majority that graveyards are places of importance, that death and the rituals of death matter. Even those who cheered when they heard about, or saw on television, the fiery collapse of the Manhattan World Trade Center Towers into a monster graveyard, showed that they knew the significance of those thousands of deaths.

In an earlier essay about emotion ("What Emotions Are About," 1990, discussed and criticized in Alanen forthcoming), I made the Freud-influenced claim that emotions typically have "depth" and tend to reenact earlier occasions for that sort of emotion—our adult loves to pick up on our infant loves and so on. Perhaps it would be more correct to say that these primal experiences of emotions on the human range set the agenda for later occurrences of that sort of emotion, sometimes by repetition, sometimes by violent rejection of earlier values. Is there a primal shoulder shrug, a primal mocking laugh? None of us began by finding it a joke to deface a gravestone, but we may as children have giggled at solemn funeral services, as an outlet for confused emotions. So we can, up to a point, understand the strange and offensive behavior of the cemetery wreckers, even of those who rejoiced at the suicide terrorists' spectacular successes. To them, as to us, death matters, has emotional charge.

For most of us the question will be not if anything matters, but rather how much various things that may compete for our attention matter. The relative strength of our reflective feelings about them, what I have called our feelings of chalance, what we give weight to, and what we shrug off, will decide that. This subjective feeling, or its absence, will not settle what if anything really matters, only what matters to us now. And as Descartes (2:103) wrote, "What is it to us that someone should make out that the perception whose truth we are so firmly convinced of may appear false to God or some angel, so that it is, absolutely speaking, false?" What matters to us is what we and those we can be in touch with take to matter. My concern here has been our everyday feelings about what matters, and our communication of such feelings. I offer for your attention what we accept or reject as having weight, what we, not Atlas or Sisyphus, let alone Zeus or Jehovah, shoulder or shrug off.

#### NOTES

A version of this paper was prepared for the conference "Passion, Thought, and Virtue" at Uppsala University, Sweden, October 2001, to mark the sixtieth birthday

of my friend and critic, Lilli Alanen. Changed circumstances, in my family and in the world, conspired to force me to cancel that journey. My talk about the important, along with its planned revisions after critical discussion at the conference, was overtaken by the indisputably important, the tragic and massively disruptive events of September 2001.

1. See Woolf year, letter dated August 13, 1936. The reference to Ethel Smyth and "important" as "the great Bloomsbury word" is in letter 3160, to Ethel Smyth (63); the reference to Moore's *Principia Ethica*, "the book that made us all so wise and good," is in letter 3610 (there's a coincidence!), to Judith Stephen (400).

## Perturbations of Desire

*Emotions Disarming Morality in the “Great Song” of  
The Mahābhārata*

PURUSHOTTAMA BILIMORIA

For Renuka

Recently there has been a good deal of interest in comparative philosophy in the treatment of emotions in non-Western cultures. An issue of *Philosophy East and West* (1991) and, closely following that, a volume of essays (Marks and Ames 1995), were dedicated to the philosophical debate surrounding the nature and ethics of emotion in South and East Asian thought. These forays have addressed the standard questions (even if somewhat alien to these traditions), notably whether emotion is comparable to or competes with reason in moral and social discourses, whether emotion plays a negative or positive role in the “good life,” whether emotion is in any sense itself “intelligent” and therefore an aid to practical judgment, and whether much significant conceptual strides have been made toward understanding the cognitive, conative, and cultural dimensions of emotion.

Such questions have preoccupied South Asian philosophers from very early times. Indeed, on the broad contours, there are immense parallels between ancient Greek thought on these matters (e.g., Stoic) and classical Indian thinking (e.g., Jaina, Buddhist). But within the large corpus of Indian philosophical and literary (but also psychology and popular folk) sources, there is vigorous rivalry in respect of the theoretic approach and systematic treatment of the complex range of issues that doubtless arise in such an inquiry. Curiously however, the competing theories (that straddle mythico-ritual aesthetics, dramaturgy, rhetoric, epic lyricism, and moral discourses) tend rather to reinforce two polarized stereotypes in the Western gaze: “Indian tradition presents no deep concept or historical traces of what

an emotion is," and "The religion and philosophy of India keep its people subsumed to an excessively emotional and therefore irrational form of life."<sup>1</sup> However, the current direction of research in South Asian philosophy points to the view that, unlike knowing, recognizing, calculating, and other cognitive occurrences, a basic emotion is to be understood not as a singular episodic experience but on a par with aesthetic reflexivity (*rasa*), as a mix variously of sentiment (*bhāva*), affect, feelings, evaluative sense or, better, "moral reserve," and—most significantly—the raw stirrings of desire (*icchā*, *kāma*, *kalpanā*, *bhāvana*), which seeks conscious expression by a transformation of the "stuff" of consciousness itself, namely *manas* (mind as "inner-sense," which in its more refined and cultivated ex[*in*]tension is called *buddhi*, or discriminating intelligence). The horizon of emotion that emerges here would appear to be somewhat more radical than those aligned to the triadic divide of the cognitive, conative, and affective, or other departures. I shall explore this refreshing theoretical position and briefly elucidate its reflective insights in part 1 of this essay. I will illustrate these using "narrativized emotion," with a paradigmatic scenario in an epic literature, the *Mahābhārata*, that highlights this line of thinking in part 2. The final section, part 3, will argue for a closer link between desire and emotion, and it draws out some key philosophical and moral ramifications of this view.

## I.

Discussions of emotion tend to begin with the body as the first stage of emotional enactments. (This is probably why drama, theater, and sublime performatives cannot make do without embodied representations on the stage.) The body is the *locus classicus* of feelings, sentiments, and affects, and life of the mind or consciousness represents the challenges, exigencies, and stresses of the external world impinging upon the inner phenomenological world. The body marks the limit of the self, if not of the world too. The senses travel out of the body (in one theory literary as agent-extensions of the *mind*, to make contact with the world of properties and objects, movements and projected affects, and so on). The mind that sits in (or pervades) the body as an inner sense-faculty (*sensus communis*) assimilates the sensations, impressions, images, and ideas, appropriating them as its own intentional objects and structural or modal transformations, and it sorts the experiences into their different categories for reflection, possible action, storage in memory, or sheer enjoyment of the self. Some encounters and the ensuing experience may give rise to physiological reactions and sensations overwhelming the body: there may be quivering, shivering, giddiness, nervousness, heaviness of breathing, blushing, and infirmity of posture. One also refers to this, especially when appraising or reporting her overall state of being as "feelings." Some feelings are engendered purely by an individual's state of mind at a particular time, expressing joy or hap-

piness—hence there is a smile and harmonious bodily gesticulations; but when there is utter confusion, the facial expression or body movements appear fear-stricken, anxious, panicky, moody, fickle, ambivalent and indecisive. The *Mahābhārata* identifies the body as the vehicle of pleasure and pain (1972–80, chap. 174, 21–22).

To be sure, however, “feelings” in themselves are not considered in Indian thought to be emotions in any definitive sense. The feelings are not of themselves intelligent or intellectual states; rather they are manifestations of “movements in [or of] the body,” not unlike the thirst or hunger that accompanies depletion of fluid and solids in the body. And these feelings might only be transitory, as the *Bhagavad Gītā* puts it: “Material touching (sense-object contact) that gives rise to cold or heat, pleasure (*sukha*) or pain (*duḥkha*), comes and goes, it is transient, and is to be endured.”<sup>2</sup>

The background taxonomy presupposed in this analysis comes from the ancient *Caraka-saṃhitā* (India’s *cyclopaedia medica*), which operates on the (*yin-yang*-like) duality of the harmony and disharmony of bodily humors (flatulence, bile, phlegm) that result in pleasure and pain. In themselves these physiological movements or modes (when stable) are not emotions but rather are their correlates or symptoms, according to the following seven “insights” into the nature of emotion.

(1) There sometimes develops, according to the theory of the ancients, an attachment to pleasure or an aversion to pain. It is then that mere feelings become transformed into a corresponding mental mode (with one or a mix of three attributes, namely *tamas*, “heaviness/dross”; *rajas*, “strident/restless”; or *sattva*, “lightness/purity”),<sup>3</sup> such as excitability in the case of pleasure and perhaps fear and despondency in the case of pain. This distinction between feelings and emotions appears to be quite categorical. In Buddhism, feelings (*vedanā*) also mark the “entry points” at which dispositional states get converted into emotions (*sankhāra*): for instance, pleasant feeling may rouse latent sensual attachment (*rāgānusaya*), painful feeling may rouse anger (*patighānusaya*). And this has implications for the analysis of their differential nature and treatment.<sup>4</sup>

(2) Emotion invariably registers a sentiment (*bhāva*) that is not unlike what June McDaniel has curiously called “inspired thought.” This *bhāva* is born from a crossing of “culture and personality.”<sup>5</sup> The *bhāvas* are almost invariably subject to refinement and cultivation, or attempted suppression, in any given cultural or ideological environment. The context-dependence and historical, or, in the case of literate cultures, intertextual development of certain “extraordinary” (*alaukika*) emotions (especially those of the refined aesthetic categories, which possibly follow from a single “ordinary” impulse, such as erotic love, quiescence, or aggression) has been increasingly iterated in Indological inquiry.<sup>6</sup> That is to say, within the context of a particular tradition, religion, or spirituality there is an evocative disposi-

tion that embeds a certain quasi-intelligent response, which is not entirely cognitive nor entirely reducible to feelings as sensations, but rather something midway between reason's thought and an inspirational "heartfelt feeling." (These are developed in the next few "insights.")

Since *bhāvas* as a range of inspired sentiments embrace or are reflective of certain values and virtues valorized in the culture, their personal experience or interiorization could lead to pleasurable or painful sensations, or the sentiments of empathy and pathos. There can of course be detached appreciation of the *bhāvas* in depersonalized context where the subject is a distanced observer identifying with a dramatic artist, as in theatrical arts or classical dance. It is instructive to draw attention in passing to the theorizing on dramatized emotion in Sanskrit aesthetic criticism for the light it throws on basic or "concrete" emotions emulated or transformed and interjected in the spectator-audience for their edification. The *bhāvas* thus dramatized from a possible personal experience of an individual, now not present, and universalized for a generic appeal powerful enough to induce or elicit a similar or synthesized emotion, are called *rasa* (literally, flavor, taste, relish, mood). Up to eight kinds of "concrete" *bhāvas* (*sthāyin*) and about a dozen synthetic modulations are enumerated with great relish in working up theories of performative aesthetics, dramaturgy and poetics.<sup>7</sup> When universalized, a *bhāva*, such as love (without or without an erotic overlay), is also said to be liberating, as the emotion does not keep the experiencer fettered or bound to one subject but rather allows her to "spread her wings" or to embrace a far wider subjecthood or domain, as in the empathic compassion for all living creatures that the Buddha elevated to a central moral emotion in his teachings.

(3) Returning to basic emotions, emotion is characterized as being *intentional*—that is, an emotion is directed toward some object or event. A mere feeling may be an elementary sensation that is not intentionally directed toward some object or other, and it may be a mere reflex or "knee-jerk" reaction at first blush.<sup>8</sup> Similarly, fear and disgust have identifiable objective correlates, even if the precise objects of these *bhāvas* remain unnamed or uncharacterized. When one hears the cry, "Wolf! Wolf! Beware . . . run!" the fear that wells up is not identifiable with the "meaning" of the utterance heard as such but is a sentiment in which this meaning transparently tends toward a reference (*viṣayinvastu*) that is in this instance an animal known to cause harm. Intentionality is a very significant feature of emotion in this theory, on the epistemic motivation that a psychic or inner disposition need not be reduced to itself, hovering without clear regard to its subjective and objective contents (*prakāratā, viṣayatā*).

(4) This affective directness or intentionality is marked by *concern*, rather than appearing as some sort of abstracted and detached apprehension, as in impersonal cognitive intentionality. This defines the ambience in which the object or situation is held or is beholden to the subject as par-

ticipant in the experience. The example in the above paragraph illustrates this aptly, as in addition to assimilating the propositional content of the utterance—that is, a linguistic-cognitive content capable of being represented in detached analytic symbolism—there is an extracognitive immediacy (*parokṣabhāvārtha*) that has the function of raising an alert, cautioning the other to be careful, to flee to safety. Another example is the aura of romance surrounding the attention of a loved one. If the object of attention is threatened, one becomes anxious in relation to the object (another subject), perhaps even paranoid. Indeed, frustration of the object of desire (*icchā* or *kamā*), to which one has attachment (*rāga*) or a bond results in dejection, or longing and pining.

(5) Implicit in emotion is a deeply subjective dimension with a strong *evaluative* sensibility, from which certain judgments about good and bad, desirability or undesirability, approval and disapproval are projected onto the object, the act, the “other” event or situation in the awareness field. One might call this evaluative aspect a *belief*—that is, the subject believes some state of affairs to obtain. However, I do not see any compelling reason why a strict *cognitive* model of belief, with the strictures of truth-conditions, has to be extended to emotions as characterized here. This might arise as a second-order judgment when passion has delivered its own verdict to reason, as when one rhetorically thinks: “I am an idiot; why did I react in anger? But I did feel right about it at the time . . .” This evaluative stance is considered to be a significant ingredient of emotion as it is an important (though not the only) way by which a culture grounds its citizens’ responsive sensibility toward personal and social values, ethical imperatives, shared experiences, and moral recollections of tradition (or transcendental valuing). For this reason moral dilemmas and conflict of values are able to evoke such strong emotions; conversely, emotions articulate these moral perturbations, and this is the theme of the next insight.

(6) The point preempted above is that emotion reflects the moral repertoire of the community or culture of which the person is a significant member, in that conflicts may begin to surface when there is clash between an emotional response or sentiment (*bhāva*) in the subject and certain moral principles enshrined in the culture (the “horizon” or “background moral knowledge” of the tradition). Thus, suppose that after the Buddha and by the time of the epic tradition (or even as late as Gandhi facing the colonialist General Dwyer), the principle of *ahiṃsā* or noninjury becomes part of the cultural self-understanding and the moral order, but the calling to war is considered to be a great virtue and the need of the day: *duty as virtue*. As a result, a citizen may feel immense tension welling up inside her because her emotional response suggests a negative moral evaluation of the consequences of war, the value of which she thought her community had valorized and which she had indeed internalized. The absent loci of a

strong corresponding *bhāva* for the latter virtue (*ahiṃsā*) and its surrogate *bhāva* of nonaggression are being usurped by the more threatening virtue of *hiṃsā*, violence, with its surrogate *bhāva* of aggression; this inversion may now stand poised as the “intimate enemy,” the scourge of the “bounds of righteous duty.”

(7) There is a show of altruistic compassion (*dayā, lokakṛpā*) reinforced by seemingly dispassionate but patently consequentialist appeal to adverse outcomes if some other sentiment or “calling” is heeded rather than the one “appropriate” to the occasion. Thus the warrior-citizen described above might be led to construe her existential state as being one of intense depression and she is therefore provoked into fearing the consequences of the pending encounter. The dispassion may be expressed through a story about the disastrous state of affairs the society will be plunged into in the aftermath of the battle. In other words, utilitarian and consequentialist considerations are searched out and appealed to in order to reinforce the fledgling moral judgment, but more importantly to determine the correct emotional response to the case in hand. How should one act under the circumstances? An engaged response might be the first step in working out the procedural directive, by allowing greater scope to “intelligent sentiment” than a mere reliance on the cool, detached, *dispassionate* aloofness of reason. The “rational,” or what would be considered reasonable in real-life responsiveness, is not a prerogative of reason alone (as in most theories of rationality, or economic rationalism).<sup>9</sup>

## 2. MORAL DILEMMAS AND EMOTIONS

I will now introduce briefly a paradigmatic illustration from the battlefield scene in the epic *The Mahābhārata*, which is narrated in the book famously known as the *Bhagavad Gītā*, which I shall render as the “Great Song.” My intent is to explore the ethical discourse undergirding the narrative in the background of the general points considered under the seven insights outlined in part I.

Moments before the assault is launched, the warrior Arjuna shows signs of fatigue and loss of strength, letting the powerful *Gāṇḍiva* bow slip from his hands. His half-muted request to brake on the wheels takes Krishna, his charioteer friend, by surprise. Arjuna is palpably troubled by something (the intentional object) and his judgment appears to be hazing over: there are more components to it than his regular cognitive percepts would indicate. It is a matter of (his) *mood*. His “inner sense” is thrown into a state of confusion, panic, and deep pity (*kṛpā*), with verbal evocation that his limbs have become weak, mouth dry, body trembling, hair standing on end, and skin erupting in a burning sensation. He confesses that the once cherished desire (*kāṅkṣe*) for conquest and aligned convictions appears shaky; he wonders aloud whether there is any joy at the end of this bloody

journey—or even in living (I.32). Expressing a deeper fear for the death of his kinsfolk at his own hands, he preaches to Krishna: “Therefore there is no justification in killing our own kinsfolk” (I.37).

Arjuna continues his disquisition, underscoring utilitarian appeals to the evils of warfare and a plea toward altruistic compassion, speaking of “the rescinding of family laws, ancestral rites, and timeless traditions, with the ultimate consequence of the collapse of society and descend into hellish chaos” (I.40–44). He can no longer stand by his earlier resolve to fight now that the “moral emotion” that he is struggling to articulate appears to be inconsistent with the “moral duty” he was brought up to believe in.

Fallen into self-pity, the despondent warrior pleads to Krishna to make sense of his woeful plight. But in this petition is Arjuna appealing to the pristine virtue of reason over emotions, or is he, instead, asking Krishna to tell him if his emotions are serving him well? Can emotions prefigure morally appropriate, “objective,” and reasonable responses, even if they appear to elude his cognitive or rational discernment? He has not yet discerned clearly whether he feels shamed, guilty, regretful, remorseful, or a combination of these; or none of these but something else. Krishna, for his part, proceeds cautiously in helping Arjuna unearth his deep perturbation.

Krishna plays the dual role of a guru and an analyst rolled into one. The guru can, with measured smirk and laughter (*hāsya*), rebuke his honored friend for losing heart at a critical moment. As an analyst, though, he implores him to search out reasons for the fragility of his judgment. Arjuna’s objections to engaging in war appear to be based on well thought out and firm ethical grounds, but when he sets out to articulate the “inspired thought” intelligently his arguments emerge as being scarcely coherent, and the appeal to his own conscience is minimally illuminating. But he is concerned that he is not able to see justice in this situation. In other words, he gives vent to a moral sentiment that he has arrived at as though intuitively (as Hume might also put it); his arguments, it will be noticed, are tangled up in his intense emotional reaction, the source of which he is not able to discern clearly. We can wonder why Arjuna remains perturbed by his emotional condition despite Krishna’s irenic response. Why would Krishna want to seemingly dismiss his friend’s condition? Is it a socially improper or morally unworthy state to be in? Perhaps it is psychologically or psychosomatically painful and therefore bereft of utility? Or is such an emotional state simply irrational because it fogs well-intended judgment and vitiates the Rawlsian equation, as Matilal has argued elsewhere?<sup>10</sup> But what if emotions have other values and efficiencies (*bhāvaka*), for example, if a “moral emotion” goes against the grain of cherished religious mores that might, in themselves, be irrational? Have not his emotions made Arjuna a little more reflective, muddled though he is now, than he might otherwise have been about his proper duties? Is he not, as a result, at least “talking it out” with his friend? Indeed, might there not be an obligation

to have such emotions, just as there is duty on Arjuna's part to engage in an action? Might it not also be a person's inalienable right—on a par with aesthetic or dramatized emotions, whose value no one really questions in the comforts of a theater seat or the civic sponsored museum and art galleries?

In the next section we shall take these reflections a stage further by looking at a theoretical ramifications in the broader context.

### 3. VARIETIES OF EMOTIONS

According to the *Mahābhārata* (1985, books 4–5, 289), there are twelve negative emotions or vices that stand in the way of self-control and so are to be avoided: “anger, desire, greed, delusion, possessiveness, non-compassion, discontent, pride, grief, lust, jealousy, and abhorrence.” It is interesting to note that anger and desire head the list. Indian yoga thought has a term stronger than emotion (as *bhāva*) for these modes; they are called *kleśas*. The Abhidharma Buddhist school has a similar theory about *kleśas* functioning as emotional predispositions or tendencies (*anusaya*), which lie “dormant” or latent at the unconscious level. The *kleśas* are regarded as forms of psychic sedimentations, which give rise to mental disturbances or excitations.

Desire as lack, then, is a prime suspect heading or enveloping all *kleśas*: it colors all our emotions from beneath, and its frustration translates more readily into obstructive anger than into pleasure (though perhaps both), only that morbid pleasure (desire as plenum) might be seen as more obstructive than “righteously felt indignation” would in some contexts. Thus, while Arjuna could be said to have harbored a *desire* for the kingdom in dispute (at least until the moment of his emotional collapse), he appears not to have expressed any anger. This is rather puzzling, given that he seems to have been overwhelmed by just about every other major negative affect and to have taken some pleasure in the positive affect of generalized compassion (*krpa*). I want to dwell on this issue a while.

It appears that the *Mahābhārata*'s “Great Song” is open to alternative perspectives to the stark ascetic or stoical tendencies of the ancients. As we saw earlier, Krishna, although apparently denigrating, did not really deny or show utter disrespect for Arjuna's revaluations of his calling to war, his duty, and so on. In point of fact, Krishna listened intently and recognized a touching concern. He rebuked Arjuna, or rather questioned him, only regarding the *grounds* on which he was making his revaluations, just as Arjuna could expect to be questioned were he making them in a perfectly regular (or “normal”) rational state. Is he sure that he is not simply projecting his own self-pity as generalized sympathy or altruistic compassion onto others? And whence did he convert to being a singular (crypto-Buddhist) utilitarian and not an Epic(urian) Kantian?

Nevertheless, Arjuna proved right in the long run in his prophylactic emotional response. Everything the beleaguered warrior suspected in his apparently confused, fearful, semimorbid and besotted state—the destruction of the kingdom, the carnage of the elders, the collapse of family and tradition, the ruin of his *Gāṇḍīva* missile, the demise of his invincible golden chariot, and so on—did actually come to pass according to his prediction during the tragic course and untriumphant conclusion of the war. It was a veritable anticlimax. Thus the evil Arjuna had portended in his seemingly tepid turpitude, and the reasons he proffered for his fears in that most despondent and emotionally charged condition (I.31, 36–40), played themselves out in the real world. Although, to be sure, the moral *judgment* he ventured in his perturbed state was not what the society, consistent with the norms of the time, was prepared to countenance, as we can observe in hindsight.<sup>11</sup> So why did Krishna believe that Arjuna’s appraisal of the impending crisis, and not his own wisdom on the matter, was misguided, erroneous, and amoral? To answer the question one must appreciate two presuppositions informing Krishna’s assessment of Arjuna’s condition. The first is derived from the following Upanishadic dictum.

A person is what he desires  
 desires affect his resolve (*kratu*)  
 this determines action . . .  
 good action makes one good  
 bad action evil. (Radhakrishnan 1975, IV, 3–4; II, 2, 12)

It is clear that consequentialist appeal, whether by the agent or the theoretician two millennia on, will not suffice to justify a particular emotional response. This is the first register. The second point here pertains to the relation between desire and karma or action. Karma is necessarily conditioned by an antecedent *kāma* (desire) and ineluctably followed by a corresponding *phala* (fruit), either in this or in a subsequent life-world (*punarjanma*) (De Smet 1977, 59). Hence all actions are binding and also delimited by their outcome: good ones to (and by) a pleasant fruit; bad ones to (and by) a painful one. Without trading in casuistry, if I helped an elderly neighbor cross the street in Melbourne, I might be helped out in Soho to find my way around the Park Place conduit to Chambers Street, although the airless tunnel might make me feel sick.

This view about karma can be rephrased in the “Great Song’ ”s terms as follows: Let a person but think (*dhyāya*) of the objects of the senses, attachment (*saṅga*) is born; from attachment springs desire (*kāma*); from desire is indignation (*krodha*, i.e., when desire is frustrated); from indignation (*krodha*) comes bewilderment (*sammohā*); from bewilderment wandering of the mind (*smṛti-vibhrama*); from wandering of the mind destruction of the intellect (*buddhi-nāśa*); once intellect is destroyed, impairment of memory and judgment . . . and the person is lost (II, 62–63).<sup>12</sup> The view expressed overlaps with the Buddhist treatment of action, with the differ-

ence that the “Great Song” situates *kāma* (desire) and *krodha* (indignation) as the paradigmatic passions and holds them in mutual tension (*kāma eṣa krodha eṣa*; III.37). In this respect also, therefore, Indian thought is closer, on the one hand, to Aristotle and the European schoolmen’s doctrine of desire, and, on the other hand, to the psychoanalytic focus on the link between libido and aggressiveness.<sup>13</sup> *Kāma* with *krodha* is characterized as the “timeless foe (*nityavairāṇā*) of the wise man” (III.39) lurking in subterranean regions of the self, to wit: “Senses, mind and intellect, they say, are the places where it lurks; through these it smothers wisdom, fooling the embodied self” (III.40–41). This is its link with unsatisfactoriness—the pervasive instability that manifests variously as doubt, confusion, deceit, distracting thoughts, and, last but not least, erroneous beliefs about reality. It is interesting to note that a vernacular term for desire is *bhāvana*, as a voluntary wish or covert expectation (*saikalpa*), achieved through adding a suffix to the root word for emotion (*bhāva*). This etymological link is a reflection of their deep epistemic and phenomenological relation.

Arjuna is said to have been lured by *kāma*, desire that accentuates attachment to things. While renouncing, in the spirit of dispassion, all remaining desires for joy and power—desires to which he is entitled as a warrior—he retains other desires, or passionate attachments, or rather *disinterested passions*, which occlude his vision. By rights, the frustration of these desires should evoke a rage of anger in him, not just resentment.

Anger, it would seem in theory at least, should come naturally to Arjuna, for in his present constitution the attributes (*gunas*) of *rajas* (movement) and *tamas* (heaviness) mutually reign in excess of *sattva* (lightness), which are key attributes of just about all mental modes with corresponding emotions that divide somewhat neatly into three clusters. This surely in part accounts for the weakness of his will and the lack of resolve in his determination (*vyavasāya*). However, nowhere in his introspective ruminations does Arjuna once give any hint of being seized by anger.

There is one apparent exception: after having sketched the dire consequences of the impending war, and as if throwing his arms up in the air “hot and cold,” he exclaims, “Ah! alas! mighty evil (*pāpam*) to perpetrate has been our resolve!” (I.45). But this posturing is followed immediately by delivery of rather fragile moral good by the desire “in the mind” (*buddhi*, as “conscience”) that stands perturbed. But this “good” is also in a sense empty of substantive contents, for its value is not in the parts or particular acts it sets off but rather in the “whole” sense of life’s accomplishment from a future standpoint. Arjuna is thus happier to be judged by this fragility rather than by the number of violent acts and beheaded “enemy trophies” he can display at the endgame. And this negative good that has emerged from the deep well of the will (*avyaktabuddhi*) as his intelligent resolve consists in disarming himself, forfeiting the valorized virtue of courage, and submitting to death by the other (I.46, 47). This is not the way

of an angry, livid person; rather, again, stoicism or *śramāṇic* resemblance aside, this appears to be the posture of a timid, humiliated, and withdrawn man.

The explanation I favor is that anger, with the concomitant expression of aggression, is theatrically or visibly a greater embarrassment and threat than the disruption or frustration of desire (which can be kept as part of one's personal or "private language," suppressed and sublimated in other ways). *But anger is a consequent rather than the antecedent of the perturbation of desire.* For this reason some traditions would rather suppress anger "in the bud" or at every opportunity. Also, anger, much like pain, is an occurrent state that need not last for too long, while desire is seen as being deeply entrenched, as advanced cancerous cells, or at least pervasive across a good part of the individual's psyche and "subtle" body corporate. True, there are numerous incidents and situations, before, during, and after the war, when rage (*kupita*) and anger expressed by members of the feuding parties appear with some prominence. But these incidents merely exemplify the theory of self-control that the *Mahābhārata* adopts from the yoga tradition. At no point does Arjuna's "blood boil over"; he seemed too "coolheaded" and in a more pensive mood to suffer any indignation or afflict the same upon others. Moreover, he was not given a voice of rightful indignation to express his sense of being morally affronted. If Arjuna had allowed himself to become angry, he might well have judged differently, and he may even have felt sufficiently empowered to resolve his dilemma in a different direction that could have resulted in a more radical action. One cannot escape the conclusion: "righteous indignation" fails to receive the regard in this tradition accorded it by, say, Aristotelian ethics, or in Buddhism, as de Silva (forthcoming) has forcefully shown.

Stocker (1996, 253) considers detachment in the exemplary lives of, say, Zen masters, and he wonders whether such affectivity can be devoid totally of passion and anger, especially when one's sensibility (such as religious identity) is slighted. If so, Aristotle might be mistaken in holding that anger is essential to a good human life. But I think the kind of dispassion that Krishna is preaching here is really in relation to personal affront rather than slight to one's group, clan, or communal identity—for which the reaction is more structured and where the respondent at large is the tradition. Hence Brahmanic and upper-caste vituperate anger was never seriously in doubt, except in the minds of marginal players, the oppressed, and outsiders (where women more often than not belonged).

Nevertheless, the discourse of desire receives a great deal more attention than anger does in the "Great Song." Krishna has not denied Arjuna's inherent capacity to make fair judgment of the situation; he has simply cast doubt on Arjuna's ability to reason and correctly evaluate his situation while in the grips of desire. This is indeed a paradox: Arjuna believed that he had come to the point of relinquishing all desires—the desire that at-

taches itself to glory, to fame, to booty from war, to a share in the disputed kingdom, to the fulfillment of caste duty and so on. His desire has been disrupted by his emotional response in the situation. Krishna, on the other hand, tells Arjuna that he is afflicted with desire, and attachment to desire causes other kinds of perturbations, not least frustration, anger, sense of unsatisfactoriness (*duḥkha*), and undignified death at the journey's cruel end. Only if a person would rid himself of all desires and remain content within himself will he be called *ceteris paribus*, a person of "steadfast constitution" (*sthitaprajña*; II.55). Time and again Krishna's sermon underscores the negative aspects of emotions such as anger, fear, passion, and egotism rooted in desire, unless moderated by the cool judgment and dispassion of *buddhi* and resolute will steeped in wisdom and ethically fine-tuned action (II.56, V.26, V.28).

In practical terms what is expected of a person is not the willingness to forfeit or relinquish these emotions in their entirety—and desire is no exception here—but rather to exercise equanimity or a mean sense of balance, an equilibrium, dispassion or disinterested passion (*niṣkāma*) toward and between the extremes, so that karma does its work. Krishna preaches even-mindedness (*samatā*) toward pleasure and pain—in general, indifference toward pairs of opposites (I.57, II.38, II.45)—as well as not being too excited when experiencing joy, nor feeling ruffled when facing sorrow (V.20).

"Balanced reason," whose mark is "steadfast constitution," is then the virtue most exalted—definitely above self-pity and self-concern, and even over altruism and self-enlargement (read self-realization)—for the objective is to let truth (not self or one's god or the aura of a beloved) shine through the emotions, as much as in cognitions and hermeneutical acts. The seat of reason is *buddhi* (or *bodhi* in Buddhist rendition), the intelligent will, and it is toward a stabilization and refinement of the latter that the qualities and virtues being inculcated here are intended. While one can use yogic methods to withdraw from the objects of sense in order to prevent further sensations from arising, it is not so very easy to curb inner and unconscious perturbations born of *kleśas* ("psychic black mirror") and other sedimented (inverted memory) traces. Tranquility, achieved through prolonged practice of concentration of the mind, or meditation, may be necessary for the "cessation of all sorrows" (II.65). Disposition toward nonviolence, veracity, absence of anger, compassion for all beings, and freedom from the thirst of either extreme are among the highest virtues inculcated through the pragmatics of yoga; but virtues can never take the place of experienced or felt affectivities themselves. The ideal is a mere surrogate for the real (thing).

If one could cultivate the alternative emotion of detachment (*asakti*; III.25), freeing oneself from the temptations of *kāma* and also anger, then one would achieve a state of reasonable intelligence (*vyavasāyātīka buddhiḥ*, II.41) and in this resolute state determine the best course of action.

Such actions would be *niṣkāma*—that is, empty of desire—and *sthitaprajñā*, “of steadfast constitution.” This is a normative heuristic, not a categorical imperative, for emotions cannot be prescribed, they can only be cultivated in a cultural setting. Actions carried out in this state do not bind one; that is to say, karma no longer accrues, for one no longer expects rewards from one’s action (II. 39; VI.14). Desire and self-interest bracketed, one is left in a state of freedom to perform actions from a sense of duty (rather than by compulsion).

The moral import of emotions is not, then, on this view, undermined in the interest of emphasizing the obligation in respect of duty. Indeed, duty is understood, appreciated, and reappraised through insightful emotional response. This response is modulated on the one hand by self-love and on the other by regard toward the wider horizon of cultural sensitivities. At the very end of their enchanting though exhaustive colloquy, Krishna does not issue an unmitigated command to Arjuna; rather, he leaves it to his own better judgment, with these telling words: “Having reflected on this [my words] in all its ramifications, *do as you desire*” (XVIII.63). In short, you are never implored to (1) eliminate all desires, (2) perform your duty regardless of desire, or (3) simply do my bidding because I am your god.<sup>14</sup> I stand by my earlier hunch that the “reasonableness” so understood means that religion can essentially be reduced to the psychological and cultural, the images of the divine presence and theophanic exuberance to little more “than a rhetorical slide show,” “a heuristic move,” which “psychoanalysis might recognize . . . as an act of positive transference” with hopes of transformation of desire and the possibility (in deference to Matilal and the *Mahābhārata*’s recalling of the trope) of “moral love.”<sup>15</sup>

## CONCLUSION

While desire as lack is everywhere sought to be bracketed and sometimes even extinguished (perhaps a legacy of the ascetic ascendancy), in the Indian approach to emotion, it nevertheless is acknowledged to be among the more significant component of the mental and spiritual life-world. And unlike the earlier śrāmaṇic or ascetic traditions, the yoga praxis does not aspire to the total eradication of desire; a complete dispassion, though desirable, becomes merely an ideal or a limiting concept after the Epics.<sup>16</sup> For this reason, the “Great Song” offers a different avenue to dealing with desire, which admittedly has not been entirely uprooted and may never be. The sermon preempts a Freudian insight into the *jouissance* of sublimation.

The energy (*śakti*) of desire can be redirected to another object of focal concentration and thereby transformed into another passion, such as love, or loving compassion (in modern Buddhist parlance). When the intentional structure of this transfigured passion is constitutive of a transcendental object as other, this particular love-passion is called *bhakti*, *prema*, or de-

votion. Sexual or erotic libido likewise can be refined and turned into a channel of love, devotion with *jouissance*, in intimate comport with the imagery of a chosen, gendered, or nongendered deity, as cultivated in Vajrayāna Buddhist and Hindu Tantric (particularly Bengali Vaiṣṇava-Sūfī or Baul) practices. The transformative praxis requires vigilance against obstructions and ego-satisfaction or fetish and boredom overrunning the passions.

In this context desire takes on a different dimension and color. The symbiotic relationship between the subject and other has the capacity to elicit the sublime enjoyment or *bhaktirasa*, and a tremendous sense of *community* or *prasāda* (*Pali passāda*) in the company of the *sagha* or congregated community. In this identification, love, the most passionate of passions, is boomeranged back to the individual: “I return my love in the same way that people bequeath theirs to me” (*bhajāmi*; IV.11). The objective of this “grace” (Irigaray’s “gift”) is to increase the individual’s self-esteem and, hence, ability to be responsible for his or her own ethical guidance. So that this transferred “love” does not become egoistic, *pace* the Buddha’s sermon to the birds; the next moral move has been to disperse much of it toward others; “moral love” has come to be called *karuṇā* or “compassion,” “loving kindness,”<sup>17</sup> and so on. But does *compassion* diminish as the circle expands where equal sympathy for umpteen hordes of sentient being vie for recycled birth? Regardless of these contingencies, compassion is indispensable as a virtuosity for the would-be *bodhisattva*, *arhant*, and *avatāra*, or sage, nearly enlightened beings. For this is a special emotion and its cultivation is a good (*śreya*) in itself. The bodhisattva would beam with the same degree of compassion in all possible worlds whether frothy to the Spinozian brim or empty as the Nāgārjunian sky.

It is important to understand the psychology here. While compassion for others is in many ways different from a one-to-one relationship between, say, two human individuals, the motivation is nevertheless somewhat similar. The net effect is that of transforming one passion, such as selfishness, into another, such as love, which makes one “intent on” the other (II.61), “trusting-and-loving” another (VII.16). Love (*bhāva*) engenders firmness (*dhṛti*) in the mind, thus returning one, integrated in mind and body, to the field of *dharma* or the ethical life-world. Alternative to theistic images, the *bhāva* that is summoned up here bears comparison to the kind of loving attitude or deeply encompassing, self-effacing, respect some Levinisian ecologists have called for in another context.

This, then, is South Asia’s way of rekindling passion by arresting emotionally depleting passions born of desire and reorienting the perceptual or symbolic field (*kṣetrajñā*) toward a more rationally balanced and passionate window on the world. The end result is to surrender neither to the dictates of a despotic orthodoxy, or orthopraxy, hitherto bridled to the fruits of sacrificial action (*yajñā*), nor to the world- and self-denying tendencies

of stoical asceticism, with its diminution of the role of emotions and passions and the feminine in the life-world. Rather, the tradition settles for something of a middle position that attempts to reconcile the two opposing strands.

But to accede to this fact is also to recognize that reflections on emotions cannot proceed without paying specific attention to the general ethical framework or moral discourse of the culture in which these have their *bāshō* or *topoi*. Likewise, reflections on the ethical norms and morality of a culture cannot proceed without reevaluation and critique that is instigated by emotions, which propels and guides rational deliberation on the pertinent and disturbing issues that might escape mere abstractions. Which emotions are to be cultivated, how many (all forty-eight as in the *rasa* dramatics?), and how these are used for adaptive and energizing purposes depends upon the mores of a culture at any given point in historical time and geocultural space.

## NOTES

1. See Masson 1981 for discussion of the first view, and Nussbaum 1995 for discussion of the second view. Nussbaum charges that non-Western traditions, such as Islamic, Chinese, and Hindu, adjudge women (among other subjects) to be given more to the emotions than to reason, and Western gender-based and ethnic denigration of emotion intersect with colonial perception, according to which people of developing countries are excessively emotional and normatively irrational, hence their economic plight.

2. *Śrīmadbhagavadgītā*, II.4. *Śrīmadbhagavadgītā*, in the shortened form *The Bhagavad-Gītā* (sometimes just *The Gītā*), is a book within the *Mahābhārata*, or the Greater Story of the Descendants of Bhārata (inhabitants of an imagined greater India), the largest continuous poem in world literary achievement. *The Bhagavad-Gītā* is usually translated in various languages and vernaculars as the “gracious god’s eulogy” (referring to Krishna’s dialogue with Arjuna), or, after the eighteenth-century Orientalism, simply as the “Lord’s Song.” The “song-line” here is sharper and rather more philosophical than much else in the *Mahābhārata*, although a certain moral turpitude is expressed throughout the entire corpus. I am taking an unpoetic license to call this mellifluous book, a moving deliberation that was purportedly “sung” by the legendary scribe Vyāsa of the *Mahābhārata*, as the “Great Song.” I provide my own translation of the verses cited, and no available translation is entirely reliable because the cultural context and moral frames are lost (see n. 16). A good bilingual rendition I have consulted is *The Bhagavadgītā in the Mahābhārata* (1985).

3. For more on these attributes, see part 3 below.

4. De Silva 1995, 110–11; Roberts; Stocker 1996.

5. See McDaniel 1995, 41–42. Although McDaniel is using this characterization only in unpacking *bhāva*’ in Hindu thought, I think it can be generalized to emotion across the board (without reducing all emotions to *bhāva*, of course, as some writers tend to).

6. Sheldon Pollock in Chicago has been working on this in the works of Bhoja, an eleventh-century cultural aesthetician. Suthar Visuvalingam writes on Abhāvanagupta's reduction of all aesthetic emotions to *śāntarasa*; see his thesis at [www.svabhinava.org](http://www.svabhinava.org).

7. See McDaniel 1995, 46–49, although McDaniel's goal seems to be to highlight development of *bhaktirasa*, devotional sentiment in Vaiṣṇava tradition.

8. In Indian theory of perception (*pratyakṣa*), sensations are described as being *nirvikalpaka* (experience): a buzzing, bubbling, inchoate, undirected, hazy cloud that floats into a barely gathered awareness field (*kṣetrajnāna*). However, even in their prelinguistic givenness it is believed that “strong” experiences that turn out to be cognitions or sentiments, or even memory recognition, are already fore-structured and therefore expressive at a later moment, in a “this-as-that” form, resembling a propositional attitude. Thus, “this” experience (*anubhāva*), phenomenologically present, is of that property marking (attached to or instantiated in such-and-such an object, whose exact identity may still be at large), as in the “blob” I see as “round red-patchiness” out there (and, in the next moment, I mutter, “Ah, it's a tomato”).

9. See Nussbaum's discussion of Posner 1981, and see Nussbaum's essay in this volume.

10. The Rawlsian type of equation that Matilal had sought between “rational frustration” and “appropriate moral repertoire” just does not apply in cross-cultural contexts (or perhaps anywhere except in Cambridge, Massachusetts). See Matilal 2003.

11. Although grief and remorse there were aplenty, but not just these. See the *Mahābhārata*, book V onward. See also essays in Matilal 1989.

12. See also *Bhagavadgītā in the Mahābhārata*, VI.41–45, IX.20–21, XVI.19–21.

13. See Matilal 1978 and Sharma 1992 and 1993.

14. I am grateful to Alan K. L. Chan's laudatory review of Marks and Ames 1995 and succinct extraction from the dense passage in an earlier version of my essay in that collection (1998, 179). This review (more than an anonymous referee's ignorant reading for the *Australasian Association of Philosophy's* proposed symposium on East-West philosophy, 1995) helped me revise and focus this treatment of emotions in Indian thought.

15. Matilal derives his argument and understanding of “moral love” from the *Śrīmadbhagavadgītā* as a corrective to the more pessimistic reading (particularly in the hands of nineteenth-century European *philosophes*, notably Schopenhauer and Hegel, who penned a commentary of his own on Humboldt's German translation of the text) of the depressive tracts between categorical exhortations to “just war” via caste duty or Bharata's vision of natural law and universalized love of others as a matter of self-duty (*svadharma*). See Matilal 2003 and the discussion in Bili-moria 2002, 166. Spivak 1999 (45, 46, 50, 310) discusses Matilal on “moral love.” (The title is an *amartyan* [immortalizing] deference to the “Third Critique,” via Sen.)

16. See Goldman 1997 and Phillips 1999.

17. The ever smiling Dalai Lama, whom Renuka and I went to hear speak

on the Buddhist philosopher Atiśā, appears to have perfected this wondrous art of beaming back the “loving compassion” he is able to generate in his audience by his majestic presence and constant reference to the pervasiveness of suffering among all sentient beings—the Chinese regime included—which can only be countered, according to the wisdom, by practical application of the values of nonviolence, compassion, and caring responsibility in all possible worlds (*bodhi-lokas*).

# VII

## On Theories of Emotion

*This page intentionally left blank*

## Is Emotion a Natural Kind?

PAUL E. GRIFFITHS

In Griffiths 1997 I argued that it is unlikely that all the psychological states and processes that fall under the vernacular category of emotion are sufficiently similar to one another to allow a unified scientific psychology of the emotions. The psychological, neuroscientific, and biological theories that best explain any particular subset of human emotions will not adequately explain all human emotions. In a slogan, emotions are not a natural kind (14–17; 241–47).<sup>1</sup> I argued that the same is probably true of many specific emotion categories, such as anger and love (17). On some occasions when a person is properly said to be angry, certain psychological, neuroscientific, and biological theories will adequately explain what is happening to that person. On other occasions of anger, however, different theories will be needed. I described my position as eliminativism about emotion, because it implies that the term *emotion* and some specific emotion terms like *anger* are examples of what philosophers of language have called “partial reference” (242). The term *jade* is the classic example of partial reference. The term *jade* is used as if it referred to a particular kind of mineral, in the same manner as *malachite* or *diamond*. In reality, however, the term covers two different stones, jadeite or nephrite. The term *jade* partially refers to each of these two minerals. Hence, for the purposes of geology or chemistry, jade cannot be treated as a single kind of thing. The properties of the two substances have to be investigated separately, their geological origins explained separately, and their abundance in unknown geological deposits predicted separately. Likewise, I argued, the sciences of the mind will have to develop separate theories of the various different kinds of emotion and also of the various different kinds of some particular emotions. In the same sense that there is really no such thing as jade, only jadeite and nephrite,

there is no such thing as emotion, only “affect programs,” “socially sustained pretences,” and other more specific categories of psychological state and process.

While the critical response to my book has been generally positive,<sup>2</sup> many philosophers remain unconvinced. I suspect that this is to a significant extent because the question of whether emotion is a natural kind has been conflated with the question of whether the concept of emotion can be given a univocal analysis. These two questions have very little to do with one another. The concept of a superlunary object can be analyzed—something is superlunary *iff* it is outside the orbit of the moon—but the consensus since Galileo has been that those things do not form a natural kind. Likewise, my claim was not that the vernacular concept of emotion is unanalyzable—I myself suggested a prototype analysis (242–45). My claim was that the things that fall under the concept do not constitute a distinct kind for the purposes of scientific enquiry. The concept of “vitamin” is a useful comparison. Vitamins are not, as was once thought, “vital amines” but a diverse group of chemicals with diverse roles in physiology sharing the feature that humans cannot synthesize them, or can synthesize them, as with Vitamin D, only under advantageous environmental conditions. Their absence leads to “deficiency diseases” with diverse etiologies and diverse prognoses. So the concept of a vitamin can be analyzed, and individual vitamins and even some groups of vitamins are natural kinds, but “vitamin” itself is a superficial descriptive category. It is not a sensible scientific project to investigate the nature of vitamins *in general*. The question “What is a vitamin?” is best answered by describing the main kinds of vitamin and how different they are from one another.

Another important line of reply has been that it is a mistake to ask whether emotion and emotions are natural kinds at all, since they are primarily normative kinds (Doris 2001). I have addressed this issue at length elsewhere (Griffiths forthcoming a). I wholeheartedly agree that the normative aspects of the semantics of emotion concepts are vital to an adequate account of those concepts, a view I endorsed in my book (7; 196–201). However, I do not accept that the normative uses of emotion concepts can be cleanly separated from the descriptive uses. Hence, the fact that emotion concepts are normative concepts does not mean that they are insulated from revision in the light of empirical discoveries about emotion.

In the remainder of this section I restate what I mean by “natural kind” and my argument for supposing that emotion is not a natural kind in this sense. In the following sections I discuss the two most promising proposals to reunify the emotion category: the revival of the Jamesian theory of emotion associated with the writings of Antonio Damasio and a philosophical approach to the content of emotional representations that draws on “multilevel appraisal theory” in psychology.

## What Do I Mean by “Natural Kind”?

I use the traditional term *natural kind* to denote categories that admit reliable extrapolation from samples of the category to the whole category. In other words, natural kinds are categories about which we can make scientific discoveries. In my book I built on the work of several other philosophers and scientists to construct an account of natural kinds in psychology and biology, an account further elaborated in Griffiths 1999 and 2001a and briefly sketched here.

The fundamental scientific practices of induction and explanation presume that some of the observable correlations between properties are “projectable” (Goodman 1954). That is, correlations observed in a set of samples can be reliably “projected” to other instances of the category. Scientific classifications of particulars into categories embody our current understanding of where such projectable clusters of properties are to be found. The species category, for instance, classifies particular organisms into sets that represent reliable clusters of morphological, physiological, and behavioral properties. Hence, these properties of the species as a whole can be discovered by studying a few members of the species.

The traditional requirement that natural kinds be the subjects of universal, exceptionless “laws of nature” is too strong and would leave few natural kinds in the biological and social sciences where generalizations are often exception-ridden or only locally valid. Fortunately, it is easy to generalize the idea of a law of nature to the broader idea that statements are to varying degrees “lawlike” (have counterfactual force). This broader conception of a lawlike generalization allows a broader definition of a natural kind. A category is (minimally) natural if it is possible to make better than chance predictions about the properties of its instances. This, of course, is a very weak condition. Very many ways of classifying the world are minimally natural. The aim is to find categories that are a great deal more than minimally natural. Ideally, a natural kind should allow *very reliable* predictions in a *large domain* of properties. The classic examples of natural kinds, such as chemical elements and biological species, have these desirable features.

It is important to note that categories are natural only relative to specific domain(s) of properties to which they are connected by background theories. The category of domestic pets is not a natural category for investigating morphology, physiology, or behavior, but might be a natural category in some social psychological theory or, of course, in a theory about domestication. Emotion, I argue, is not a natural kind relative to the domains of properties that are the focus of investigation in psychology and the neurosciences. It is not the case that the psychological states and pro-

cesses encompassed by the vernacular category of emotion form a category that allows extrapolation of psychological and neuroscientific findings about a sample of emotions to other emotions in a large enough domain of properties and with enough reliability to make emotion comparable to categories in other mature areas of the life sciences, such as biological systematics or the more robust parts of nosology.

### Why Emotion Is Not a Natural Kind in This Sense

On some occasions when a person “has an emotion” they are producing an affect program response—a “basic emotion” in Paul Ekman’s sense. The response is short lived, highly automated, triggered in the early stages of processing perceptual information, and realized in anatomically ancient brain structures that we share with many other vertebrates. It is found in all human cultures and closely related to responses in other primates. Suppose, for example, that you are waiting in line outside a nightclub. After twenty minutes, someone unexpectedly pokes you sharply in the small of the back. You spin around, making a threat expression, probably the “square-mouthed” variety, your body adjusts physiologically for violent action, and your attention is entirely on your assailant. If the situation is rapidly defused (you are male and an attractive young woman has tripped against you and is smiling apologetically) then this will be a pure case of affect program anger. On other occasions, however, a person “having an emotion” is responding in a more cognitively complex way to more highly analyzed information. The episode may or may not involve the occurrence of one or more affect program responses. Suppose, for example, that you are locked into a dysfunctional pattern of interaction with your spouse involving continual fault finding and put-downs, this pattern emerging without any intention from the particular patterns of relationship management you both bring to the marriage. The pattern has resisted your occasional attempts to consciously improve your behavior and, as you reflect one day on what appears to be the inevitable degeneration of the relationship, you experience a deep sense of guilt and self-loathing. I referred to such states as “higher cognitive emotions.” Finally, on some occasions, “having an emotion” may centrally involve an internalized cultural model of appropriate behavior. In my book, I suggested that people who respond to losing their job by “going postal”—going on a killing spree followed by attempted suicide—could be simultaneously “out of control” and following a “script” derived from real and fictional incidents that are culturally salient for them. I presented a tentative analysis of such cases as “socially sustained pretences.”

My argument that emotion is not a natural kind rested mainly on the first two cases, the affect programs, or “basic emotions,” and the more

cognitively complex emotions, which I termed “higher cognitive emotions.” I regret using the latter phrase, as it gives the impression that I have a substantive theory of those emotional responses. In fact, I discussed the views of leading theorists like Antonio Damasio (102–6) and Robert Frank (117–22), whose ideas clearly do something to illuminate these more complex emotions, but my only firm conclusion was that these states and processes are unlikely to be reduced to the basic emotions or understood as blends or elaborations of them. I accepted, however, that these other emotions may involve basic emotions as parts, depend on basic emotions for their development in the child, and interact with basic emotions in typical ways in real-life situations. To better indicate how little we know about nonbasic emotions, I now prefer to call them “complex emotions” (Griffiths, 2003, forthcoming b). Finally, I rejected the view that the basic emotions are not emotions or that they are merely proto-emotions (26, 164). This sort of procrustean treatment is inevitable as long as we insist that emotions are a single kind of thing. Instead, I suggested we should accept that there are two or more different kinds of psychological process involved in emotion, and if research into complex emotions suggests that, like basic emotions, they can be classified into emotion types, then there are two or more different kinds of emotions.<sup>3</sup>

The other major element of my case for “eliminating emotion” was my specific account of natural kinds in biology. I have defended that account in numerous places (Griffiths 1994, 1996a, 1996b, 1997, 1999, 2001a, 2001b) and can only briefly sketch it here. Evolution leads to the existence of two fundamental sets of biological categories—homologies and analogies. A homologue is “the same organ in different animals under every variety of form and function.” (Owen 1843, 374), a definition interpreted since Darwin to mean that these organs are descended from a common ancestral form. Analogies are cases where two unrelated structures resemble one another because natural selection has adapted them for the same ecological role. My hips are homologous to those of a horse, but they are analogous to the articulation of an arthropod limb-pair with the rest of the arthropod segment. These two concepts have been routinely applied to behavior since the rise of ethology in the 1920s.

I argued that some basic emotions are analogous to complex emotions that fall under the same vernacular category. The different kinds of fear, for example, are all responses to danger. Any psychobiological theory of emotions *in general*, I suggested, 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 neuroscience. They would support induction and explanation on the wrong domain of properties (230–41). Suppose that two animals have homologous psycho-

logical traits, such as the basic emotion of fear in humans and fear in chimpanzees. We can predict that, even if the function of fear has been subtly altered by the different meaning of “danger” for humans and for chimps, the computational methods used to process danger-related information will be very similar and the neural structures that implement them will be very similar indeed. After all, Joseph LeDoux’s widely accepted account of fear processing in the human brain is largely, and legitimately, based on the study of far more distantly homologous processes in the rat (LeDoux 1996). Now suppose that two animals have psychological traits that are analogous—fear in the rat and fear in the octopus, for example. It is a truism in comparative biology that similarities due to analogy (shared adaptive function) are “shallow.” The same problem can be solved in different ways, and so the deeper you dig, the more likely it is that mechanisms will diverge. Bat wings and bird wings, for example, have similar aerodynamic properties but their internal structure is radically different, reflecting their different ancestries. In contrast, similarities due to homology (shared ancestry) are notoriously “deep”: even when function has been transformed, the deeper you dig, the more similarity there is in the underlying mechanisms. Threat displays in chimps look very different from anger in humans, but when their superficial appearance is analyzed to reveal the specific muscles whose movement produces the expression and the order in which those muscles move, it becomes clear that they are homologues of one another. The same is almost certainly true of the neural mechanisms that control those movements.

Now, psychology is in the business of uncovering the mechanisms that produce behavior. This is even more evident in the case of neuropsychology. Hence these disciplines seek categories that are heuristically valuable for the study of underlying mechanisms. Inferences to shared mechanism based on homology are not 100 percent reliable, but they are reliable enough to build good science with—the criterion I outlined above—and they are more reliable than inferences to shared mechanism based on analogy. I concluded that replacing the categories of basic emotion—which are explicitly categories of psychobiological homology—with more general categories that included any mechanism that performs the same broad adaptive function would reduce the inductive and explanatory power of cognitive psychology and the neurosciences. It would be a move from a more natural category to a less natural category in the sense defined above. The aims of these sciences are better served by recognizing that there is more than one kind of emotion and by investigating each on its own terms.

Having summarized my arguments for the view that emotions do not form a natural kind, I now go on to consider the most promising proposals for the opposite view.

Are Complex Emotions Blends  
of Basic Emotions?

Innumerable theorists have suggested that complex emotions are blends of basic emotions. In my book I argued that some of the features that characterize complex emotions cannot be accounted for merely by supposing that several basic emotions occur simultaneously or in sequence. First, it is characteristic of some complex emotions that they occur in response to complex properties of the stimulus situation and so need a more sophisticated appraisal of the environment than would be obtained by adding together the appraisal criteria for the basic emotions: "Situations that elicit sexual jealousy or moral indignation do not differ from each other merely in the proportions of danger, conspecific challenge, noxiousness and loss that they involve" (102). Second, some complex emotions endure much longer than the basic emotions. What is more, they endure as real psychological processes, not mere dispositions. When a woman's feeling of guilt explains her behavior through a long session of negotiation with her husband and their lawyers, it does more than dispose her to intermittently display affect-program sadness and affect program fear. Third, basic emotions have behavioral consequences of the sort detected by the facial affect coding system. I denied that all complex emotions result in blends of the facial action associated with the known basic emotions. Finally, while basic emotions have reciprocal interactions with more complex cognitive processes, some complex emotions are more directly involved in the control of long-term, planned action. The woman's guilt in the example just given is arguably an intimate part of the thought processes by which she arrives at a decision on which demands to concede and which to resist.

The proposal that complex emotions involve blends of basic emotions is more promising when conjoined to the idea that complex emotions involve additional cognitive activity. The idea that complex emotions are elaborations of basic emotions resulting from the integration of activity in phylogenetically ancient brain structures with activity in the neocortex is currently the most popular proposal to reintegrate the domain of emotion. This is largely because of the work of the neuroscientist Antonio Damasio (1994, 1999). Jesse Prinz has captured the spirit of this approach with the analogy that basic emotions are shots of hard liquor and complex emotions are cocktails in which specific hard liquors are mixed with specific non-alcoholic ingredients (Prinz 2000). The basic emotion is the motivational "kick" in each complex emotional cocktail.

## Emotion and Affective Neuroscience

The “cognitive elaboration” view is a promising approach to complex emotions, provided it is borne in mind that basic emotions are themselves emotions and not only constituents of complex emotion, and that basic emotions and complex emotions are very different from one another. It is all too easy to gloss over these differences. Simon Blackburn, for example, has argued that Damasio’s work refutes my claim that emotion is not a natural kind because, “Empirically, the suggestion that we split the operation of the affect program from ‘higher cognitive emotion’ seems to ignore the most fascinating result of Damasio’s work, which is the extent to which ‘higher-order’ decision making has to harness the limbic system in order to work at all” (Blackburn 1998, 129).

But, empirically, the operation of two kinds of emotions *can* be split, whether this occurs in pathologies like phobias or in the phenomena reported in the literature on “affective primacy” (Öhman 1999, 2002; Zajonc 1980). Basic emotions can occur without the accompanying activity in the neocortex, and basic emotion appraisal processes can reach conclusions contradictory to those reached by full-blown cognitive appraisal of the stimulus situation. Furthermore, there is a longstanding consensus in the literature on the evolution of emotion that this is why humans still have basic emotions. Basic emotions are rapid acting, failsafe devices that produce evolved behavioral, physiological, *and cognitive* responses tailored to certain critical features of the environment. They are faster and more reliable than the slower, but arguably more accurate, responses that we make using our recently evolved neocortical resources. This, of course, is entirely consistent with Damasio’s claims about the role of the emotions in rational decision making. That role makes use of one of the outputs of the basic emotional response, which, Damasio argues, is a representation of the physiological changes that have been produced by that initial emotional response. Other outputs to cognition from the basic emotion process include orienting sensory systems to the emotional stimulus and allocating attentional resources to that stimulus.

I suspect Blackburn is interested less in whether basic emotions can occur without complex emotions than in whether complex emotions can occur without basic emotions and thus whether all emotions share a set of core processes. In fact, according to philosopher Louis Charland, a fair-minded reading of contemporary affective neuroscience makes it quite clear that all emotions do, indeed, share a single kind of core process. Affective neuroscience, he argues, provides a general theory of the nature of “emoters,” or “affective systems.” These are “a distinct class of biological systems whose behavior is largely governed by emotion and only explainable in those terms. This is the neurobiological version of the hypothesis that emotion is a natural kind term” (Charland 2001: 151–52).

Charland is referring in part to the idea elaborated by the neuroscientist Paul D. Maclean, in a series of publications from the 1950s to the 1980s, that emotion is a distinctive form of information processing which we have inherited from our distant evolutionary ancestors, and which manifests itself to us as what we call emotion, or “feelings.” These processes are realized in phylogenetically ancient anatomical regions surrounding the brain stem, regions that Maclean termed the “limbic brain.” Charland notes that the leading contemporary neuroscientist of emotion Joseph LeDoux regards the limbic brain concept as more or less anatomically and functionally meaningless, but he sets against this the fact that LeDoux’s experimental work has dealt with only one emotion—fear—and that his views are not shared by all other neuroscientists. The views of another leading neuroscientist of emotion, Jaak Panksepp, are more congenial to Charland. Panksepp accepts that MacLean’s concept of an “emotional brain” is oversimplified but defends the underlying concept that emotion represents an ancient form of information processing that we share with many other species. In Charland’s view, the work of Panksepp, Damasio, and others suggests very strongly that there is a single, natural kind of psychological process that generates affect. He takes it to be obvious that we should identify the category of emotion with this class of “affective” processes.

I am sympathetic to the view that the basic emotions represent a distinctive form of information processing that humans share with many other animals (91–97) but I have argued that identifying the emotions with that form of processing would amount to a substantive and stipulative revision of the vernacular concept of emotion (230–34). This revision would both exclude some things that are in the vernacular category and include some things that are not in the vernacular category. In my book I was concerned to leave room for the possibility that some complex emotions may not involve basic emotions, or may involve them only peripherally.<sup>4</sup> I still think it is important not to foreclose that issue, but here I will concentrate on the opposite problem, namely that the new category is too inclusive to be simply identified with the vernacular category. There is, of course, nothing wrong with scientists deciding to use the term *emotion* in a revised sense, and most leading emotion neuroscientists seem to be aware that they are in the business of productive stipulation, not conceptual analysis. In his own response to my claim that emotion is not a natural kind, Damasio remarks: “At this point, my preference is to retain the traditional nomenclature, clarify the use of the terms, and wait until further evidence dictates a new classification, my hope being that by maintaining some continuity we will facilitate communication at this transitional stage. I will talk about three levels of emotion—background, primary and secondary. This is revolutionary enough, given that background emotions are not part of the usual roster of emotions” (1999, 341).

Thus, rather than vindicating the vernacular category of emotion, Damasio is recruiting the old term for a new and more general category. This category covers phylogenetically ancient “primary emotions” (basic emotions), “secondary emotions” that are elaborations of these ancient responses into complex and variegated responses that involve substantial areas of the neocortex and uniquely human cognitive abilities (similar to my “complex emotions”), and background emotions, which are the ever present awareness of our own body.

This perspective on how affective neuroscience uses the emotion category is reinforced by the way in which the discipline treats basic drives like hunger and thirst and hedonic states like pain and pleasure. Panksepp remarks that he will follow tradition in distinguishing between emotions proper and drives that regulate some specific state of the body, and thus that he will not initially consider hunger, thirst, or disgust to be emotions. But this distinction, he remarks, will become “less defensible” in later chapters of his book (Panksepp 1998, 47, 342 n. 17). In the same vein, Damasio remarks, “I will refer to drives and motivations and pain and pleasure as triggers or constituents of emotions, but not as emotions in the proper sense. No doubt all these devices are intended to regulate life, but it is arguable that emotions are more complex” (341). These remarks exemplify a research strategy I take to be central to much recent work on the emotions. Panksepp and Damasio take the domain of affective neuroscience to be all mental processes that involve affective feelings. Panksepp has described this as “the basic psychological criteria that emotional systems should be capable of elaborating subjective feeling states that are affectively valenced” (Panksepp 1998, 48). But this domain is both much broader than the vernacular category of emotion and somewhat difficult to work with in practice, because subjective feeling states “have so far defied neural specification” (48). In practice, then, emotions are defined as that class of affective processes that have certain distinctive performance features: “(1) Various sensory stimuli can unconditionally access emotional systems; (2) Emotional systems can generate instinctual motor outputs, as well as . . . (3) modulate sensory inputs. (4) Emotional systems have positive feedback components that can sustain emotional arousal after precipitating events have passed. (5) These systems can be modulated by cognitive inputs and (6) can modify and channel cognitive activities” (Panksepp 1998: 48).

A comparison of these criteria with descriptions of the distinctive features of affect programs or basic emotions in the Tomkins-Izard-Ekman tradition reveals striking similarities (Griffiths 1997, 77–99). In both research traditions, emotion is being conceived as information processing that is somehow simpler than paradigm examples of cognition, perhaps because it involves bodily feelings rather than explicit representations of external states of affairs, but which is nevertheless more complex than tropisms, reflexes, and homeostatic drives. In my book I used Paul Ekman’s argument

that the startle reaction is not an emotion as a paradigm of what it takes to establish that some reaction is a basic emotion (Ekman, Friesen, and Simons 1985). Ekman and his collaborators argue that the startle reaction is too reflex-like to count as an affect program; I commented that "this suggests that the information processing arrangements underlying startle are not of the same sort as those underlying affect programs. . . . Extending the concept of an affect program state to cover it would not be a positive step in theory construction, since findings about startle may not be true of the affect program states and vice-versa" (241; see also Robinson 1995).

In my view, what Charland fails to appreciate is that the category of felt affective states is very large, and certainly larger than the existing vernacular category of emotion. In addition to classic emotions like anger and joy, it will include drive states like hunger and thirst, hedonic states like pain and pleasure, and the ubiquitous awareness of bodily states that Damasio calls "background emotion." Furthermore, given the strong connection between motivation and affect made by writers like Damasio, the category of felt affective states probably includes desires and preferences. Blackburn, for example, has argued that the motivational power of our long-term goals reflects an emotional commitment to those goals and cites with approval Damasio's idea that our awareness of the body is the motive power of practical reason (Blackburn 1998, 129). Because the category of felt affective states is so broad, it is natural within this framework to seek distinctive *kinds* of processes involving affective feelings. Hence, I suggest that rather than demonstrating that emotions are a natural kind, the empirical success of current approaches to affective neuroscience would establish that there is a very large domain of affective and motivational phenomena, within which we could distinguish categories of state and process such as Damasio's background emotions, primary emotions and secondary emotions, homeostatic drives, and pure hedonic states like pain and pleasure. The scientific domain of affective neuroscience would be like the scientific domain of chemistry, where atoms and their bonds are at the basis of everything, but where lumping together mixtures, compounds, alloys, pure elements, and pure isotopes on the grounds that they are all "chemical substances" is not very helpful. Making and exploring those distinctions was essential to the development of modern chemistry and I suggest that the same will be true of affective neuroscience.

#### MULTILEVEL APPRAISAL THEORY AND THE CONTENT OF EMOTIONAL REPRESENTATIONS

##### Appraisal and the Philosophy of Emotion

Appraisal theories are the closest scientific equivalents of the theories that have dominated philosophy of emotion since the 1960s. Like appraisal the-

orists, philosophers have sought to understand emotions in terms of the situations that elicit them. Emotions have been analyzed as states with specific “formal objects” (Kenny 1963), as evaluative judgments (Solomon 1976), as evaluative judgments that cause bodily arousal (Lyons 1980), as feelings of comfort or discomfort directed toward an evaluative thought (Greenspan 1988), and as the results of either true belief or uncertainty about emotion-inducing situations (Gordon 1987). For these and many other authors, the central aim of a philosophical theory of emotion is to identify the *content* of an emotion: the actual or imagined state of affairs in the world that corresponds to that emotion.

Appraisal theories aim to identify the features of an emotion-eliciting situation that lead to the production of one emotion rather than another (Scherer 1999). Typical appraisal theories include a set of dimensions against which potentially emotion-eliciting situations are assessed. The dimensions of the emotion hyperspace might include, for example, whether the eliciting situation fulfills or frustrates the subject’s goals or whether an actor in the eliciting situation has violated a norm. Richard Lazarus’s well-known model of emotional appraisal has six dimensions, and the regions of the resulting hyperspace that correspond to particular emotions are summarized by Lazarus as the “core relational themes” of those emotions. Anger, for example, is elicited by the core relational theme “a demeaning offence against me and mine,” sadness by “having experienced an irrevocable loss,” and guilt by “having transgressed a moral imperative” (Lazarus 1991). These themes correspond to each emotion’s “content” in philosophical theories of emotion.

The ongoing effort to test appraisal theories has produced a consensus that actual emotional responses do not walk in step with subjects’ evaluation of stimulus situations unless the notion of “cognitive evaluation” is broadened to include subpersonal processes (Teasdale 1999). Many appraisal theorists have come to accept that even such apparently conceptually complex appraisals as Lazarus’s core relational themes can be assessed (1) without the information evaluated being available to other cognitive processes, (2) before perceptual processing of the stimulus has been completed, and (3) using only simple, sensory cues to define where the eliciting situation falls on the dimensions. These conclusions are congruent both with the older literature on “affective primacy” (Öhman 1999, 2002; Zajonc 1980) and with the recent mapping of multiple neural pathways to the same emotional response (LeDoux 1996, 1993).

One of the reasons philosophers remain convinced that emotion is a natural kind is their belief that, however psychologically different the various instances of anger or any other emotion may be, every instance shares the same or similar *content*. Moreover, the general category of emotion is unified by the fact that all emotions have a certain broad kind of *evaluative content*. Thus, to take a prominent recent example, Martha Nussbaum ar-

gues that emotions are all of one kind because they are all “intelligent responses to the perception of value” (Nussbaum 2001, 1). In defending this view she draws on psychologist Richard Lazarus’s classical presentation of appraisal theory, which she describes as “in all essentials the view of emotions I have defended in Chapter 1” (109). Nussbaum’s treatment of the emotions in children and animals also makes use of something like multilevel appraisal. Animals as well as humans make the evaluative judgments that constitute emotions according to Nussbaum’s theory, but they do so without self-conscious awareness and in such a way that the content of their judgments cannot be rendered in language without distortion. Nevertheless, she argues, emotion remains primarily an intentional phenomenon. Despite the existence of low-level appraisal that cannot be expressed in language, “emotions include in their content judgments that can be true or false, and good or bad guides to ethical choice” (1). “What we need, in short, is a multifaceted notion of cognitive interpretation or seeing-as, accompanied by a flexible notion of intentionality that allows us to ascribe to a creature more or less precise, vaguer or more demarcated, ways of intending an object and marking it as salient” (129).

Nussbaum’s description of what is needed for an adequate account of emotional cognition is compelling, but levels of appraisal do not just differ *between* organisms. They also differ *within* a single organism. Multilevel models of emotional appraisal suggest that the same stimulus can be represented in several places in the human brain by different representations. Hence it is vital to understand not only what these multiple appraisals have in common, but also how they *differ* and how they *interact*. The existence of multiple representations in a “hierarchical” emotional architecture (Delancey 2001) violates a key assumption of most philosophical reasoning about emotion, which is that emotional cognition manipulates emotional representations on the basis of their content, and thus that emotional processes can be explored via the semantic “logic” of emotions: “All emotions presuppose or have as their preconditions . . . certain sorts of cognitions—an awareness of danger in fear, recognition of an offense in anger, appreciation of someone or something as lovable in love. Even the most hard-headed neurological or behavioral theory must take account of the fact that no matter what the neurology or the behavior, if a person is demonstrably ignorant of a certain state of affairs or facts, he or she cannot have certain emotions” (Solomon 1993, 11).

Multilevel models imply that this picture is too simple. How emotional and other representations interact, if they interact at all, depends on details of cognitive architecture as well as on the content of the representations. This architecture, of course, cannot be determined by studying the conceptual relations between the contents of emotional representations. Phobias and affective primacy phenomena provide insight into the architecture of the emotion system by revealing that certain information, such as partially

analyzed visual data, is available to low-level appraisal but not to high-level appraisal.<sup>5</sup> A complimentary insight is provided by people with “flattened affect,” who are apparently able to carry out high-level appraisal but not low-level appraisal and who do not experience the physiological components of normal emotional response. The possibility of flattened affect without intellectual impairment reveals that only low-level appraisal has direct connections to the effector systems for the automated components of rapid emotional response.

Normal human emotion involves several subsystems that interact, and interact with other cognitive subsystems, in ways that reflect the particular cognitive architecture in which they are embedded. Nussbaum suggests that we can cope with this phenomenon using a “flexible notion of intentionality.” This is supposed to allow us to identify what is in common between animal and human emotion. The suggestion is presumably that there is some degree of isomorphism between the way in which high-level representations relate to one another on the basis of their content and the way in which low-level representations relate to one another on the basis of their “content.” Thus, the “logic” of the emotion will be the same in the two cases. But in my view, there are likely to be more radical differences between the representational states involved in low-level and high-level appraisal.

Ruth Millikan (1996) has suggested that mental representations in simple organisms may unite the functions of beliefs and desires. Low-level appraisal in humans seems to manifest the same “collapse of the attitudes.” Consider the low-level appraisal of the core relational theme “a demeaning offence to me and mine” that presumably occurs when a soccer player is dribbling the ball down the field, another player grabs his jersey causing him to lose the ball, and the first player turns *angrily* toward the second. It is misleading to say that the relevant brain region *believes* that the core relational theme has been instantiated. Beliefs are mental states that represent how things are and which produce action in conjunction with desires—representations of how the world should be. But in low-level appraisal for anger there is no question as to what action will be taken. The frustrated player in our example will orient to the stimulus, produce the pan-cultural facial expression of anger, and undergo physiological changes to prepare for aggressive action. The “affective computation” in this example is simultaneously the belief that the world is a particular way and the intention to act in a particular way. Likewise for the better understood case of affective computing of fear in the amygdala, and, presumably, for any emotion that has a clear behavioral signature and can be induced to exhibit affective primacy. I suggest that it is simply misleading to describe low-level appraisal as evaluative judgment, or using any other locution derived from a psychology that presumes a fundamental distinction between data and goals. Instead, low-level emotional appraisal seems to involve

action-oriented representation (Griffiths forthcoming b; Scarantino forthcoming).

Another way in which low-level emotional appraisal may differ from high-level is in terms of the narrow inferential role imposed on low-level representations by the task-specific architecture in which they occur. The inferential role of these representations is impoverished in three ways. First, low-level appraisal processes do not have access to most of what is represented elsewhere in the brain, which is why knowledge that the cockroach in my drink has been completely sterilized does not eliminate the disgust response. Hence, many inferences that would seem to follow from the content we ascribe to this low-level appraisal—"I am taking in or being too close to an indigestible object"—are not actually made by subjects because they cannot recombine that content appropriately with their other contentful states. Second, the processes of affective computing, as opposed to their final output, are not available for inspection by other cognitive subsystems. Once again, architectural barriers to information flow block inferences that follow from what would otherwise seem the natural content to ascribe to those states. Finally, the inferential principles used in affective computing are not truth-preserving, but heuristically survival-enhancing. It does not follow by any reasonable deduction that if I have been poked hard and unexpectedly in the small of my back then I have suffered "a demeaning offence to me and mine" but the automatic appraisal mechanism for anger will reliably draw that conclusion.<sup>6</sup>

If the concepts that figure in the content ascribed to a representation do not have their usual inferential role, then what is meant by attributing that content? The differences between the role of the representations involved in low-level appraisal and the inferential role of the content-sentences with which we ascribe those appraisals strongly suggests that, at least when applied to these low level processes, appraisal theories are not theories of cognitive content. I have suggested elsewhere that they are theories of the ecological significance of the environment to the organism (Griffiths forthcoming b). That significance is tracked by multiple cognitive subsystems using different environmental cues and *different psychological and neurological mechanisms*. This, of course, returns us to the main theme of this chapter, which is that the states and processes we call "emotion" are not all of the same kind.

#### 4. CONCLUSION: WHY EMOTION IS STILL NOT A NATURAL KIND AND WHY IT MATTERS

I have defended the thesis that the psychological states and processes that fall under the vernacular concept of emotion are unlikely to be a single "natural kind." I believe that the same is probably true of some specific

vernacular categories of emotion, although I have not defended that view here. I have considered two main alternatives. The first is that the emotions are simply those of states and processes studied in “affective neuroscience.” I have argued that, if empirically successful, current approaches to affective neuroscience will define a much larger class of affective and motivational processes, united by the role in all these processes of felt affect. I also think it is important not to foreclose the possibility that some emotions do not involve the basic emotions or involve them only peripherally, although I have not argued for this here.

The second suggestion was that emotions are appraisals of the significance of the environment to the organism. I agree that multilevel appraisal theory brings many different kinds of emotion under a single taxonomic scheme, but I have argued that it does so precisely by abstracting away from the kind of psychological processes that constitute those emotions. A theory of emotions based on their content is not a psychological theory, but an ecological theory, as I have argued at more length elsewhere (Griffiths forthcoming b).

It is reasonable to ask why the claim that emotion is not a natural kind matters to philosophy. The simple answer is that many philosophers still take it to be their role to provide an account of the genesis, development, and consequences of a “typical” human emotion (e.g., Wollheim 1999). If I am correct, then there is no such thing as a *typical* emotion. Instead, there are different kinds of emotion, or of emotional process, each of which should be treated in its own terms and whose various possible interactions should be studied. Similarly, the idea that all emotions are intentionally directed at aspects of the environment *in the same sense* is a core methodological assumption of much current philosophical work on the emotions. If I am correct, then we should be more concerned with the distinctive properties of the different kinds of emotional intentionality and with how these different kinds of emotion process interact in real emotional episodes. Cases in which people have an emotion in one sense and do not have it in another should be as illuminating for the philosophy of psychology as they have been for psychology itself.

## NOTES

1. Parenthetical page numbers refer to Griffiths 1997 throughout.

2. See, e.g. *Nature* no. 391, 1998; *Times Literary Supplement*, July 17, 1998; *Philosophy in Review* 18 (1998); *Australian Review of Books*, April 1998; *Metascience* 8.1 (1999); *Australasian Journal of Philosophy* 77.4 (1999); *Philosophical Review* 108.1 (1999); *American Journal of Psychology* (fall 2000); *Dialogue* 38.4 (2000).

3. I take this to be an open question. Many psychologists of emotion reject the idea that emotions come in discrete types and support instead a dimensional account of the emotion system. Even if basic emotions form discrete types, the

best theory of complex emotions might still be dimensional rather than typological.

4. One reason I take this possibility seriously is that contemporary research in evolutionary psychology does not proceed under the assumption that all emotions are composed of or have at their core one of the basic emotions. In fact, so-called Santa Barbara school evolutionary psychologists make considerable play of the claim that “all emotions are equally basic” (Buss 2000; Cosmides and Tooby 2000; Gaulin and McBurney 2001). These researchers define an emotional adaptation as any motivational mechanism designed to influence behavior in some specific problem domain and whose operation cannot be understood as the application of domain-general processes to that problem. Their commitment to the “massive modularity hypothesis” makes them suspicious of attempts to explain all specific emotions via the interaction of a smaller number of general-purpose mechanisms, such as the basic emotions and our capacity for cognitive evaluation of stimulus situations.

5. I do not think it is yet clear how many appraisal levels are needed for an adequate representation of emotional appraisal. When I talk of “low” and “high” levels I mean not to imply that there are just two levels of appraisal, but rather to mark the existence of a dimension of difference.

6. Some would say that the appraisal mechanism has “innate knowledge” that this cue reliably predicted conspecific aggression in ancestral environment (Tooby and Cosmides 1992), but the scientific substance of this claim is simply that the appraisal mechanism consistently makes certain inferences.

## Emotion as a Subtle Mental Mode

AARON BEN-ZE'EV

Over a hundred years ago, William James asked, "What is an emotion?" Today, the answer still remains unclear. Several replies have been proposed in response to this question, but they all seem to be unsatisfactory. After briefly describing some of these replies, I propose considering emotion as a subtle general mental mode. I then describe this mode by referring to its logic, psychological mechanism, and psychological nature.

### PREVAILING CHARACTERIZATIONS OF EMOTIONS

Several suggestions have been made concerning the characterization of the nature of emotions. Major suggestions propose that (a) emotions are either mental states or mental dispositions; (b) emotions are mental capacities; (c) emotions are types of intentional reference; (d) emotions are feelings; (e) emotions are brain states. After indicating the inadequacy of these suggestions, I describe my own proposal: emotions are a general mode of the mental system.

### Is an Emotion a Mental State or a Mental Disposition?

Richard Wollheim divides mental phenomena into mental states, such as perceptions, sensations, dreams, images, and thoughts, and mental dispositions, such as beliefs, desires, knowledge, abilities, virtues, and vices. He considers emotions to be mental dispositions. Mental states are transient

events that make up the lived part of the life of the mind and that occur at a particular time. Mental dispositions are more or less persistent modifications of the mind that underlie this sequence of mental states (Wollheim 1999, 1).

My own use of the term *state* is broader and refers to “a combination of circumstances or attributes belonging for the time being to a person or thing.”<sup>1</sup> In this sense, we can speak about passive, active, dispositional, and actualized states; *states* can refer to both attitudes and activities. Whereas Wollheim limits the use of *state* to an actualized state, my usage also includes dispositional states. Thus far, the discrepancy between us may be merely one of terminology.

Is an emotion an actualized, transient state? Although emotions include such states, they are much more than these states. Emotions also include dispositions and they have some duration and persistence in time. We can speak about emotional states, but we cannot identify an emotion with a certain state. Emotions are more complex than simple actualized, transient states.

Is an emotion a disposition? There is no doubt that emotions include dispositions. Thus, love involves a disposition to behave in a certain manner and to have certain feelings toward the beloved. However, although emotions include dispositions, they are not simply dispositions; they also involve actualized and transient states.

Wollheim argues that emotions manifest themselves in mental states, but that they are not mental states and therefore they do not have the feeling component of a mental state.<sup>2</sup> However, it seems peculiar to maintain that emotions, including love, sexual desire, and anger, lack the feeling component. Some people even identify an emotion with its feeling component. Although emotions are not mere feelings, feelings are a necessary component of emotions. Wollheim may argue that when emotional dispositions are manifested in mental states, they may have the feeling component. However, in light of the crucial role of feeling in emotions, it is implausible to consider emotions as mere dispositions; emotions are richer and more complex than that.

### Is an Emotion a Mental Capacity?

Traditional descriptions of mental phenomena indicate the existence of a few mental capacities (faculties)—for example, sensation (or feeling), perception, memory, imagination, and thought. These capacities are expressed in actualized states, such as having a headache, seeing a tree, remembering one’s first date, imagining the next date, or thinking about the environment.

Although the list of mental capacities is not so disputable, the nature of a mental capacity is far from being clear. For example, it is doubtful

whether each of these capacities can be described as a single, unitary capacity. Thus, memory is probably not a single capacity but actually consists of various learning systems.

Without discussing the nature of a mental capacity, I argue that an emotion is not on the same conceptual level as each of these capacities. A typical emotion involves the activation of sensation, perception, memory, imagination, and thought; hence, an emotion should be conceptualized as a higher-level construct than any of these capacities. An emotion is not like any of these capacities; rather, these capacities are like elements of the emotional experience.

### Is an Emotion a Type of Intentional Reference?

We can discern several types of intentional reference: cognition, evaluation, and motivation. Not all mental capacities involve these types of intentional reference. Sensation, which is the most primitive mental capacity, lacks any of these types; it is not intentional (Ben-Ze'ev 2000, chap. 3). The more complex mental capacities, such as perception and memory, have the cognitive type of reference; the evaluative and motivational types are of lesser importance, if present at all. Imagination and thought often include all these types. These types of intentional reference are essential components of typical emotions, but an emotion is not identical to any of them. There is not a cognitive, an evaluative, or a motivational feature that is present in emotions and not in any other mental mode. Accordingly, I do not think that there is a necessary and sufficient cognitive, evaluative, or motivational condition in light of which emotions can be defined. Emotions should be characterized by referring to more than one of these components.

### Is an Emotion a Mere Feeling?

The feeling dimension is a primitive mode of consciousness associated with our own state. It is the lowest level of consciousness; unlike higher levels of awareness, such as those found in perception, memory, and thinking, the feeling dimension has no significant cognitive content. It expresses our own state, but is not in itself directed at this state or at any other object. Since this dimension is a mode of consciousness, one cannot be unconscious of it; there are no unfelt feelings. In the intentional domain we play a more active role; feelings, on the other hand, just seem to surface and can overcome us when they are intense (Ben-Ze'ev 2000, 64–67).

In contradiction to the philosophical and psychological traditions identifying the emotions with feelings, there are also scholars who argue that feelings are not essentially involved in the concept of emotion. I reject both views: for the same reasons that the identification of emotions with feelings should be rejected, the attempt to detract from the importance of the feeling

dimension in emotions should be rejected (Ben-Ze'ev 2002, 2003a). Feelings are more than a frequent accompaniment of emotions; they are a constitutive element, as they play a crucial role in expressing the urgency of emotional situations. Despite the importance of feelings in emotions, equating the two is incorrect since emotions have an intentional component in addition to the feeling component. Unlike the complex experience of emotions, mere feelings are more localized in space and time and are not intentional. Mere feelings are more passive states than emotions are. Moreover, people can experience and identify feelings that are typical of a certain emotion without experiencing the emotion itself.

### Is an Emotion a Brain State?

It has become popular these days to be “modern” and “scientific” and to discard “psychological nonsense” and concentrate on “real” science, namely neuroscience. According to this view, emotions are nothing but brain states—or other neurological structures—and hence a psychological research of the emotions is of little explanatory value. The difficulties of such a view go back to the fundamental issue of the mind-body problem, which I cannot discuss here. I merely want to cite Aristotle’s claim that anger can be described on two major levels. A scientist may describe anger as a boiling of the blood and the presence of heat around the heart, and a philosopher may describe anger as the desire to retaliate by returning evil for evil (Aristotle 412b19).

The desire to retaliate cannot be found in the boiling blood—which is, however, a necessary supporting basis for that desire. To explain that desire, we have to refer to the evil that was inflicted and not to the boiling blood. We cannot understand the nature of higher-level phenomena—for instance, emotions—by merely studying discrete lower-level entities, such as neurological activities; nevertheless, understanding the latter may be relevant to understanding the former. A clear distinction between the various levels of description is essential for dealing with complex phenomena such as the emotions.

Traditional descriptions of mental phenomena are not suitable for describing emotions because they fail to take into account the complexity of emotions. In my view, an emotion is neither a mental capacity nor an intentional mode of reference; it is also not a mere feeling or a mere brain state. It consists of various mental elements, such as dispositions, states, capacities, types of intentional references, attitudes, activities, and feelings.

In light of this complexity, it is preferable to replace the substantial notion of “emotion” with a functional concept. An emotion is not a noun, but rather an adjective. For the purpose of an initial explanation, people may refer to an emotion as an entity, but when a more precise and scientific explanation is required, a functional explanation is in order (Cassirer

1923). Instead of considering an emotion as a single entity, we should consider it as a mode of the mental system that involves all the above mental elements.

The preceding discussion did not attempt to refute classical answers concerning the nature of emotions. Its purpose was to indicate some of the difficulties of such views and to pave the way for considering the emotions as a general mental mode.

### Mental Modes

An emotion is a general mode (or style) of the mental system. A general mental mode includes various mental elements and expresses a dynamic functioning arrangement of the mental system. The kinds of elements involved in a certain mode and the particular arrangement of these elements constitute the uniqueness of each mode. The emotional mode involves the activation of certain dispositions and the presence of some actualized states. It also includes the operation of various mental capacities and the use of different kinds of intentional references. This mode involves cognition, evaluation, motivation, and feeling.

Other possible modes are, for example, the perceptual and the intellectual modes. The perceptual mode is the most basic mental mode. It involves being aware of our immediate environment without being engaged in a complex intellectual activity and without being in the midst of a stormy emotional experience. Perception is the major capacity used in the perceptual mode; other capacities involved in this mode are memory and imagination. A more extensive use of the latter capacities is found in the intellectual mode, where thinking is the basic capacity.<sup>3</sup>

A given mental mode is not necessarily the complete opposite of another mode; they may differ in a few, but not all features. For example, perception is found in all modes—although in different forms. Similarly, while feelings are intense in the emotional mode, they are not essential for the perceptual and intellectual modes. Thinking dominates the intellectual mode, but not the perceptual and emotional modes.

Each mental mode is more or less dominant in various periods of our lives and of the evolutionary development of human beings. The perceptual mode is dominant in our early childhood and in the primitive stages of human development. The emotional mode may also have been present in these periods, but in a less complex manner—this mode is dominant during our late childhood and adolescence. The intellectual mode is characteristic of more advanced stages of personal and evolutionary development.

Mental modes are not isolated entities, but prototypes of various mental phenomena: one prototype is typical of our usual everyday situation, the second is typical of abstract thinking associated with scientific, detached calculations, and the third is typical of intense, stormy emotions. In be-

tween these, there are many types of phenomena that do not fit perfectly into any of the three prototypes. Moreover, some of the features constituting a mental mode admit of degrees and hence no precise borderline is possible between the various modes. Nevertheless, the description of the prototypes of mental modes has a great explanatory value.

A distinction may be made between general and specific mental modes. A general mode is comprehensive in the sense that it involves most types of mental elements; a specific mode involves only a few mental elements. I believe that the emotional mode is the most comprehensive mode since it typically involves more types of mental elements than any other mode. The intellectual mode is less comprehensive; it may not include, for instance, mental elements such as perception, feelings, motivation, and evaluation. Certainly, these elements may be found in some occurrences of the intellectual mode, but they are not constitutive or even typical of this mode. The perceptual mode is probably even less comprehensive, as it does not involve the activation of several mental capacities.

We may distinguish between elements that are constitutive of a certain mode and elements that are frequently associated with it. Examples of constitutive elements of the emotional mode are cognition, evaluation, motivation, feeling, instability, great intensity, partiality, and brief duration. Specific actions are typically associated with emotions but are not constitutive elements of it; action readiness is such a constitutive element. Thinking is a constitutive element of the intellectual mode; feeling, motivation, and evaluation are sometimes associated with this mode, but this mode is sometimes present without them.

A mental mode is typically complex, structured, episodic, and dynamic (Goldie 2000, 12–13). It is complex, since it involves many elements; it is structured, in the sense that the elements are arranged in a certain organized manner; and it is dynamic, as it typically undergoes changes in the particular manifestations of its constitutive elements, in the kinds of associated elements involved, or in the relationships among them. A mental mode is also episodic as its duration is limited.

I suggest distinguishing the various mental modes on the basis of a few categories:

- a. Basic *psychological features*;
- b. Basic types of *information-processing mechanisms*;
- c. Basic *logical principles* of information processing.

The category of psychological features may be divided into characteristics, such as complexity, instability, and intensity, and components, such as evaluation and motivation. Schematic mechanisms and deliberative mechanisms are examples of information-processing mechanisms. Logical principles are those determining the significance of events—for example, whether change or stability is of greater significance.

In what follows, I describe the emotional mode in light of the above categories. In doing this, I will compare the emotional attributes to their corresponding attributes in the intellectual and, to a lesser extent, in the perceptual mode. Since my focus is on the emotional mode, the description of the other modes is brief and has mainly illustrative value; more detailed discussions may formulate these features somewhat differently.

#### PSYCHOLOGICAL FEATURES OF THE EMOTIONAL MODE

In characterizing an emotion as a mental mode, I limit myself to the mental level of description. Emotions can also be explained on other levels, such as the biological and neurological. Accordingly, I could have said that an emotion is a general mode of the person and not merely of the mental system of a person. In this case, I would have to analyze other relevant levels of discussions—for instance, the neurological, biological, and sociological level. Since my discussion here is restricted to the mental level, I can characterize an emotion as a mode of the mental system.

The basic psychological features of the emotional mode may be divided into basic characteristics and components. Among the major characteristics, we may refer to complexity, stability, intensity, partiality, and duration. The basic components are cognition, evaluation, motivation, and feeling. Whereas emotional characteristics are properties of the whole emotional experience, emotional components express a conceptual division of the elements of this experience. It is arguable that one could perhaps find a few relevant characteristics other than those mentioned above; however, the conceptual division of emotions into four components is more comprehensive and is intended to cover all possible components.

#### Typical Characteristics

In describing the basic characteristics of a mental mode, I refer to the following features: complexity of the mode, the stability and dynamism of the mode, the extent (intensity) of using the various capacities, the focus of the various modes of reference, and the duration of a typical state in such a mode. The emotional mode can be characterized as highly complex, greatly unstable, highly intense, using a partial perspective, and lasting for a relatively brief period. In contrast, the intellectual mode is typically somewhat less complex, more stable, not so intense, having a broader perspective and longer duration. In comparison with the emotional mode, the perceptual mode is less complex, more stable, less intense, has more restricted perspective, and can last longer.

*Complexity.* The complexity and subtlety of the emotional mode is associated with the fact that it consists of elements belonging to various ontological levels; hence, its explanation requires reference to various levels of discussion. Thus, the basic emotional components—cognition, evaluation, motivation, and feeling—are on a different ontological, as well as epistemological, level than the basic emotional characteristics, such as instability, great intensity, partiality, and brief duration. Even when discussing the basic components, we can find such ontological and epistemological differences; the intentional components of cognition, evaluation, and motivation differ from the feeling component. All of these are still on a different epistemological and ontological level than that of mental capacities, such as memory, perception, imagination, and thought, which are also elements of the emotional mode. Similarly, dispositions and states are on yet another level of discussion, and they have a different ontological status. Moreover, even the component of evaluation (appraisal) occurs at multiple levels. In light of such complexity, one cannot characterize the emotional mode as a mere collection of components.

The complexity and subtlety of the emotional mode are also related to its highly *structured* nature. This mode, which consists of diverse kinds of mental elements, is not a theoretical mode; on the contrary, it is highly practical. Action readiness is an important constitutive element within it. In order to maintain its practical value, the various elements should be highly structured. For example, anger includes a negative evaluation of the other's action, negative types of actions toward this person, and unpleasant feelings. The link between these components has obvious evolutionary advantages: pleasant feelings are an important motivating force for doing things we positively evaluate. There are also such correlations between the basic emotional characteristics. The more intense the emotion is, the more partial is the emotional attitude and the more unstable is the mental mode.

Another feature related to the complexity and subtlety of the emotional mode is its comprehensiveness. More than any other mental mode, the emotional mode involves the activation of most mental elements. In light of the importance we attach to emotional situations, the emotional mode usually involves the activation of all mental capacities, for example, sensation, perception, memory, imagination, and thought. In the other mental modes, a certain capacity is clearly more dominant than others, some of which may not even be present at all. Low intensity of the feeling dimension, as well as of other mental components, usually prevails in the perceptual mode. Although the intellectual mode is characterized by low intensity of the feeling component, it involves the activation of some other capacities. Thinking, and especially abstract thinking, is more active and complex in the intellectual mode than in the emotional and perceptual modes. Thinking is also present in the emotional mode but is less complex

and has a more limited role. Perception usually plays a greater role in the emotional mode than it does in the intellectual mode. The capacities of imagination and memory may be activated in both modes, but usually they are activated in a different form.

The intellectual and perceptual modes are less complex, structured, and comprehensive than the emotional mode. This is particularly true of the perceptual mode, since perception is typically part of the emotional and intellectual modes.

*Stability.* Emotions typically occur when we perceive positive or negative significant changes in our personal situation—or in that of those related to us. In light of the crucial role that changes play in generating emotions, and the fact that the system does not know yet how to cope with this change, the mental mode is highly unstable and dynamic (Nussbaum 2001). Emotions indicate a transition during which the preceding context has changed, but no new context has yet stabilized. Hence, there is a continuous updating of the mental mode until the system can cope with the new situation.

The perceptual and intellectual modes are much more stable, as changes are less significant to them. Sometimes, the mere presence for a long time of the perceptual mode is typical of indifferent people who are unresponsive to and detached from changes in their situation; they remain stable in the face of such changes. Endurance, evenness, and lack of fluctuation characterize the life of such people. Abruptness, changeability, and volatility characterize the life of people in whom the emotional mode is central.

*Intensity.* Emotions are intense experiences. In emotions, the mental system has not yet adapted to the given change, and due to its significance, the change requires the mobilization of many capacities and resources. Accordingly, emotions are perceived as associated with urgency and heat. One basic evolutionary function of emotions is indeed that of immediate mobilization of resources and of focusing them upon the perceived significant change. Focusing upon fewer objects increases the resources available for each and hence increases emotional intensity. It is like a laser beam that focuses upon a very narrow area and consequently achieves high intensity on that point.

The other modes, which do not necessarily involve significant personal concerns, are, generally speaking, of lesser intensity. No doubt thinking is more intense in the intellectual mode than in other modes. But all other mental capacities and elements are used in a more intense manner in the emotional mode. The perceptual mode is the least intense, as it is the mode prevailing when everything is normal and no significant changes or problems are perceived.

*Partiality.* Emotions are partial in two basic senses: they are focused on a narrow target, such as one person or a very few people; and they express a personal and interested perspective. Emotions direct and color our attention by selecting what attracts and holds our attention; they make us preoccupied with some things and oblivious to others. We cannot assume an emotional state toward everyone or those with whom we have no relation whatsoever.

The important role played by the personal evaluative concern in emotions is a major reason for their partiality; this role is less evident in other modes. The perceptual mode is partial in the sense that it is focused on a relatively narrow target—that is, the environment that is present at that moment. However, it is less partial in the sense that it does not express a personal and interested perspective. The intellectual mode is less partial in both senses.

*Duration.* Typical emotions are essentially transient states. The mobilization of all resources to focus on one event cannot last forever. A system cannot be unstable for a long period and still function normally; it may explode due to a continuous increase in emotional intensity. A change cannot persist a very long time; after a while, the system construes the change as a normal and stable situation. The association of emotional intensity with change causes the intensity to decrease steadily due to the transient nature of changes. The transient nature of emotions does not imply that emotions can only last a few seconds: sometimes the transition from one stabilized state to another takes longer. Such a transition is not just a switch from one state to another; it involves profound changes in our plans and concerns and, as such, it may occupy us for some time.

The above factors, which limit the duration of the emotional mode, are absent in the other modes, and hence they can last for a longer period of time.

### Basic Components

In addition to the above basic characteristics, I suggest describing the emotional mode as consisting of four basic components: cognition, evaluation, motivation, and feeling. The first three components are intentional; the fourth is not. The cognitive component consists of information about the given circumstances; the evaluative component assesses the personal significance of this information; the motivational component addresses our desires, or readiness to act, in these circumstances. When a person is in love with someone, the feeling dimension surfaces in a particular feeling, say a thrill, that is experienced when they are together; the intentional dimension is expressed in the person's knowledge of her beloved, her evaluation of his attributes, and her desires toward him.

The presence of all components in the emotional mode stems from the central place in emotions of the evaluative component, which appraises the factual information presented by the cognitive component, in terms of its implications for personal well-being. Since emotions involve evaluative attitudes, including a positive or a negative stance toward the object, they also entail being disposed to act—and often also taking action—in a manner compatible with the evaluation. The extreme evaluation, and hence the strong urge to act, which are typical of emotions, entail an intense feeling component as well.

Other mental modes may not include all four basic components. Thus, an intense feeling component is typically absent from both the perceptual and intellectual modes. The same is true for the motivational component, and to a lesser extent also for the evaluative component. Thinking, which is the dominant activity of the intellectual mode, can lack intense feeling, the motivational component, and sometimes also the evaluative component. In the perceptual mode, the evaluative and motivational components are also of lesser importance.

#### INFORMATION-PROCESSING MECHANISMS OF THE EMOTIONAL MODE

The second feature that distinguishes the various mental modes concerns its information-processing mechanisms. In light of the different features of each mental mode, it is plausible that each mode relies upon a different information-processing mechanism. In this regard, a distinction can be made between two major kinds of mechanisms: schematic and deliberative. Whereas a schematic mechanism is most typical of the perceptual mode, and to a lesser extent also of the emotional mode, a deliberative mechanism is most typical of the intellectual mode.

A schematic mechanism is a kind of dispositional mechanism expressing past knowledge. It typically involves spontaneous responses depending on a more tacit and elementary evaluative system. Schematic activity is typically fast, automatic, and is accompanied by little awareness. It is based upon ready-made structures or schemes of appraisal that have already been set during evolution and personal development; in this sense, history is embodied in these structures. Since these schemes are part of our psychological constitution, we do not need time to create them; we simply need the appropriate circumstances to activate them. A deliberative mechanism typically involves slow and conscious processes, which are largely under voluntary control. Such processes usually function on verbally accessible, semantic information and operate in a largely linear, serial mode. The dependence of such mechanism on past knowledge is much more limited.<sup>4</sup>

The two mechanisms may conflict. Thus, we may persist in being afraid even when our conscious and deliberative judgment reveals that we are no

longer in any peril. We can explain such cases by assuming that certain schematic evaluations become constitutive to a degree where no intellectual deliberation can change them. This corresponds to situations in which intellectual knowledge fails to influence illusory perceptual content. Spontaneous evaluations are similar to perceptual discriminations in being immediate, meaningful responses. They entail no deliberative mediating processes, merely appearing as if they were products of such processes.

The spontaneous and schematic nature of emotions does not imply that deliberative thinking has no role in generating emotions: we may think about death and become frightened, or think about our mates and become jealous. Deliberative thinking, however, has a preparatory rather than constitutive role in emotions. Thinking may prepare the system for the activation of schematic evaluations: it brings us closer to the conditions under which evaluative patterns are spontaneously activated. Deliberative thinking may be the immediate stimulus for the activation of an evaluative pattern, but the emotional evaluation itself is nondeliberative. This gives the mental system, while being in the emotional mode, the ability to react almost instantaneously to significant events, and yet to draw upon past knowledge.

The details of information-processing mechanisms underlying various mental modes must be determined by empirical research. Concerning this issue, I am unsure that the above characterization is entirely adequate. I have tried to point out, however, certain conceptual considerations that should be taken into account when analyzing this matter. I have indicated that different psychological features of each mental mode should be expressed in different information-processing mechanisms. It is probably the case that the distinction between deliberative and schematic processing does not exhaust all types of processing underlying the various mental modes.

#### THE LOGIC OF THE EMOTIONAL MODE

The differences between the various modes should also be expressed in the logical principles underlying the information processing of each mode. If indeed such logical principles can be described, we may be able to speak about the logic of emotions and hence to substantiate their rationality.

The logical principles described below should be conceived as the principles of information processing that determine the meaning of the events around us. As is the case with other psychological generalizations, these principles are used by the majority of people in most circumstances, but not in all of them. The implementation of these principles may be influenced by personal or cultural factors.

I divide the logical principles underlying the emotional and intellectual modes into three groups, each concerned with a different type of information.<sup>5</sup> Those types refer to (a) the nature of reality, (b) the impact of the

given event upon the agent, and (c) the background circumstances of the agent.

A. The Nature of Reality

*The emotional mode*

The emotional world consists of the environment I actually perceive or in which I imagine myself to be;

Changes are more significant than stability;

A personal event is more significant than a nonpersonal event.

*The intellectual mode*

The environment that I actually perceive or in which I imagine myself to be constitutes a small portion of the intellectual world;

Changes are not more significant than stability; on the contrary, we should assume that there are stable regularities in the world;

A personal event is not necessarily more meaningful than a nonpersonal event.

B. The Impact of the Given Event

*The emotional mode*

The perceived strength of an event is most significant in determining its impact;

The more real an event is perceived to be, the more significant it is;

Those who are relevant and close are more significant than those who are irrelevant and remote.

*The intellectual mode*

The objective strength of an event is what is most significant;

The significance of an event is not always connected to its perceived reality;

My psychological distance from a certain person is of no relevance in evaluating this person.

C. Background Circumstances of the Agent

*The emotional mode*

The more responsible I am for a certain event, the more significant the event is;

The less prepared I am for a certain event, the more significant the event is;

The issue of whether the agent deserves a certain event is greatly significant in evaluating this event.

*The intellectual mode*

My responsibility for a certain event is in many cases not relevant to its present significance;

My preparedness for a certain event is in many cases not relevant to its present significance;

The issue of whether the agent deserves a certain event is not always significant in evaluating this event.

## A. The Nature of Reality

1. *The scope of reality.* Emotional and intellectual reasoning often refer to different types of realities: emotions are concerned with the immediate and personal reality perceived by the individual subject, while abstract thinking—the prototype of intellectual reasoning—is concerned with a more detached and objective world, which is common to all subjects. Sense perception and vivid imagination increase emotional significance, but they are almost irrelevant for increasing intellectual significance.

The nature of the world is different for the perceptual, emotional, and intellectual agents. Emotional reality consists of immediate, unstable, personal events, which could have been otherwise, whereas intellectual reality is more detached, stable, and deterministic in nature. Perceptual reality is more stable and less personal than emotional reality, but it is more limited in its scope than the scope of the other two types of reality.

2. *Change and stability.* Emotions typically occur when we perceive positive or negative significant changes in our personal situation, or in the situation of those related to us. Like burglar alarms going off when an intruder enters, emotions signal that something needs attention. Accordingly, change is emotionally much more significant than stability (Ben-Ze'ev 2000, 13–17; Spinoza 1677, IIIp6; IIIdef.aff.; Vp39s). The intellect is concerned with the general and the stable, whereas emotions are engaged with the personal and the volatile. The aim of the intellect is to see a specific event as a specific case of general stable regularities; the foundations of intellectual thinking are features common to individual cases. Change and instability are often taken to be the surface phenomena that are governed by stable regularities; the intellectual search is often a search for such regularities.

The emotional system is concerned with a change that is not merely an actual event, but also a potential event. Accordingly, the issue of the availability of an alternative is crucial for the emergence of emotions: the

more available the alternative, the more emotionally significant the event is. In cool, intellectual thinking, the availability of an alternative should have no relevance to the significance of an event. The fact that circumstances could have been different is of no importance in analyzing the present impact of a given event. We may be sad at missing a close opportunity, but our “objective” situation now is the same as it would be if no such opportunity had been available.

3. *Personal concerns.* Emotions address practical concerns from a personal and interested perspective. In the emotional mode, we look at the world from our own personal perspective, applying our own sense of personal importance to various events (Nussbaum 2001, I.V). Intellectual deliberations do not necessarily put the thinker at the center of the world; nor do they necessarily address personal or practical concerns. The thinker is more of a detached, objective observer who is looking for regularities independent of her existence. The perceiver is at the center of the perceptual environment, but its personal characteristics have less influence upon that environment. The perceptual environment is more stable and less personal than the emotional environment.

## B. The Impact of the Given Event

4. *The event's strength.* The strength of an event is an important factor in determining its significance; this is true for both the emotional and intellectual modes. However, whereas emotions are concerned mainly with perceived strength—the strength as the subject perceives it—intellectual deliberations are more often concerned with the objective strength of the event, independent of the subject's attitude toward it. This difference is in accordance with the greater role that personal concerns play in emotions. Perceived strength refers to the subject's personal perception, whereas objective strength is not related to the way a certain individual perceives it.

5. *The degree of reality.* The more we perceive the event as real, the more intense the emotion. In analyzing the notion of “emotional reality,” two major senses should be discerned: (a) an ontological sense referring to whether the event actually exists or is merely imaginary, and (b) an epistemological sense that is typically expressed in the vividness of the content. The degree of reality is highest when the object is real in both senses. In the emotional realm, vividness is often more significant than mere existence. Accordingly, movies often generate more intense emotions than does dry information in a newspaper. The vividness of an event is of little relevance in determining its significance in the intellectual realm. Vividness expresses the subject's personal perspective: something is vivid when it is

close to the subject. In order to determine the actual reality of a certain event, the vividness of the subject's perception is of little relevance.

6. *Relevance and closeness.* The principle of relevance restricts the emotional impact to areas that are particularly significant to us. Emotions express our values and preferences; hence, they cannot be indiscriminate. Emotional relevance is related to emotional closeness. Events close to us in time, space, or effect are usually emotionally relevant and significant. Greater relevance and closeness imply greater significance and greater emotional intensity. In intellectual deliberations, the distance from the observer is typically of no relevance in determining the significance of a certain event. The significance of an event is determined by analyzing its properties and not by considering its distance to a certain person. Personal concerns are not a factor in such analysis.

The logical principles underlying the perceptual mode are close to the emotional mode in the sense that the strong and real event is the one that we perceive; however, as in the intellectual mode, the issue of relevance is of little significance.

### C. Background Circumstances of the Agent

Background circumstances of the agent are more significant in the emotional mode than in the perceptual and the intellectual modes. Such circumstances may influence to a certain degree the perceptual content, but the influence is limited: perception of a given event is quite similar among various perceivers—the similarity is much higher than in the emotional mode, where the event's significance is highly dependant on personal and background circumstances. Background circumstances are also of little value in the intellectual mode. The intellectual negative attitude toward background circumstances is part of the more general negative attitude toward taking the past into consideration.

7. *Responsibility.* Responsibility (or accountability) refers to the nature of the agency generating the emotional encounter. The major issues relevant in this regard are: (a) degree of controllability, (b) invested effort, and (c) intent. The greater the degree of controllability we have, the more effort we invest, and the more intended the result, the more significant the event usually is and the greater emotional intensity it generates. In the intellectual and perceptual modes, the issue of the agent's responsibility is typically of less importance in evaluating the significance of a certain event—in those modes the significance is derived more from present circumstances (and in the intellectual mode, from future implications as well).

8. *Preparedness.* Preparedness refers to the cognitive change in our mind; major factors here are unexpectedness and uncertainty. Since emotions are

generated at times of sudden change, unexpectedness and uncertainty typically generate emotions and are usually positively correlated with their intensity, at least up to a certain point. Like the issue of responsibility, the issue of the subject's preparedness is of little significance in the perceptual and intellectual modes, which are less concerned with past circumstances.

9. *Deservingness.* The perceived deservingness (equity, fairness) of our situation or that of others is of great importance in determining the emotional significance of a certain event. People do not want to be unjustly treated, or to receive what is contrary to their wishes. Accordingly, the feeling of injustice is hard to bear—sometimes even more so than the actual hardship caused. The issue of deservingness, which entails considerations of the subject's past situation and rights, is again of less importance in the perceptual mode, which is mainly focused on present circumstances, and in the intellectual mode, which focuses on both the present and future implications of a certain event.

#### IMPLICATIONS OF THE PROPOSED VIEW

I have characterized the emotional mode as having certain psychological features, certain information-processing mechanisms, and some logical principles of information processing. Are there connections among these categories?

The connection between the first category and the other categories is evident. Thus, I have suggested characterizing the emotional mode as greatly complex and unstable, highly intense, using a partial perspective and lasting for a relatively brief period. These features, which express the urgency of the emotional circumstances, make it natural to have a quick and spontaneous information-processing mechanism based upon ready-made responses proven to be beneficial during evolution and personal development. It is also natural that in such urgent and unstable circumstances, where profound personal concerns are at stake, the logical principles that give meaning to various events will be primarily focused upon personal concerns and significant changes.

The connection between the second and third categories—namely, between the information-processing mechanism and the logic of such information processing—is less obvious. Using a deliberative mechanism can also take care of personal concerns and significant changes, which are typical of the logic of emotions. It is more difficult, though probably not impossible, to design a schematic mechanism responding to stable and non-personal characteristics, as those consist of more abstract features that are typically accessible only to intellectual considerations. The first feature referring to the psychological characteristics of the emotional mode provides

a strong reason for connecting the second and third features; there are also some contingent connections between these two features.

Characterizing emotions as a mental mode may change some of the traditional disputes concerning the nature of emotions. The following are a few examples.

*The dichotomy between emotions and thinking.* The intellectual mode often lacks many features typical of the emotional mode. Nevertheless, emotions and thinking are not two opposites, as they refer to different levels of description. Thinking may take place—although in a more limited and less complex manner—in the emotional mode also. Although the intellectual mode is different from the emotional mode, it is inadequate to assume a clear-cut dichotomy between the two, as some features are common to both.

*The essence of emotions.* A central dispute in the study of emotions is whether the essence of emotions is a kind of feeling, a type of appraisal, or action readiness. If an emotion is not a single capacity but a complex mode of the whole mental system, we are less likely to find a single essence of that mode. Hence, it is more adequate to describe emotions by referring to various mental elements.

*Basic and nonbasic emotions.* A heated dispute in the philosophy and psychology of emotions is whether there are certain emotions that can be regarded as basic. Above all, “basic” means simple, as opposed to complex. Accordingly, any emotion is a simple irreducible emotion, or it can be analyzed into a simple emotion plus  $x$ , where  $x$  is either another emotion or some nonemotional element. When emotions are considered as a mode of the mental system, they become more complex and a simple reduction of them to a few basic forms is less plausible. Indeed, criteria for simple or basic emotions vary from one theory to another, and such differences cast doubt on the existence of basic emotions.

Considering an emotion to be a mode of the whole mental system, rather than a mere mental element—such as a disposition, capacity, feeling, or type of intentional reference—is a novel suggestion that provides a more comprehensive and precise description of emotions. Although this description is subject to modifications, I believe that it is basically correct. Above all, this description is able to depict the subtlety of emotions.

## NOTES

1. *Oxford English Dictionary*; see also Ben-Ze'ev 1993, chap. 2.

2. In Wollheim's terms, emotions do not possess subjectivity; see Wollheim 1999, 9.

3. I use the term *mode*, which may be considered archaic, for the lack of a better term. In a sense somewhat similar to the one suggested here, Margaret Donaldson (1993) also proposes that the mind consists of various modes. She distinguishes between four modes, which are defined by their loci of concern: (1) Point mode—locus “here and now”; (2) Line mode—locus “there and then”; (3) Construct mode—locus “somewhere/sometime” (no specific place or time); (4) Transcendent mode—locus “nowhere” (that is, not in space-time). Donaldson further distinguishes four major mental components: perception, action, emotion, and thought. In my view, I do not define the various modes in light of their focus of concern, but in light of their psychological nature. Nevertheless, the first two modes in Donaldson’s view are related to the perceptual mode in my view, and the other two modes in her view to the intellectual mode suggested here. While I believe that perception and thought are mental capacities that are on the same conceptual level, emotions (and probably action as well) are on a different level.

4. For further discussion of this distinction, see, e.g., Ben-Ze’ev 1993, chap. 4, and Smith and Kirby 2000.

5. Due to space limitations, I do not compare the logic of emotion to that of perception; for further discussion on the logic of emotions, see Ben-Ze’ev forthcoming.

## Enough Already with “Theories of the Emotions”

AMÉLIE OKSENBERG RORTY

Notoriously, emotions do not form a natural kind distinguished from motives, moods, and attitudes.<sup>1</sup> To be sure, occurrent emotions like anger, fear, awe, lust, and pride are typically phenomenologically distinguishable from propositional attitudes like astonishment, suspicion, and amusement; but so are many desires (hunger, lust) and moods (depression, anxiety), and many propositional attitudes (astonishment, suspicion, amusement). What have we learned if we are told that (for example) anger, like fear, is an emotion, in contrast to the motivational desires of lust and hunger? Or that nostalgia and depression are moods, while doubt and surprise are propositional attitudes? After all, nostalgia and depression also have propositional content; and anger and fear can, like nostalgia and depression, pervade the psychological field, affecting a person’s perceptions and interpretations. And what about other attitudes? Where does distrust fit? Is greed a motive, an emotion, a character trait? Is joy a mood? A feeling? Is pride a character flaw, a sin, a social construction (whatever that is)? Is benevolence a condition of the will? A feeling? A virtue? Is love a passion, an emotion, a sentiment? And *schadenfreude*? A sense of devotion, dedication, or ambition? Japanese *amae*? Awe, piety? Respect for human rights? What about a sense of vulnerability? Aesthetic delight? Moral indignation? Of course any sensible person would refuse to answer these questions, and both the friends and skeptics of the emotions would rightly argue that the emptiness of the questions—in their bald forms—only shows how subtle, how complex the “emotions” are. The more subtle among the philosophical friends of the emotions recognize that greed, nostalgia, pride, and love designate generic families of quite different attitudes. They are ready to distinguish the pride of self-respect from that of arrogance or

vanity, the nostalgia of the dying from aestheticized sentimental nostalgia, righteous indignation from self-righteous indignation. But once these refinements are given place, what remains of the initial bold distinction between the “emotions” and motives or propositional attitudes? In deference to the richness of folk speech, let’s agree that in speaking of the “emotions” in this chapter, we shall refer to the nonclass of attitudes that also encompasses motives, moods, evaluations, prejudgments, and so on.

In the zeal of theory construction, philosophers and psychologists can of course set these phenomenological refinements aside and define categories as suits them. Emotions have been characterized as a species of evaluative judgments, as erroneous or incomplete partial judgments based on *phantasiai*, as specific sorts of sensations caused by visceral disturbances prompted by a disturbing perception, as exogenous modifications/deflections of endogenous desires, as willful strategies designed to block the realization of the non-being of the self, as perceptions of tertiary qualities, or as indicating what seems important to a person. While staying within the generous and hospitable bounds of the phenomena characterized by folk psychology and folk speech, each of these theories can, despite their dramatic differences, find some plausible purchase for their competing views. The rich variety of words for emotions in contemporary Anglo-American folk speech furthers this hospitality. Their connotations carry distinctive theoretical implications: *passion* suggests fervor; *feeling* connotes sensation, *affect* implies a change, *sentiment* indicates a cognitive attitude, *emotion* (*ex-motu*) suggests a motivational charge. The utility and plausibility of any of these various theories of the emotions—its explanatory power—depends on its integration within a relatively complete theory of mental functioning. The meaning and import—the claims—of the views of Aristotle on *pathe*, Seneca on *ira* and *passio*, Spinoza on *affectus*, Hume on the passions, Rousseau on *sentiment*, Sartre on *emotion* are deeply embedded in their metaphysics and philosophy of mind, on the force of their distinctions between activity and passivity, their theories of the essential or individuating properties of persons. The presumptive category of emotion can play a role in the philosophy of mind only when its function is coordinated or opposed to other psychological activities: sensation, perception, imagination, belief, desire, choice, propositional attitudes. Theories of the emotions whose ambitions stretch to neurophysiology or cognitive science are obliged to extend their integrative scope yet further: their plausibility depends on the explanatory success of their integration within full-blown physiological theories, with their own budgets of distinctions between neurological and hormonal phenomena. It is idle to compare and evaluate these various theories of the emotions as if they were competing accounts of the same phenomena. Presented independently, out of the context of its role in a more encompassing theory, a theory of the emotions can—like its close competitors—at best be shown compatible with current folk psychology, one story among

others, with its particular budget of insight and blindness. But this is a weak requirement; many theories can plausibly claim such compatibility.

To be sure, the explanatory tasks of philosophical psychology can in principle be segmented, parceled out for specialized piecework: theories of emotion, theories of perception, action theory, and so on. Specialization can even go further: experts on envy, regret, or revenge, on pity, guilt, or shame insist on carrying on without serious attention to the historical and sociological contexts of their topics. In any case, piecework is only as good as its successful integration within a functioning whole. I believe this is also true of the segmentation of moral theory: piecework analyses of (as it may be) the virtues, rights, obligations, moral norms, and motivations presuppose a general moral theory. Sometimes, in the face of a glut of particular analyses of (for example) minority rights or cases of lifeboat survival choices, we long for a solid general moral theory. But when we are flush with competing moral theories, we long for good case studies. These two approaches are mutually dependent: the development and defense of general theories initially rests on reflective analyses of specific cases, and the classification and analysis of cases presuppose a set of theoretical categories and principles. Continental philosophers call this kind of mutual dependence “the hermeneutic circle”; the Anglo-American tradition calls it “bootstrapping.” Rawls—referring to Aristotle’s philosophical practice—calls it “the method of reflective equilibrium.” It is the working condition of theory construction.

As long as pieceworkers are under the direction of one plan with shared assumptions and purposes (say, a Thomistic factory or a research team in cognitive science), their distinctive contributions can be evaluated for their ability to merge smoothly into the entire program. Of course specifically focused theories of the emotions can in principle stand in isolation, as in themselves beautiful or startlingly novel museum pieces that, like other works of art, enable acute observers to see patterns in experience they had not previously discerned. A Dogon mask can enable us to find new delight in patterns in the human face, even if we do not have a clue about the place of such masks in Dogon life, or even whether it was even meant to be seen except in (what we call) religious contexts. But we cannot respond to piecemeal theories without placing them in context, as, for instance, within Aristotle’s distinction between *orexis* and *pathos*, of Kant’s analysis of the morality of respect (*Achtung*), Nietzsche on the will to power and *ressentiment*, or Heidegger on *Dasein* and *Angst*. We can, in principle, evaluate piecemeal theories of the emotions, as we do paintings, for the ingenuity and boldness of their construction, the wit of their response to their predecessors and opponents, the way they enlarge, illuminate, or enrich our perceptions. If their claims to *truth* amount to this, and no more than this, fine; but if so, then their competing cousin-theories can also claim that privilege. In that case, polemical discussions among serious com-

petitors are pointless: "Aristotle's account of *pathe* is more subtle than that of the early Stoics" is as out of place as "Piero's gravitas outshines Titian's colors" or "You can safely ignore Fragonard but not Rothko."

Flush with the enthusiasm of theory construction, philosophers and psychologists can nevertheless use two strategies to respond to apparent counterexamples: species qualification and astute gerrymandering.

The *tactic of species qualification* runs like this:

"Emotions are evaluative judgments."

"What sort of evaluative judgments? What about stock market evaluations? The evaluative judgments of realtors and art dealers? Are they emotions?"

"Emotions are a *species* of erroneous or incomplete evaluative judgments."

"So 'Socrates was a vulgar, ugly layabout' is an emotion?"

"Emotion-judgments are a species of incomplete evaluations that are presumptively motivating."

"So motivating desires that embed incomplete evaluations—for instance, 'I want that juicy red apple'—is an emotion?"

"Well, desire-emotions are accompanied by feelings of a certain sort."

And there follows a set of further conditions and qualifications that diminish the bold rhetorical force of the original proclamation. Bit by bit, the whole cohort of psychological attitudes is reintroduced in a theory that was meant to privilege one of them.

It is not a peculiar defect of theories of the emotions that their explanatory power (in contrast to their boldness or wit) depends on their place in an integrated theory of mental functioning. It is equally true of theories of perception or theories of the imagination or of the will that their evaluation—their explanatory interest—lies beyond their being minimally compatible with a solid portion of current folk psychology and folk speech. Considering how nuanced, how hospitable it is in representing the variety of experience, folk psychology can accord basic voting legitimacy to a large quorum of competing theories. As much can be said for a host of philosophic theories that (to our boredom and dismay) treat one another as in direct competition for the Truth of the Matter. Of course not everything goes. "Emotions are the crystallized breath of Satan; beliefs are the soft murmurs of angels" may have limited space as weak poetry; it has no space as a weak theory of the emotions. "Emotions are a species of magic" is an attempt to overcome the distinctions between poetry, rhetoric, and philosophy, and its claim to meaning and plausibility depends on an encompassing philosophy of mind.

The *tactic of gerrymandering* locates the emotions within a larger frame of distinctions between (as it may be) mental states and mental activities, or between cognition and motivation or between perceptions and proprio-

ception, or between physical and psychophysical conditions. Decontextualized theories of the emotions are then evaluated for their relative elegance, simplicity, richness, and completeness in encompassing what are (currently) taken to be the relevant phenomena of the field, as well as for the plausibility of their ingenuity in absorbing or excluding objections and counterexamples. (Of course each of these parameters can vary independently. A brilliantly simple theory can fail to be robust, a perspicuous theory can have few ramifications.) But they are also essentially evaluated for their success in meshing with—and strengthening—theories within the larger domain in which they presumptively function. The relative dominance of each mode of evaluation varies contextually: what is at issue? An NIMH grant for the study of autism, an interdisciplinary ACLS grant for a book on the irony of Dickens's sentimentalism? Yet another dutiful APA session on the emotions? In this, as in other evaluations, there is no bottom line. But it doesn't follow that NIMH review committees are at a loss in distinguishing better and worse theories or grant proposals. In order to evaluate NIMH grant proposals for studies of the effects of war-related trauma on children, review committees need not determine whether (for example) enuresis or attention deficit express emotional or motivational disorders.

For the sake of further inquiry, let's ignore the familiar theory-laden distinctions among psychological attitudes—that is, between beliefs, fantasies, motives, perceptions, decisions, reasons, and so on. Like all other mental phenomena, emotions occur within a system of related attitudes. Independently of any particular role assigned them in this or that philosophical or psychological theory, particular emotions are systematically connected with one another. Without regard to their theoretical classification, emotions always occur and are identified and individuated by reference to other attitudes. Just as every belief presupposes a set of other beliefs, so too every frisson of fear, stab of jealousy, sense of indignation occurs in—and is analyzed by—a set of other supportive and opposed perceptions, beliefs, desires, imaginings. So, for instance, current folk psychology has it that particular instances of fear or jealousy presuppose a sense of vulnerability, that resentment or indignation embeds a sense of pride or fairness, that affection and piety is typically combined with a feeling of gratitude and obligation. What is true in experience is also true in theory: generic classifications of psychological functions/attitudes are defined and identified within a system of mutually implicated categories. Further differentiation of these general attitudes—for example, distinctions between beliefs about matters of fact in contrast to those about values, negative in contrast to positive emotions, motivational attraction in contrast to motivational aversion—are also systematically specified in a more comprehensive frame. And of course generic emotion-types (envy, greed, love, hate, contempt) not only

occur but are also conceptually defined in a system of presuppositions, contrasts, and consequences. Consider the ways in which many philosophers, psychologists, and biologists have mapped the relations among the passions on a Euclidean model, deriving the analysis of complex and compounded passions from those that are (in terms of a comprehensive theory) classified as basic. Although Descartes's *Treatise on the Passions* provides one of the most explicit examples of this sort of analysis, it is implicit in Hobbes's *Leviathan*, Spinoza's *Ethics*, and Hume's *Treatise* as well as in a good deal of contemporary psychology.<sup>2</sup>

Emotions are further identified and individuated by their place in a dramatic narrative. "It is not," Hume remarks, "the present sensation alone or some momentary pain or pleasure which determines the character of any passion, but the whole bent and tendency of it from beginning to end, described by enumeration of the circumstances that attend them" (1973, 385). Proud Achilles became wrathful when his trophy-woman was denied him; trusting Othello became suspicious and then jealous and then enraged after Iago stage-managed apparent scenes of Desdemona's betrayal; newly come to kingship after a devastating civil war, Creon was threatened by Antigone's defiance, and his antagonism mounted to a devastatingly self-destructive rage as her opposition became more pronounced and explicit. For good or ill, emotions and motives are notoriously often erratic and unstable: in the standard scenarios in which they occur, they can sometimes carry the dialectic of their own undoing. Loyalty turns to ambition and envy (Iago, Macbeth); love turns to hate (Medea, Saul); fear turns to love (Jane Eyre); friendship and brotherhood turn to rivalry and sometimes to treachery (Absalom and Amnon, Pandu and Dhritarastra). The question of whether the convoluted narratives of such unstable emotional impulses are built into the very ambivalent nature of devotion, benevolence, and loyalty remains an unresolvable source of anxiety in reflective experience, of subtlety in fiction and of controversy in philosophy.

These are not the stories only of high drama, but also of the melodrama of everyday life. "Jack and Jill adored each another: they seemed the perfect couple. His sense of himself increasingly depended on her admiring support; she was ready to transform her life for his sake. His dependency made him hypersensitive to the tonalities and directions of her attention. When she became absorbed in their newborn child, he lost his bearings and was easily propelled into the arms of Teresa, his welcoming and ambitious boss (administrative assistant, graduate student). Left alone all day to care for their child, Jill lost her enchantment, became chronically desolate, resentful, and eventually enraged, and she . . ." It is no accident that pulp fiction and soap operas are universally gripping, and that their plots do not dramatically differ from those of high literature.<sup>3</sup> Shift the pronouns, add a subplot of political intrigue or religious fervor to the story of

Jack and Jill, express it in superb imagery and language, and you have the fundamental plot of immortal tragedy.

The long history of the distrust of the emotions begins with these narratives of dramatic instability and ambivalence; they tend toward melodrama. The critics of “the emotions” charge that they are reactions to contingent events over which we have little control; that they are attitudes whose impetuous strength and intensity are more likely than not to be irrational and so to endanger well-being. For good and ill, they are primarily prompted by—and respond to—present circumstances, rarely checked by long-range “all-things considerations.” Because emotions guide what is perceived as salient, because they provide parameters of interpretation, they are difficult to correct.

It is no accident that motivating emotions are regarded with suspicion, to be repressed or controlled, at best to be managed, brought under the domain of virtue. Even when they are endorsed as necessary supplements to high devotion or nobility, the dangers of their moving to their own excess remains. Ardent benevolence runs the risk of crippling its recipients; parental devotion may become possessive; patriotic fervor readily moves to bigotry. Yet for all of their excesses and instability, benevolence, love, loyalty remain. With a neat gerrymandering twist, Stoics categorize reliably constructive attitudes as *eupatheiai* instead of *pathe*, Descartes treats them as *habitudes* instead of *passions*, Spinoza analyzes them as active instead of passive affects. Benevolence that debilitates will be disqualified as benevolence, possessive patriotism that prohibits criticism will be seen as dangerous fanaticism; love and devotion that cripple will be suspect as love. The cognitive/intentional components of benign attitudes (*eupatheiai*, *habitudes*, active affects) carry their own normative regulators: benevolence and generosity that do not debilitate their recipients are classified as virtues rather than emotions. But when the cognitive components of these attitudes are not normatively screened, they are classified as emotions, presumed to be erratic, and their exercise unreliable. Of course this gerrymandering preserves the neatness of a theory and the subtlety of folk speech, but it is not much help to well-intentioned citizens and lovers. But then no one, not even the Stoics, thought that philosophic analysis could by itself, without the support of sound political, social, and economic arrangements, succeed in structuring a well-lived life.

The history of the distrust of the emotions is nevertheless matched by a history of insistence on their essential roles in engendering or expressing morality. The Platonic tradition—as expressed by Augustine, Dante, Spenser, Iris Murdoch, C. S. Lewis—makes a variety of love the pivotal point of the transformation of the fallen condition of self-absorption. In a less mystical vein, eighteenth-century moral theorists—Shaftesbury, Hume, Adam Smith—made the education of the sentiments the basis of the sense—and the practices—of justice. More recently, post- and anti-Kantian philosophers

(Annette Baier, Alan Gibbard) have argued that moral attitudes express emotions; yet others analyze the moral weight of various emotions (shame, guilt, loyalty, friendship, respect).

Why does the map—the relative dominance and recessiveness—of psychological attitudes change? Why—and when—do the emotions receive bad press? When are their benign forms seen as humane and humanizing, a condition of morality? Why does the focus of psychological/philosophic attention shift, say from pride and shame to piety and guilt, or from patriotism and loyalty to ambition and greed? Certainly some “basic” emotions—fear, love, hate, anger, startle/wonder—are presumptively universal, fundamental survival equipment for a vulnerable species whose existence depends on maintaining a fragile balance of hierarchical power, mutual dependence, and trust/distrust. Given (what we take to be) our natural situation, it is no surprise that the psychological and logical relations among anger, loyalty, pride-ranking, love and hate, perhaps also the obligations of piety and gratitude, are always somewhere near the fore of our experience and attention, the subjects of our most searching and enduring epics, dramas, and novels. But the interpretations—the presumptive causes, tonalities, and expressions of the “basic” attitudes—do not remain constant. Beyond those basic attitudes, the focus of phenomenological and analytic attention shifts with changes in primary social, political, and economic activities. Indeed the prime examples of emotions change: the list expands and contracts. For from being a pathos, Aristotelian wonder (*thauma*) is a fundamental activity of the mind; for Descartes *admiratio* is the first of the passions; for Augustine and early Christian theologians, pride is a sin; for Hume it is a key to a person’s motivational structure, revealing what he prizes as central to his continued identity. For Rousseau amour propre is a mark of social corruption, the dependence of self-regard on social standing.

Journalistic sociological banalities are sometimes illuminating: the merits and dangers of pride, ambition, and “the achievement motive” are central psychological topics when political and economic development are at the forefront of attention (Machiavelli, Mandeville, Smith); “benevolence” and “compassion” come to the fore when “natural” bonding or altruistic motives seem eclipsed (Butler and post-Hobbesian ethical theory); anxiety and *angst* emerge in times of stress and anomie (Kierkegaard and Heidegger); free floating rage is a matter of concern during periods of traumatic social and political unrest (Seneca, eras of religious and ethnic/national strife).

A person’s—a society’s—characteristic emotional repertoire, its pattern of dominant and recessive emotions and attitudes, is structured by, and in turn reinforces, political and economic arrangements.<sup>4</sup> While officially condemning envy as a socially undesirable emotion, societies use and even

induce envious traits to encourage the development of what they regard as useful talents and abilities. Market-based consumer-oriented economic systems generate invidious comparisons as a way of increasing consumption. The mass media, television dramas, songs, and advertisements present riveting and reigning fantasies and models of desirability and success. They are brilliantly designed to affect patterns of consumption, to present images of satisfied desire, reinforcing the range of emotion-based motives that sustain the very social and political arrangements that structure them.

Workplaces, religious institutions, banks, schools, courts, museums, mass media, armies, advertising and insurance companies, clinics and hospitals all model, direct, and constrain the psychological/emotional repertoire of a culture. They frame what is to be feared and locate the terms and patterns of a sense of security. They valorize some attitudes (loyalty, gratitude, kindness, respect, ambition); stigmatize others (hate, contempt, jealousy); generate ambivalence for yet others (greed, charity, rivalry). They set the standards for excess and deficiency; they structure the narratives of love and devotion, the boundaries of self-regard and altruism, of parental and filial affections. Emotional nonconformists are regarded with suspicion and charged with irrationality; they have difficulty eliciting cooperation and they suffer the pressure that is intended to produce guilt, or at the very least, shame. Social institutions provide the models for the feelings of responsibility and accountability; they set norms for the tenor of social interactions, finely attuned for status and power, formality or intimacy, empathically tactful or aggressively confrontational. They form the patterns and the habits of aggression and cooperation that are exercised in generating and resolving ordinary conflicts.

A culture whose institutions conflict with one another—banks and workplaces endorsing one set of attitudes, while the arts and religious institutions affirm and promote a contrary set—is likely to be divided; and its members are likely to suffer *akrasia* and other sorts of psychological conflicts. Social and psychological compartmentalization can sometimes diminish the appearance of such conflicts. It does so all the more effectively and securely when socioeconomic institutions separate “distinct” domains: work, family, recreation, citizenship. The habits and mentality of impersonal cost accounting that govern many occupations are cordoned off from those that govern affectional relations. Ironically, sometimes a person’s attempt to integrate his repertoire of psychological attitudes—to import cost accounting into friendships, for instance—may supply the occasion for further conflicts. It can also go the other way: a judge who attempts to import the psychological attitudes of her personal life into the courtroom can violate the principles she thinks ought to govern her judicial decisions. The persona that strives for emotional integration sometimes adds yet another voice to the cacophony that is the endeavoring self.

In short: theories of the “emotions” (1) do not “cut at the joints”: their subject matter encompasses a heterogeneous set of attitudes, not sharply distinguished from motives, moods, propositional attitudes; (2) are comprehensible only within the larger frame of a relatively complete philosophy of mind/philosophical psychology.

Specific “emotional” attitudes are individuated and identified (1) within a nexus of supportive and opposed attitudes that are characteristically (2) within the context of a narrative scenario. (3) A culture’s repertoire of “emotions” is structured by its economic, political, and social arrangements.

#### NOTES

1. See the introduction to Rorty 1980.

2. For contemporary taxonomies of the emotions, see Ekman and Sherer 1984; Ekman and Davidson 1994; Ortony, Clore, and Collins 1988; Izard 1992; and Panskepp 1998.

3. For a gripping account of the transformations of the basic plots of narrative development, see Frye 1957.

4. See Rorty 1998 and 1997. I have imported some passages from these essays to this chapter.

## BIBLIOGRAPHY

- Alanen, L. Forthcoming. "What Are Emotions About?" *Philosophy and Phenomenological Research* (Sept. 2003).
- Alston, W. P. 1967. "Emotion and Feeling." In *The Encyclopaedia of Philosophy*, ed. P. Edwards. Vol. 2. New York: Macmillan.
- Aristotle. 1985. *Nicomachean Ethics*. Trans. T. Irwin. Indianapolis, IN: Hackett.
- Aristotle. 1991. *Rhetoric*. Trans. George A. Kennedy. New York: Oxford University Press.
- Armstrong, D. M. 1968. *A Materialist Theory of Mind*. London: Routledge.
- Arnold, M. B. "Human Emotion and Action." In *Human Action: Conceptual and Empirical Issues*, ed. T. Mischel. New York: Academic Press, 1969.
- Arrow, K. 1971. "Political and Economic Evaluation of Social Effects and Externalities." In *Frontiers of Quantitative Economics*, ed. M. Intriligator. Amsterdam: North-Holland.
- Austin, J. L. 1946. "Other Minds." *Proceedings of the Aristotelian Society*, supp. vol. 20, reprinted in *Philosophical Papers*, 3d ed. Oxford: Oxford University Press, 1979.
- . 1961. "Agathon and Eudaimonia in the Ethics of Aristotle." In *Philosophical Papers*, 3d ed. Oxford: Oxford University Press, 1979.
- Averill, J. 1982. *Anger and Aggression: An Essay on Emotion*. New York: Springer.
- Ayer, A. J. 1952. *Language, Truth, and Logic*. New York: Dover.
- Baier, A. 1986. "Extending the Limits of Moral Theory." *Journal of Philosophy* 83: 538–45.
- . 1990. "What Emotions Are About." *Philosophical Perspectives*, vol. 4: *Theory and Philosophy of Mind*, ed. J. E. Tomberlin. Atascadero, CA: Ridgeview.
- Becker, G. 1996. *Accounting for Tastes*. Cambridge: Harvard University Press.
- Becker, L. C., and C. B. Becker, eds. 2001. *The Encyclopedia of Ethics*. London: Routledge.

- Bedford, E. "Emotions." *Proceedings of the Aristotelian Society*, 1957; reprinted in *Essays in Philosophical Psychology*, ed. D. Gustafson. New York: Doubleday-Anchor, 1963.
- Ben-Ze'ev, A. 1993. *The Perceptual System: A Philosophical and Psychological Perspective*. New York: Peter Lang.
- . 2000. *The Subtlety of Emotions*. Cambridge: MIT Press.
- . 2002. Emotions Are Not Feelings. *Consciousness and Emotion* 3: 81–89.
- . 2003a. Emotions Are Not Mere Judgments. *Philosophical and Phenomenological Research* (forthcoming).
- . 2003b. "The Logic of Emotions." In *Philosophy and the Emotions*, ed. A. Hatzimoysis. Cambridge: Cambridge University Press.
- Bhagavadgītā in the Mahābhārata: Text and Translation*. 1985. Trans. and ed. J.A.B. van Buitenen. Chicago: Chicago University Press.
- Bilimoria, P. 2002. "A Postcolonial Critique of Reason: Spivak between Kant and Matilal." *Interventions Journal of Postcolonial Studies* 42.2: 160–67.
- Bilimoria, P., and J. N. Mohanty, eds. 1997. *Relativism, Suffering, and Beyond: Essays in Honour of Bimal K Matilal*. Delhi: Oxford University Press.
- Blackburn, S. 1998. *Ruling Passions: A Theory of Practical Reasoning*. Oxford: Oxford University Press.
- Bollas, C. 1987. *The Shadow of the Object: Psychoanalysis of the Unthought Known*. London: Tavistock.
- Bovens, L. 1999. "The Two Faces of Akratics Anonymous." *Analysis* 59.4: 230–36.
- Bowlby, J. 1973. *Separation: Anxiety and Anger*. New York: Basic.
- . 1980. *Loss: Sadness and Depression*. New York: Basic.
- . 1982. *Attachment*. New York: Basic.
- Bratman, M. 2001. "Two Problems about Agency." *Proceedings of the Aristotelian Society* 101.3: 309–26.
- Brewer, B. 1995. "Bodily Awareness and the Self." In *The Body and the Self*, ed. J. L. Bermúdez, A. Marcel, and N. Eilan. Cambridge: MIT Press, 1995.
- . 2002. "Emotions and Other Minds." In *Understanding Emotions: Mind and Morals*, ed. P. Goldie. Aldershot, U.K.: Ashgate.
- Briggs, J. 1970. *Never in Anger*. Cambridge: Harvard University Press.
- Budd, M. 1985. *Music and the Emotions*. London: Routledge and Kegan Paul.
- Buss, D. M. 2000. *The Dangerous Passion: Why Jealousy Is as Essential as Love and Sex*. New York: Simon and Schuster.
- Calder, A. J., A. D. Lawrence, and A. W. Young. 2001. "Neuropsychology of Fear and Loathing." *Nature Reviews Neuroscience* 2: 352–63.
- Calhoun, C. 1984. "Cognitive Emotions?" In *What Is an Emotion?*, ed. C. Calhoun and R. C. Solomon. New York: Oxford University Press.
- Campbell, S. 1998. *Interpreting the Personal: Expression and the Formation of Feeling*. Ithaca, NY: Cornell University Press.
- Cannon, W. B. 1927. "The James-Lange Theory of Emotion: A Critical Examination and an Alternative Theory." *American Journal of Psychology* 39: 106–24.
- . 1929. *Bodily Changes in Pain, Hunger, Fear, and Rage*. New York: Harper and Row.
- Cassirer, E. 1923. *Substance and Function*. New York: Dover, 1953.
- Chalmers, D. 1996. *The Conscious Mind*. Oxford: Oxford University Press.

- Chan, A. K. L. 1998. "Ethics of Emotion." *Philosophy East and West* 48.1: 176–87.
- Charland, L. C. 1997. "Reconciling Cognitive and Perceptual Theories of Emotion: A Representational Proposal." *Philosophy of Science* 64.4: 555–79.
- . 2001. "In Defence of 'Emotion.'" *Canadian Journal of Philosophy* 31.1: 133–54.
- . 2002. "The Natural Kind Status of Emotion." *British Journal for the Philosophy of Science* 53.4 (Dec.): 511–37.
- Chodorow, N. 1980. *The Reproduction of Mothering*. Berkeley: University of California Press.
- Churchland, P. M., and P. S. Churchland. 1998. "Recent Work on Consciousness: Philosophical, Theoretical, and Empirical." In *On the Contrary: Critical Essays, 1987–87*. Cambridge: MIT Press.
- Churchland, P. S. 1986. "Consciousness: The Transmutation of a Concept." *Pacific Philosophical Quarterly* 64: 80–93.
- Chwalisz, K., E. Diener, and D. Gallagher. 1988. "Autonomic Arousal Feedback and Emotional Experience: Evidence from the Spinal Cord Injured." *Journal of Personality and Social Psychology* 54: 820–28.
- Coleman, J. 1990. *Foundations of Social Theory*. Cambridge: Harvard University Press.
- Cooper, J. 1980. "Aristotle on Friendship." In *Essays on Aristotle's Ethics*, ed. A. Rorty. Berkeley: University of California Press.
- Cosmides, L., and J. Tooby. 2000. "Evolutionary Psychology and the Emotions." In *Handbook of the Emotions*, ed. M. Lewis and J. M. Haviland-Jones. 2d ed. New York: Guilford.
- Crane, T. 1998. "Intentionality as the Mark of the Mental." In *Current Issues in Philosophy of Mind*, ed. A. O'Hear. Oxford: Oxford University Press.
- Crane, T. 2001. *Elements of Mind*. Oxford: Oxford University Press.
- Critchley, H. D., C. J. Mathias, and R. J. Dolan. 2001. "Neural Correlates of First and Second-Order Representation of Bodily States." *Nature Neuroscience* 4: 207–12.
- D'Arms, J., and D. Jacobson. 2000. "The Moralistic Fallacy: On the 'Appropriateness' of Emotion." *Philosophy and Phenomenological Research* 61: 65–90.
- Damasio, A. R. 1994. *Descartes' Error: Emotion, Reason, and the Human Brain*. New York: G. P. Putnam's Sons.
- . 1999. *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. New York: Harcourt Brace.
- Damasio, A. R., T. J. Grabowski, A. Bechara, H. Damasio, L.L.B. Ponto, J. Parvizi, and R. D. Hichwa. 2000. "Subcortical and Cortical Brain Activity during the Feeling of Self-Generated Emotions." *Nature Neuroscience* 3: 1049–56.
- Darwin, C. 1871. *The Descent of Man and Selection in Relation to Sex*. New York: Modern Library.
- . 1998. *The Expression of the Emotions in Man and Animals*. Intro., afterword, and commentary by Paul Ekman. 3d ed. New York: Oxford University Press.
- Davidson, D. 1970. "How Is Weakness of the Will Possible?" In *Essays on Actions and Events*. Oxford: Oxford University Press, 1980.
- . 1976. "Hume's Cognitive Theory of Pride." In *Essays on Actions and Events*. Oxford: Oxford University Press, 1980.

- Delancey, C. 2001. *Passionate Engines: What Emotions Reveal about Mind and Artificial Intelligence*. New York: Oxford University Press.
- de Silva, P. 1984. "The Ethics of Moral Indignation and the Logic of Violence." V. F. Gunaratne Memorial Lecture, Public Trustee Department, Colombo, Sri Lanka.
- . 1995. "Theoretical Perspectives on Emotions in Early Buddhism." In *Emotions in Asian Thought*, ed. J. Marks and R. Ames. Albany: State University of New York Press.
- de Sousa, R. 1974. "The Good and the True." *Mind* 83: 534–51.
- . 1980. "The Rationality of Emotions." In *Explaining Emotions*, ed. A. Rorty. Berkeley: University of California Press.
- . 1987. *The Rationality of Emotion*. Cambridge: MIT Press.
- . 1990. "Emotions, Education, and Time." *Metaphilosophy* 21: 434–46.
- . 2000. "Learning to Be Natural." In *Being Humans*, ed. Neil Roughley. Berlin: Walter de Gruyter.
- . 2002a. "Twelve Varieties of Subjectivity." In *Proceedings of the Sixth International Colloquium on Cognitive Science*, ed. Jesus M. Larrazabal. Dordrecht: Kluwer.
- . 2002b. "Emotional Truth." *Proceedings of the Aristotelian Society*, supp. vol. 76.1: 247–63.
- de Waal, F. 1996. *Good Natured: The Origins of Right and Wrong in Humans and Other Animals*. Cambridge: Harvard University Press.
- Deigh, J. 1994. "Cognitivism in the Theory of Emotions." *Ethics* 104: 824–54.
- . 2000. "Nussbaum's Defense of the Stoic Theory of Emotions." *Quinnipiac Law Review* 19: 293–307.
- . 2001. "Emotions: The Legacy of James and Freud." *International Journal of Psychoanalysis* 82: 1247–56.
- Descartes, R. [1649] 1989. *The Passions of the Soul*. Indianapolis, IN: Hackett.
- . 1967. *Meditations on First Philosophy*. In *The Philosophical Works of Descartes*. Trans. E. Haldane and G.R.T. Ross. Cambridge: Cambridge University Press.
- De Smet, R. 1977. "A Copernican Reversal: The Gītā's Reformation of Karma." *Philosophy East and West* 27: 53–63.
- Dewey, J. 1984. "The Theory of Emotion." In *What Is an Emotion?*, ed. C. Calhoun and R. C. Solomon. New York: Oxford University Press.
- Donaldson, M. 1993. *Human Minds: An Exploration*. London: Penguin.
- Doris, J. M. 2001. Review of P. E. Griffiths, *What Emotions Really Are: The Problem of Psychological Categories*. *Ethics* 110.3 (Apr.): 617–19.
- Downing, G. 2001. "Emotion Theory Revisited." In *Heidegger, Coping, and Cognitive Science: A Festschrift for Hubert Dreyfus*, vol. 2. Cambridge: MIT Press.
- Dretske, F. 1981. *Knowledge and the Flow of Information*. Cambridge: MIT Press.
- . 1988. *Explaining Behavior*. Cambridge: MIT Press.
- Drori, I. 2002. "Palestinian Suicide Missions in Israel: A Briefing Paper." Prepared for All Souls Suicide Missions Workshops, 2002.
- Ekman, P. 1972. *Emotions in the Human Face*. New York: Pergamon.
- . 1973. *Darwin and Facial Expression*. New York: Academic Press.
- . 1980. "Biological and Cultural Contributions to Body and Facial Move-

- ment in the Expression of Emotion." In *Explaining Emotions*, ed. A. Rorty. Berkeley: University of California Press.
- . 1984. "Expression and the Nature of Emotion." In *Approaches to Emotion*, ed. K. Scherer and P. Ekman. Hillsdale, NJ: Erlbaum.
- . 1994. "Strong Evidence for Universals in Facial Expression: A Reply to Russell's Mistaken Critique." *Psychological Bulletin* 115: 268–87.
- . 1999. "Basic Emotions." In *The Handbook of Cognition and Emotion*, ed. T. Dalgleish and T. Power. New York: John Wiley and Sons.
- Ekman, P., and R. Davidson, eds. 1994. *The Nature of Emotion*. New York: Oxford University Press.
- Ekman, P., and W. V. Friesen. 1971. "Constants across Cultures in the Face and Emotion." *Journal of Personality and Social Psychology* 17: 124–29.
- Ekman, P., W. V. Friesen, and R. C. Simons. 1985. "Is the Startle Reaction an Emotion?" *Journal of Personality and Social Psychology* 49:5: 1416–26.
- Ekman, P., and K. Scherer, eds. 1984. *Approaches to Emotion*. Hillsdale, NJ: Erlbaum.
- Ellsworth, P. C. 1994. "William James and Emotion: Is a Century of Fame Worth a Century of Misunderstanding?" *Psychological Review* 101: 222–29.
- Elster, J. 2000. *Alchemies of the Mind*. Cambridge: Cambridge University Press.
- Elster, J., ed. 1999. *Addiction: Entries and Exits*. New York: Russell Sage.
- Emery, N. J., and D. G. Amaral. 2000. "The Role of the Amygdala in Primate Social Cognition." In *The Cognitive Neuroscience of Emotion*, ed. R. Lane and L. Nadel. New York: Oxford University Press.
- Esteves, F., and A. Öhman. 1993. "Masking the Face: Recognition of Emotional Facial Expression as a Function of the Parameters of Backward Masking." *Scandinavian Journal of Psychology* 34: 1–18.
- Fairbairn, W.R.D. 1952. *Psychoanalytic Studies of the Personality*. London: Tavistock.
- Farrell, D. 1980. "Jealousy." *Philosophical Review* 89: 527–59.
- Fodor, J. A. 1990. "A Theory of Content, I and II." In *A Theory of Content and Other Essays*. Cambridge: MIT Press.
- Forgas J., ed. 2000. *Feeling and Thinking: The Role of Affect in Social Cognition*. Cambridge: Cambridge University Press.
- Frank, R. H. 1988. *Passions within Reason: The Strategic Role of the Emotions*. New York: W. W. Norton, 1988.
- Frankfurt, Harry. 1976. "Identification and Externality." In *The Identities of Persons*, ed. Amélie O. Rorty. Berkeley: University of California Press.
- . 1979. "The Problem of Action." *American Philosophical Quarterly*.
- . 1988a. "The Importance of What We Care About." In *The Importance of What We Care About: Philosophical Essays*. New York: Cambridge University Press.
- . 1988b. "Rationality and the Unthinkable." In *The Importance of What We Care About: Philosophical Essays*. New York: Cambridge University Press.
- . 1999. "On the Usefulness of Final Ends." In *Necessity, Volition, and Love*. New York: Cambridge University Press.
- Freud, S. 1915. *Project for a Scientific Psychology*. 1895. In *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, trans. and ed. James Strachey, vol. 3. London: Hogarth.
- Frijda, N. 1986. *The Emotions*. Cambridge: Cambridge University Press.

- Frye, N. 1957. *The Anatomy of Fiction*. Princeton, NJ: Princeton University Press.
- Gallagher, S. 1995. "Body Schema and Intentionality." In *The Body and the Self*, ed. J. L. Bermúdez, A. Marcel, and N. Eilan. Cambridge: MIT Press.
- Gaulin, S.J.C., and D. H. McBurney. 2001. *Psychology: An Evolutionary Approach*. Upper Saddle River, NJ: Prentice Hall.
- Goldie, P. 2000. *The Emotions: A Philosophical Exploration*. Oxford: Clarendon.
- . 2002. "Emotions, Feelings, and Intentionality." *Phenomenology and the Cognitive Sciences* 1: 235–54.
- . Forthcoming. "Emotion, Reason, and Virtue." In *Emotion, Evolution, and Rationality*, ed. P. Cruse and D. Evans. Oxford: Oxford University Press.
- Goldman, R. "Esa Dharma Sanātānah: Shifting Moral Values and the Indian Epics." In *Relativism, Suffering, and Beyond: Essays in Honour of Bimal K Matilal*, ed. P Bilimoria and J. N. Mohanty. Delhi: Oxford University Press, 1997.
- Goodman, N. 1954. *Fact, Fiction, and Forecast*. London: Athlone Press.
- Gordon, R. M. 1986. "The Passivity of Emotions." *Philosophical Review* 95: 339–60.
- . 1987. *The Structure of Emotions: Investigations in Cognitive Philosophy*. Cambridge: Cambridge University Press.
- Gould, S. J. 2002. *The Structure of Evolutionary Theory*. Cambridge: Harvard University Press.
- Gould, S. J., & Vrba, E. 1982. "Exaptation: A Missing Term in the Science of Form." *Paleobiology* 8: 4–15.
- Green, O. H. 1972. "Emotions and Belief." *American Philosophical Quarterly, Monograph* 6. Oxford: Blackwell.
- Greenspan, P. 1980. "A Case of Mixed Feelings: Ambivalence and the Logic of Emotions." In *Explaining Emotions*, ed. A. Rorty. Berkeley: University of California Press.
- . 1988. *Emotions and Reasons: An Inquiry into Emotional Justification*. New York: Routledge, Chapman and Hall.
- . 1995. *Practical Guilt: Moral Dilemmas, Emotions, and Social Norms*. New York: Oxford University Press.
- . 2000. "Emotional Strategies and Rationality." *Ethics* 110: 469–87.
- . 2004. "Practical Reasoning and Emotions." In *The Handbook of Rationality*, ed. A. Mele and P. Rawlings. New York: Oxford University Press.
- Greenspan, S., with B. L. Benderly. 1997. *The Growth of the Mind and the Endangered Origins of Intelligence*. Reading, MA: Perseus.
- Griffiths, P. E. 1994. "Cladistic Classification and Functional Explanation." *Philosophy of Science* 61.2: 206–27.
- . 1996a. "Darwinism, Process Structuralism, and Natural Kinds." *Philosophy of Science* 63 (Supplement: PSA 1996 contributed papers), S1–S9.
- . 1996b. "The Historical Turn in the Study of Adaptation." *British Journal for the Philosophy of Science* 47.4: 511–32.
- . 1997. *What Emotions Really Are: The Problem of Psychological Categories*. Chicago: University of Chicago Press.
- . 1999. "Squaring the Circle: Natural Kinds with Historical Essences." In *Species: New Interdisciplinary Essays*, ed. R. A. Wilson. Cambridge: MIT Press.

- . 2001a. "Emotion and the Problem of Psychological Categories." In *Emotions, Qualia, and Consciousness*, ed. A. W. Kazniak. Singapore: World Scientific.
- . 2001b. "From Adaptive Heuristic to Phylogenetic Perspective: Some Lessons from the Evolutionary Psychology of Emotion." In *The Evolution of Minds: Psychological and Philosophical Perspectives*, ed. H.H.I. Holcolmb. Dordrecht: Kluwer.
- . 2003. "Basic Emotions, Complex Emotions, Machiavellian Emotions." In *Philosophy and the Emotions*, ed. A. Hatzimoysis. Royal Institute of Philosophy no. 52. New York: Cambridge University Press.
- . Forthcoming a. "Emotions as Natural Kinds and Normative Kinds." *Philosophy of Science*, supplement, *Proceedings of the 2002 Biennial Meeting of the Philosophy of Science Association*. Part II: Symposia Papers.
- . Forthcoming b. "Towards a Machiavellian Theory of Emotional Appraisal." In *Emotion, Evolution and Rationality*, ed. P. Cruise and D. Papineau. Oxford: Oxford University Press.
- Guntrip, H. 1969. *Schizoid Phenomena, Object Relations, and the Self*. New York: International Universities Press.
- Gupta, Anil. 1998. "Tarski's Theory of Truth." *Routledge Encyclopedia of Philosophy*, ed. E. Craig. 9:265–69. London: Routledge.
- Hardin, C. 2000. "Red and Yellow, Green and Blue, Warm and Cool: Explaining Colour Appearance." *Journal of Consciousness Studies* 6.6–7: 113–22.
- Harlow, C. M., ed. 1986. *From Learning to Love: The Selected Papers of H. F. Harlow*. New York: Praeger.
- Harré, R. 1986. *The Social Construction Theory of Emotions*, ed. R. Harré. Oxford: Blackwell.
- Heelas, P. 1986. "Emotion Talk across Cultures." In *The Social Construction Theory of Emotions*, ed. R. Harré. Oxford: Blackwell.
- Helm, B. 2001. *Emotional Reason: Deliberation, Motivation, and the Nature of Value*. New York: Cambridge University Press.
- Hill, P. 2002. "Kamikaze Briefing Paper." Prepared for All Souls Suicide Missions Workshops, 2002.
- Hinde, R. A. 1985a. "Expression and Negotiation." In *The Development of Expressive Behavior*, ed. G. Zivin. New York: Academic Press.
- . 1985b. "Was 'The Expression of Emotions' a Misleading Phrase?" *Animal Behaviour* 33: 985–92.
- Hobbes, T. 1994. *Leviathan*. Indianapolis, IN: Hackett.
- Hoffman, D. D. 1998. *Visual Intelligence: How We Create What We See*. New York: Norton.
- Hohmann, G. W. 1966. "Some Effects of Spinal Cord Lesions on Experienced Emotional Feelings." *Psychophysiology* 3: 143–56.
- Hookway, C. 1998. "Doubt: Affective States and the Regulation of Inquiry." *Canadian Journal of Philosophy*, supp. vol. 24, ed. C. Misak, 203–26.
- Hume, D. 1973. *A Treatise of Human Nature*, ed. L. A. Selby-Bigge. Oxford: Clarendon.
- Hutto, D. 2002. "The World Is Not Enough: Shared Emotions and Other Minds." In *Understanding Emotions: Mind and Morals*, ed. P. Goldie. Aldershot, U.K.: Ashgate.

- Izard, C. E. 1971. *The Face of Emotion*. New York: Appleton-Century-Crofts.
- . 1977. *Human Emotions*. New York: Plenum.
- . 1989. *Encyclopedia Britannica*, s.v. "Emotions."
- . 1992. "Basic Emotions, Relations among Emotions, and Emotion-Cognition Processes." *Psychological Review* 99.3: 561–65.
- Jackson, F. 1982. "Epiphenomenal Qualia." *Philosophical Quarterly* 32: 127–36.
- Jaggard, A. M. 1989. "Love and Knowledge: Emotion as an Epistemic Resource for Feminists." In *Gender/Body/Knowledge: Feminist Reconstructions of Being and Knowing*, ed. Alison M. Jaggard and Susan R. Bordo. Rutgers University Press.
- James, W. 1884. "What Is an Emotion?" Reprinted in *What Is an Emotion?*, ed. C. Calhoun and R. C. Solomon. New York: Oxford University Press.
- . [1890] 1950. *The Principles of Psychology*. New York: Dover.
- . 1894. "The Physical Basis of Emotion." *Psychological Review* 1: 516–29.
- Keller, E. F. *Reflections on Gender and Science*. New Haven, CT: Yale University Press.
- Kenny, A. 1963. *Action, Emotion, and Will*. London: Routledge and Kegan Paul.
- Kinsbourne, M., and R. Hicks. 1978. "Functional Cerebral Space: A Model for Overflow, Transfer, and Interference Effects in Human Performance. A Tutorial Review." In *Attention and Performance VII*, ed. J. Requin, 345–62. Hillsdale, NJ: Erlbaum.
- Klein, M. 1984. *Love, Guilt, and Reparation and Other Works, 1921–1945*. London: Hogarth.
- . 1985. *Envy and Gratitude and Other Works, 1946–1963*. London: Hogarth.
- Krystal, H. 1959. *Integration and Self-Healing: Affect, Trauma, and Alexithymia*. New York: Basic Books.
- Kunst-Wilson, W. R., and R. Zajonc. 1980. "Affective Discrimination of Stimuli Cannot Be Ignored." *Science* 207: 557–58.
- Lamb, S., and J. Murphy, eds. 2002. *Before Forgiving: Cautionary Views of Forgiveness in Psychotherapy*. New York: Oxford University Press.
- Lane, R. 1999. *The Cognitive Neuroscience of Emotion*. New York: Oxford University Press.
- Lange, C. G. 1885. *Om Sindsbevaegelser: Et Psyko-fysiologisk Studie*. Copenhagen: Jacob Lunds. Reprinted in *The Emotions*, ed. C. G. Lange and W. James, trans. I. A. Haupt. Baltimore: Williams and Wilkins Company, 1922.
- Lazarus, R. S. 1991. *Emotion and Adaptation*. New York: Oxford University Press.
- . 1994. "Appraisal: The Long and the Short of It." In *The Nature of Emotion*, ed. P. Ekman and R. Davidson. New York: Oxford University Press.
- . 1999. "The Cognition-Emotion Debate: A Bit of History." In *The Handbook of Emotion and Cognition*, ed. T. Dalgleish and M. J. Power. Chichester, NY: John Wiley and Sons.
- Lazarus, R. S., J. Averill, and E. Opton. 1970. "Towards a Cognitive Theory of Emotion." In *Feelings and Emotions*, ed. M. B. Arnold. New York: Academic Press.
- Lazarus, R. S., J. C. Coyne, and S. Folkman. 1984. "Cognition, Emotion, and Motivation: Doctoring Humpty Dumpty." In *Approaches to Emotion*, ed. K. Scherer and P. Ekman. Hillsdale, NJ: Erlbaum.
- Lear, J. 1990. *Love and Its Place in Nature*. New Haven, CT: Yale University Press.

- LeDoux, J. E. 1993. "Emotional Networks in the Brain." In *Handbook of Emotions*, ed. M. Lewis and J. M. Haviland. New York: Guilford.
- . 1996. *The Emotional Brain: The Mysterious Underpinnings of Emotional Life*. New York: Simon and Schuster.
- Levenson, R. W. 1994. "The Search for Autonomic Specificity." In *The Nature of Emotions*, ed. Paul Ekman and Richard J. Davidson, 252–57. New York: Oxford University Press.
- Levenson, R. W., P. Ekman, and W. V. Friesen. 1990. "Voluntary Facial Action Generates Emotion-Specific Autonomic Nervous System Activity." *Psychophysiology* 27: 363–84.
- Levy, R. 1973. *The Tahitians*. Chicago: University of Chicago Press.
- Lloyd, G. 1985. *The Man of Reason: "Male" and "Female" in Western Philosophy*. Minneapolis: University of Minnesota Press.
- Locke, J. 1975. *An Essay Concerning Human Understanding*, ed. P. Nidditch. Oxford: Clarendon.
- Loewenstein, G. 1996. "Out of Control: Visceral Influences on Behavior." *Organizational Behavior and Human Decision Processes* 65: 272–92.
- . 1998. "A Visceral Theory of Addiction." In *Getting Hooked: Rationality and the Addictions*, ed. J. Elster and O.-J. Skog. Cambridge: Cambridge University Press.
- Lovejoy, A. O. 1961. *Reflections on Human Nature*. Baltimore: Johns Hopkins University Press.
- Lutz, C. 1988. *Unnatural Emotions: Everyday Sentiments on a Micronesian Atoll and Their Challenge to Western Theory*. Chicago: University of Chicago Press.
- Lyons, W. 1980. *Emotion*. Oxford: Oxford University Press, 1980.
- . 1996. *Modern Philosophy of Mind*. London: Penguin, 1996.
- MacLean, P. D. 1960. "New Findings Relevant to the Evolution of Psychosexual Functions of the Brain." *Journal of Nervous and Mental Diseases* 135: 289–301.
- Mahābhārata*. 1933–71. Critical edition. 19 vols. Poona: Oriental Research Institute.
- Mahābhārata*. 1972–80. Trans. J.A.B. van Buitenen. Books I–IV, 3 vols. Chicago: Chicago University Press.
- Mangan, B. 2001. "Sensation's Ghost: The Non-sensory 'Fringe' of Consciousness." *Psyche* 7. [psyche.cs.monash.edu.au/v7/psyche-7-18-mangan.html](http://psyche.cs.monash.edu.au/v7/psyche-7-18-mangan.html).
- Marks, J., and R. T. Ames, eds. 1995. *Emotions in Asian Thought: A Dialogue in Comparative Philosophy*. With a discussion by R. C. Solomon. Albany: State University of New York Press.
- Martin, M. 1995. "Bodily Awareness: A Sense of Ownership." In *The Body and the Self*, ed. J. L. Bermúdez, A. Marcel, and N. Eilan. Cambridge: MIT Press.
- Mason, W. A. 1985. "Experiential Influences on the Development of Expressive Behaviors in Rhesus Monkeys." In *The Development of Expressive Behavior*, ed. G. Zivin. New York: Academic Press.
- Masson, J. M. 1981. *The Oceanic Feeling: The Origins of Religious Sentiment in Ancient India*. Studies of Classical India, vol. 3. Dordrecht: Reidel.
- Matilal, B. K. 2003. 1978. *Logical and Ethical Issues of Religious Beliefs*. Oxford: Oxford University Press.
- . "Dharma and Rationality." In *A Companion to Indian Ethics: Traditional and Modern*, ed. P. Bilimoria, J. Prabhu, and R. Sharma. Aldershot: Ashgate.

- Matilal, B. K., ed. 1989. *Moral Dilemmas in the Mahābhārata*. Delhi: Motilal Banarsidass.
- McDaniel, J. 1995. "Emotion in Bengali Religious Thought: Substance and Metaphor." In *Emotions in Asian Thought*, ed. J. Marks and R. Ames. Albany: State University of New York Press.
- McDougal, J. 1991. *Theaters of the Mind: Illusion and Truth on the Psychoanalytic Stage*. New York: Brunner Mazel.
- McIntyre, A. 1990. "Is Akratic Action Always Irrational?" In *Identity, Character, and Morality*, ed. A. Rorty and O. Flanagan. Cambridge: MIT Press.
- Mellers, B. A., A. Schwartz, and I. Ritov. 1999. "Emotion-Based Choice." *Journal of Experimental Psychology: General* 128: 1–14.
- Millikan, R. G. 1984. *Language, Thought, and Other Biological Categories*. Cambridge: MIT Press.
- . 1989. "In Defense of Proper Functions." *Philosophy of Science* 56: 288–302.
- . 1996. "Pushmi-Pullyu Representations." In *Mind and Morals: Essays on Cognitive Science and Ethics*, ed. L. May, M. Friedman, and A. Clark. Cambridge: MIT Press.
- Moldoveanu, M. C., and N. Nohria. 2002. *Master Passions: Emotion, Narrative, and the Development of Culture*. Cambridge: MIT Press.
- Moore, G. E. 1903. *Principia Ethica*. Cambridge: Cambridge University Press.
- Morris, J. A., A. Öhman, and R. J. Dolan. 1999. "A Subcortical Pathway to the Right Amygdala Mediating 'Unseen' Fear." *Proceedings of the National Academy of Science* 96: 1680–85.
- Mosbach, H., and W. J. Tyler. 1986. "A Japanese Emotion: *Amae*." In *The Social Construction of Emotion*, ed. R. Harré, 289–307. Oxford: Blackwell.
- Murphy, J. G., and J. Hampton. 1988. *Mercy and Forgiveness*. New York: Cambridge University Press.
- Nagel, T. 1979. *Mortal Questions*. New York: Cambridge University Press.
- Neese, R. M. 1990. "Evolutionary Explanations of Emotions." *Human Nature* 1.3: 261–89.
- Neu, J. 1977. *Emotion, Thought, and Therapy: A Study of Hume and Spinoza and the Relationship of Philosophical Theories of the Emotions to Psychological Theories of Therapy*. London: Routledge and Kegan Paul; Berkeley: University of California Press.
- . 2000. *A Tear Is an Intellectual Thing*. New York: Oxford University Press.
- . 2002. "To Understand All Is to Forgive All—Or Is It?" In *Before Forgiving: Cautionary Views of Forgiveness in Psychotherapy*, ed. S. Lamb and J. Murphy, 17–38. New York: Oxford University Press.
- Nisbett, R. E., and L. Ross. 1980. *Human Inference: Strategies and Shortcomings of Social Judgment*. Englewood Cliffs, NJ: Prentice-Hall.
- Noë, A. 1999. "Experience and Experiment in Art." *Journal of Consciousness Studies* 6.6–7: 123–36.
- Noë, A., ed. 2002. "Is the Visual World a Grand Illusion?" Special issue. *Journal of Consciousness Studies* 9.5–6.
- Nozick, R. 2001. *Invariances: The Structure of the Objective World*. Cambridge: Harvard University Press.
- Nussbaum, M. 1993. "Equity and Mercy." *Philosophy and Public Affairs* 22.

- . 1994. *The Therapy of Desire: Theory and Practice in Hellenistic Ethics*. Princeton, NJ: Princeton University Press.
- . 1995. "Emotions and Women's Capabilities." In *Women, Culture, and Development: A Study of Human Capabilities*, ed. M. Nussbaum and J. Glover. WIDER Studies in Development Economics. Oxford: Clarendon.
- . 2000. "Reply." *Quinnipiac Law Review* 19: 349–70.
- . 2001. *Upheavals of Thought: The Intelligence of Emotions*. Cambridge: Cambridge University Press.
- Nussbaum, M., and H. Putnam. 1992. "Changing Aristotle's Mind." In *Essays on Aristotle's De Anima*, ed. M. Nussbaum and A. Rorty. Oxford: Clarendon.
- O'Donoghue, T., and M. Rabin. 1999. "Addiction and Self-Control." In *Addiction: Entries and Exits*, ed. J. Elster. New York: Russell Sage.
- Oatley, K. 1992. *Best-Laid Schemes: The Psychology of Emotions*. Cambridge: Cambridge University Press.
- Öhman, A. 1986. "Face the Beast and Fear the Face: Animal and Social Fears as Prototypes for Evolutionary Analyses of Emotion." *Psychophysiology* 23: 123–45.
- Öhman, A. 1999. "Distinguishing Unconscious from Conscious Emotional Processes: Methodological Considerations and Theoretical Implications." In *Handbook of Emotion and Cognition*, ed. T. Dalgleish and M. J. Power. Chichester, NY: John Wiley and Sons.
- . 2002. "Automaticity and the Amygdala: Nonconscious Responses to Emotional Faces." *Current Directions in Psychological Science* 11.2: 62–66.
- Öhman, A., and J.J.F. Soares. 1994. "Unconscious Anxiety: Phobic Responses to Masked Stimuli." *Journal of Abnormal Psychology* 102: 121–32.
- Orphanides, A., and Zervos, D. 1998. "Myopia and Addictive Behavior." *Economic Journal* 108: 75–91.
- Ortony, A., G. Clore, and A. Collins. 1998. *The Cognitive Structure of Emotions*. New York: Cambridge University Press.
- Owen, R. 1843. *Lectures on the Comparative Anatomy and Physiology of the Vertebrate Animals, Delivered at the Royal College of Surgeons in 1843*. London: Longman, Brown, Green and Longmans.
- Panksepp, J. 1992. *Affective Neurology*. New York: Oxford University Press.
- . 1998. *Affective Neuroscience: The Foundations of Human and Animal Emotions*, ed. R. Davidson, P. Ekman, and K. Scherer. Oxford: Oxford University Press, 1998.
- Papineau, D. 2002. *Thinking about Consciousness*. Oxford: Oxford University Press.
- Parfit, D. 1984. *Reasons and Persons*. Oxford: Clarendon Press.
- Philosophy East and West*. 1991. 41.1.
- Phillips, S. 1999. "Review of Bilimoria and Mohanty, eds., *Relativism, Suffering and Beyond*." *Journal of the American Oriental Society* 119.2.
- Pinker, S. 1997. *How the Mind Works*. New York: W. W. Norton.
- Pitcher, G. 1965. "Emotion." *Mind* 74: 324–46.
- Plutchik, R. 1980. *Emotions: A Psychoevolutionary Synthesis*. New York: Harper and Row.
- Posner, R. 1981. *The Economics of Justice*. Cambridge: Harvard University Press.
- . 1990. *The Problems of Jurisprudence*. Cambridge: Harvard University Press.

- . 1992. *Sex and Reason*. Cambridge: Harvard University Press.
- Prichard, H. A. 1935. "The Meaning of *Agathon* in the Ethics of Aristotle." *Philosophy* 10: 27–39.
- Prinz, J. 2000. "Getting to the Heart of Emotions." Paper delivered at the American Philosophical Association annual meeting, New York.
- . 2004. *Gut Reactions: A Perceptual Theory of Emotion*. New York: Oxford University Press.
- Pugmire, D. 1998. "Rediscovering Emotion." Edinburgh: Edinburgh University Press.
- Radhakrishnan, V. 1975. *The Principal Upanisads*. London: Allen and Unwin.
- Ramachandran, V., and S. Blakeslee. 1998. *Phantoms in the Brain: Human Nature and the Architecture of the Mind*. New York: William Morrow.
- Rapaport, D. 1967. *The Collected Papers of David Rapaport*, ed. Merton M. Gill. New York: Basic Books.
- Rawls, John. 1971. *Gut Reactions: A Perceptual Theory of Emotion*. Cambridge, MA: Harvard University Press.
- Reisenzein, R. 1983. "The Schachter Theory of Emotion: Two Decades Later." *Psychological Bulletin* 94: 239–64.
- Reisenzein, R., W.-U. Meyer, and A. Schützwohl. 1995. "James and the Physical Basis of Emotion: A Comment on Ellsworth." *Philosophical Review* 4: 757–61.
- Rizzolatti, G., L. Fadiga, V. Gallese, and L. Fogassi. 1996. "Premotor Cortex and the Recognition of Motor Actions." *Cognitive Brain Research* 3: 131–41.
- Roberts, R. C. 1988. "What An Emotion Is: A Sketch." *Philosophical Review* 97: 183–209.
- Robinson, J. 1995. "Startle." *Journal of Philosophy* 92.2: 53–74.
- Rorty, A. O. 1988. *Mind in Action: Essays in the Philosophy of Mind*. Boston: Beacon.
- . 1997. "The Political Sources of Akrasia." *Ethics*.
- . 1998. "The Political Sources of the Emotions: Greed and Anger." *Philosophical Studies*, 143–59. Reprinted in *Mid-West Studies* 20 (1998).
- Rorty, A. O., ed. 1980. *Explaining Emotions*. Berkeley: University of California Press.
- Ross, S. 1984. "Painting the Passions: Charles LeBrun's *Conférence sur l'expression*." *Journal of the History of Ideas* 45: 25–47.
- Royzman, E. R., and J. J. Sabini. 2001. "Something It Takes to Be an Emotion: The Interesting Case of Disgust." *Journal for the Study of Social Behavior* 31.1: 29–60.
- Rozin, P. 1976. "The Evolution of Intelligence and Access to the Cognitive Unconscious." In *Progress in Psychobiology and Physiological Psychology*, ed. J. M. Sprague and A. N. Epstein. Vol. 6. New York: Academic Press.
- Salovey, P., B. T. Bedell, J. B. Detweiler, and J. D. Mayer. 2000. "Current Directions in Emotional Intelligence Research." In *The Handbook of Emotions*, ed. M. Lewis and J. M. Haviland-Jones. 2d ed. New York: Guilford.
- Sartre, J.-P. 1948. *The Emotions: Sketch of a Theory*. Trans. B. Frechtman. New York: Citadel.
- Schachtel, E. 1959. *Metamorphosis*. New York: Basic Books.
- Scarantino, A. 2004. "Affordances Explained." *Philosophy of Science*, supplement, *Proceedings of the 2002 Biennial Meeting of the Philosophy of Science Association: Contributed Papers*.

- Schachter, S., and J. E. Singer. 1962. "Cognitive, Social, and Physiological Determinants of Emotional State." *Psychological Review* 69: 379–99.
- Schelling, T. C. 1980. *The Strategy of Conflict*. Cambridge: Harvard University Press.
- Scherer, K. R. 1999. "Appraisal Theory." In *The Handbook of Emotion and Cognition*, ed. T. Dalgleish and M. J. Power. Chichester, NY: John Wiley and Sons.
- Searle, J. R. 1983. *Intentionality: An Essay in the Philosophy of Mind*. Cambridge: Cambridge University Press.
- . 1995. *The Construction of Social Reality*. New York: Free Press.
- Seligman, M. 1975. *Helplessness: On Development, Depression, and Death*. New York: W. H. Freeman.
- Sharma, R. M. 1992. "Empathy: A Retrospective on Its Development in Psychotherapy." *Australian and New Zealand Journal of Psychiatry* 26.3: 377–90.
- . *Understanding the Concept of Empathy and Its Foundations in Psychoanalysis*. Lewiston, NY: Edwin Mellon.
- Sibley, F. 1965. "Aesthetic and Non-aesthetic." *Philosophical Review* 74: 135–59.
- Skorupski, J. 2000. "Irrealist Cognitivism." In *Normativity*, ed. J. Dancy. Oxford: Blackwell.
- Smith, B. 2002. "Keeping Emotions in Mind." In *Understanding Emotions: Mind and Morals*, ed. P. Goldie. Aldershot: Ashgate.
- Solomon, R. C. 1973. "Emotions and Choice." *Review of metaphysics* 28.1: 20–41.
- . 1976. *The Passions*. New York: Doubleday.
- . 1980. "Emotions and Choice." In *Explaining Emotions*, ed. A. Rorty, 103–26. Berkeley: University of California Press.
- . 1988. "On Emotions as Judgments." *American Philosophical Quarterly* 25.2.
- . 1993a. *The Passions: Emotions and the Meaning of Life*. 2d ed. Indianapolis, IN: Hackett.
- . 1993b. "The Philosophy of Emotions." In *The Handbook of Emotions*, ed. M. Lewis and J. M. Haviland. New York: Guilford.
- . 1998. "The Politics of Emotion." In *Midwest Studies in Philosophy XXII*, ed. Peter French et al. Notre Dame: University of Notre Dame Press.
- Spinoza, B. [1677] 1985. *Ethics*. In *The Collected Works of Spinoza*, ed. E. Curley. Princeton, NJ: Princeton University Press.
- Spinoza, B. 1981. *Ethics*. Malibu, CA: J. Simon.
- Spivak, G. C. 1999. *A Critique of Postcolonial Reason: Toward a History of the Vanishing Present*. Cambridge: Harvard University Press.
- Stich, S. 1983. *From Folk Psychology to Cognitive Science*. Cambridge: MIT Press.
- Stocker, M. 1976. "The Schizophrenia of Modern Ethical Theories." *Journal of Philosophy* 73.
- . 1983. "Psychic Feelings: Their Importance and Irreducibility." *Australian Journal of Philosophy* 61: 5–26.
- . 1989. *Plural and Conflicting Values*. New York: Oxford University Press.
- Stocker, M., with E. Hegeman. 1996. *Valuing Emotions*. Cambridge Studies in Philosophy. Cambridge: Cambridge University Press, 1996.
- Strack, F., L. L. Martin, and S. Stepper. 1988. "Inhibiting and Facilitating Conditions of Facial Expressions: A Nonobtrusive Test of the Facial Feedback Hypothesis." *Journal of Personality and Social Psychology* 54: 768–77.

- Tanner, M. 1976–77. “Sentimentality.” *Proceedings of the Aristotelian Society* 77: 127–47.
- Taylor, C. 1989. *Sources of the Self*. Cambridge: Harvard University Press.
- Teasdale, J. D. 1999. “Multi-level Theories of Cognition-Emotion Relations.” In *The Handbook of Cognition and Emotion*, ed. T. Dalgleish and T. Power. New York: John Wiley and Sons.
- Thalberg, Irving. 1978. “Mental Activity and Passivity.” *Mind*.
- Tice, D., E. Braslavsky, and R. Baumeister. 2001. “Emotional Distress Regulation Takes Precedence over Impulse Control.” *Journal of Personality and Social Psychology* 80: 53–67.
- Tomkins, S. S. 1962. *Affect, Imagery, and Consciousness*. New York: Springer.
- Tompkins, J. 1987–88. “Me and My Shadow.” *New Literary History* 29: 169–78.
- Tooby, J., and L. Cosmides. 1992. “The Psychological Foundations of Culture.” In *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*, ed. J. H. Barkow, L. Cosmides, and J. Tooby. Oxford: Oxford University Press.
- Watson, J. B. 1929. *Psychology from the Standpoint of a Behaviorist*, 3rd ed. Philadelphia: Lippincott.
- Wierzbicka, A. 1985. *Lexicography and Conceptual Analysis*. Ann Arbor, MI: Karoma.
- . 1999. *Emotions across Languages and Cultures: Diversity and Universals*. Cambridge: Cambridge University Press.
- Williams, B. 1973. “Egoism and Altruism.” In *Problems of the Self*. Cambridge: Cambridge University Press.
- . 1985. *Ethics and the Limits of Philosophy*. Cambridge: Harvard University Press.
- Wilson, J. Q. 1993. *The Moral Sense*. New York: Free Press.
- Wollheim, R. 1999. *On the Emotions*. New Haven, CT: Yale University Press.
- Woolf, V. *Leave The Letters Till We’re Dead: Collected Letters*. Vol. 6, 1936–41, ed. Nigel Nicholson. London: Hogarth.
- Zajonc, R. B. 1980. “Feeling and Thinking: Preferences Need No Inferences.” *American Psychologist* 35: 151–75.
- . 1984a. “On the Primacy of Affect.” In *Approaches to Emotion*, ed. K. Scherer and P. Ekman. Hillsdale, NJ: Erlbaum.
- . 1984b. “The Interaction of Affect and Cognition.” *American Psychologist* 39: 117–29.
- . 1994. “Evidence for Nonconscious Emotions.” In *The Nature of Emotions*, ed. Paul Ekman and Richard J. Davidson, 293–97. New York: Oxford University Press.
- Zajonc, R. B., S. T. Murphy, and M. Inglehart. 1989. “Feeling and Facial Efference: Implications of the Vascular Theory of Emotion.” *Psychological Review* 96: 395–416.
- Zillman, D., and J. Cantor. 1976. “Effect of Timing of Information about Mitigating Circumstances on Emotional Responses to Provocation and Retaliatory Behavior.” *Journal of Experimental Social Psychology* 12: 38–55.

## INDEX

- action, 154, 160  
action readiness, 255, 257  
action tendencies, 151-53, 157-59  
action theory, 135, 138  
adrenalin, 51  
aesthetics, 70, 73  
affect, 46, 78-79, 84-88, 87-88, 130-  
32, 137, 145, 241, 243, 248  
affect program, 19, 24, 77-78, 234, 236,  
239-40, 242-43  
*ahimsa*, 218  
*akrasia*, 66-67  
*alanes*, 212  
alexithymia, 146  
*amae*, 39  
amygdala, 35-36, 42, 46, 67, 246  
anger, 39, 142-43, 223-5  
animals, emotions of, 9-10, 12, 14, 16,  
19, 25, 35, 63  
appraisal, 34, 38, 46, 85, 128, 186  
affective, 34, 36-43  
cognitive, 32, 36, 38-39, 41, 43, 57  
emotional, 36  
theory, 57-58, 243-48  
Aristotle, 3-4, 127-28, 131, 141, 147,  
175, 186, 188, 191, 203, 222, 224,  
263, 270, 271, 272, 276  
Arjuna, 219-6  
Asian philosophy, 214-15  
auditory thalamus, 35  
Augustine, Saint, 170, 276  
autonomic nervous system, 31, 85-86,  
93  
axiology, 69, 73-75  
Baier, Annette, 112  
basic emotion, 35, 236-43, 248  
systems of, 31, 41  
Bedford, Errol, 116  
behavior, 68  
behaviorism, 32  
belief, 28, 79, 83, 116, 175, 178, 188-  
91, 204, 246  
Bell, Charles, 15-16  
Bell, Clive, 210  
Bell, Vanessa, 209  
Bhagavad Gita, 216, 219  
*bhava*, 215-19, 223, 227  
Bishop Juxon, 207  
Blackburn, Simon, 240, 243  
Blake, William, 81, 178  
Bovens, Luc, 67  
Bratman, Michael, 206  
Brentano, Franz, 107  
Buber, Martin, 107  
Buddha, 217, 227  
Buddhism, 214, 216, 221, 227, 222,  
224-6  
Calhoun, Cheshire, 79  
Cannon, Walter B., 48  
Capgras syndrome, 65, 67-68

- Caraka-samhita*, 216  
 cause of emotion, 20–21. *See also* object  
 of emotion *and* intentionality  
 chalance, 203–8, 210–11  
 Charland, Louis, 240–41, 243  
 Chomsky, Noam, 126  
 Christianity, 167, 172–73  
 cingulate cortex, 46  
 cognate emotions, 17–18  
 cognition, 32–34, 36–38, 42, 61, 67–68,  
 73, 76, 78–81, 84–85, 87–88, 116,  
 128–29, 158–59, 192, 195  
 cognitive appraisal theory, 54  
 cognitive theories, 10–11, 14, 19, 25, 61–  
 62, 73, 76–77, 79, 81, 84, 88, 91,  
 103, 108, 110, 116, 186  
 colliculuous, pulvinar and superior, 46  
 consciousness, 64, 252  
 continental philosophy, 3, 271  
 core relational themes, 53–57, 244, 246  
 Crick, Francis, 140–41, 145  
 culture and social interaction, 63, 167–  
 70, 277–78
- Damasio, Antonio, 48, 54–55, 57, 78,  
 125–27, 130, 137, 234, 237, 239–  
 43  
 Darwin, Charles, 10, 14–24, 45, 126,  
 202–3, 211, 237  
 Davidson, Donald, 66  
 de Silva, Padmasiri, 224  
 De Sousa, Ronald, 70, 80–81, 118, 131  
 de Waal, Frans, 63  
 Descartes, René, 11–15, 44, 64, 67, 69,  
 79–80, 85, 111, 117, 126, 171, 175,  
 200, 203–4, 211, 212, 275, 276  
*Descartes' Error*, 125, 137  
 desire, 135–36, 139, 153–54, 177  
 in Indian philosophy, 222–25  
 determinism, 66, 172–73, 175  
 de Waal, Frans, 63  
 Dewey, John, 81, 107, 117, 120  
 dualism, 125, 126  
 Dwyer, General, 218
- Eibesveldt, Eibl, 211  
 Ekman, Paul, 31, 45, 53, 93, 202, 204,  
 211, 236, 242–43  
 elicitors, of emotion, 54–57, 244  
 eliminativism, 233  
 Ellsworth, Phoebe, 38  
 emotion, types of  
 appropriate, 118–20, 130, 132  
 basic, 35, 236–43, 248  
 cerebral, 49–50  
 cognate emotions, 17–18  
 complex, 236–37, 239–43  
 intellectual, 141–43, 145–46, 148  
 moral (Indian), 220  
 primitive, 10–11, 13–14, 55  
 short-term, 93  
 emotional truth, 70, 72  
 emotion-proper properties, 94, 97–100  
 emotivism, 3  
 empathy, 125–26, 159  
 empiricism, 13–14, 17, 186  
 engagement in the world, 77, 84, 86–88  
 Enlightenment, the, 111, 112  
 envy, 164–65. *See also* jealousy  
 Epics, the, 226  
 Epicurus, 3, 62  
 epistemology, 98–100, 102–3, 107–9,  
 111–15, 210, 257  
 ethics, 112, 147  
 eudaimonia, 190, 193  
 eudaimonistic ethical theory, 189–90  
*Eustace Diamonds, The*, 205  
 evolutionary function of emotions, 257–  
 58  
 existentialism, 83–84  
*Expression of the Emotions in Man and  
 Animals, The*, 10  
 expressions, 14, 23, 179–80, 204  
 conventional, 16–17  
 facial, 15, 17–19, 21–22, 24, 31, 34,  
 43, 46, 53, 86, 204, 239, 246  
 involuntary, 15–16
- fago*, 9  
 fatalism, 66  
 fear, 35, 41, 42, 55  
 feeling, 84–86, 91–97, 101–2, 133, 178,  
 195, 252–53, 260  
 feeling theory, 84  
 in Indian philosophy, 215–17  
 feminism, 108, 126  
 fight or flight reaction, 30  
 folk psychology, 28, 39, 41, 43, 270,  
 272, 273  
 forgiveness, 170, 172–73, 176  
 Frank, Robert, 237  
 Frankfurt, Harry, 138, 203, 205–6, 210  
 Freud, Sigmund, 25, 64, 83, 101, 107,  
 115, 166, 211–12, 226

- Friege, Gottlob, 133  
 Friesen, W. V., 45, 53  
 Frijda, Nico, 32, 151  
 frontal cortex, 55
- galvanic skin response, 33, 36, 65, 85–86  
 Gandhi, M., 218  
 Goldie, Peter, 84  
 Gordon, Robert, 51  
 Greenspan, Patricia, 41, 43, 79  
 grief, 193–96  
 Griffiths, Paul, 77, 80, 84, 88, 127, 204, 233, 235  
 Guntrip, Henry, 146
- Heidegger, Martin, 4, 84, 87, 135, 139, 271  
 Helm, B., 204  
 Hobbes, Thomas, 63  
 Hohmann, G. W., 47–48  
*Horace*, 70  
*How the Mind Works*, 6  
 hubris, 169  
 Hume, David, 107, 140, 152, 168–69, 200–204, 206–7, 210, 211, 220, 270, 274, 276  
 Hutcheson, Francis, 107, 111
- impatience, 153, 155. *See also* urgency  
 Indian philosophy, 214–48  
   on the body, 215  
   on feelings, 215–17  
   on intentionality, 217  
   on moral emotions, 220  
   on rationality, 219  
 information processing, 255, 266  
 inheritance of emotions, 10, 15–17, 19, 23  
 insular cortex, 46  
*Integration and Self-healing: Affect, Trauma, and Alexithymia*: “Affects are familiar to everyone,” 137  
 intellect, as compared with emotion, 258–59, 262–64  
 intellectual activity, 136–40, 143–46, 148  
 intellectualization, 144–45  
 intentional states, 9–15, 19–20, 22, 25  
 intentionality, 54–57, 63, 77, 79, 93, 97, 131, 177, 188, 190, 195, 203, 245–46, 248, 252–53, 259  
   in Indian philosophy, 217  
   *See also* object of emotion *and* cause of emotion  
 Izard, Carroll, 78, 202, 211, 242
- Jain, 214  
 James, William, 25, 30, 37, 44–57, 64–65, 76–78, 81, 84–85, 88, 107, 234, 250  
 James-Lange Theory, 47, 49, 51–52, 54, 56–57  
 jealousy, 109, 131–32, 163–66. *See also* envy  
 judgment, 10, 28, 37–38, 43–44, 46, 51, 62, 76–79, 81–85, 87, 98, 100, 107–8, 184, 187, 191, 194–96, 211  
   evaluative, 29, 30, 32, 37, 40, 54–57, 76, 79, 127–29, 131–33, 141, 146, 185, 190, 192–93, 244–46, 260–61, 272  
   kinesthetic, 82, 86  
   moral, 116, 222  
 judgment theory, 28–29, 32, 40, 43
- kama*, 222–23, 225  
 Kant, Immanuel, 107, 111, 200, 271  
 karma, 222, 225–26  
 Kenny, Anthony, 3–4  
 Kierkegaard, Soren, 107, 135, 139  
 King Charles I, 207  
 Kinsbourne, Marcel, 70  
*klesas*, 221, 225  
 knowledge, 111, 112, 113  
   extraspective, 92–94, 96, 99, 101, 103  
   introspective, 92, 95–96, 101, 103  
 Krishna, 219–22, 224–26  
 Krystal, Henry, 137
- La Rochefoucauld, 159  
 Laing, R. D., 171  
 Lange, Carl, 44–46, 48, 50, 52–57  
 language, 13–14, 117  
 Larkin, Philip, 62  
 Lazarus, Robert, 32–33, 35, 53–54, 244–45  
*le bourgeois gentilhomme*, 7  
 Le Brun, Charles, 15  
 Le Doux, Joseph, 35–36, 41–42, 78, 238, 241  
 Lear, Jonathon, 83  
 Levenson, R. W., 45, 53  
 limbic system, 68, 240, 241

- Locke, John, 12–14  
 logic of emotions, 261–62, 265–66  
 Luther, Martin, 206  
 Lutz, Catherine, 39  
 Lyons, William, 30, 37, 40
- Machiavelli, Nicolo, 63  
 MacIntyre, Alastair, 206  
 MacLean, Paul, 68, 241  
*Mahabharata*, 215–16, 219, 221, 224, 226  
 Malcolm, Norman, 20  
 Matilal, 220, 226  
 McDaniel, June, 216  
 McDougal, Joyce, 146  
 McIntyre, Alasdair, 67  
 medieval philosophy, 3–4  
 memory, 67  
   declarative, 42  
   emotional, 42  
 Mencius, 74  
 mental capacity, 251–54, 257  
 mental mode, 254–55, 257–67  
   emotional, 256–67  
 mental states, emotions as, 203–4, 250–51  
*Metamorphosis*, 7  
 Mill, John Stuart, 135  
 Millikan, Ruth, 74, 246  
 mirror neurons, 48–49  
 Moliere, 137  
 mood, 40, 99  
 Moore, G. E., 210  
 moral theory, 111, 112  
   British, 115, 126  
   ancient Greek, 189  
 morality, 70, 111, 112, 180, 210, 154, 220  
   and judgment, 116  
 motivation, 65  
 motor cortex, 48  
 Murdoch, Iris, 136  
 Murphy, Jeffrey, 79
- Nagel, Thomas, 211  
 natural kind, 233–38, 240–41, 244, 247–48  
 naturalism, 130–31, 210  
 neocortex, 239–40  
 Neo-Stoics, 186–87, 190–93, 203, 207, 275
- Neu, Jerome, 79–82  
 neurophysiological mechanism, 16, 18, 20–22, 24–25  
 neurophysiology, 71  
 Nietzsche, Friedrich, 271  
 Noe, Alva, 64  
 norms  
   moral, 157  
   social, 156–58  
 Nussbaum, Martha, 61, 244–46
- object of emotion, 20–23, 54, 56, 71–72, 99, 186–89, 203–4. *See also* cause of emotion *and* intentionality  
 obstacle theory of emotion, 117  
 ontology, 211, 257
- Panksepp, Jaak, 78, 241–42  
 paradigm scenario, 80–81, 83, 118–19, 131–32  
*Passions of the Soul, The*, 6  
 perceptions, sensory, 69, 263. *See also* sensory experience  
 perspectival account, 125, 128, 130  
 phenomenology, 70–71, 84–86, 95, 177, 201  
 phobia, 41  
 physiological change, as emotions, 30–31, 33, 39, 44–49, 51–53, 55–57  
 physiological response, 37, 40  
 physiology of emotions, 65, 76, 85–86  
 Pinker, Stephen, 126  
 Pitcher, George, 3  
 Plato, 107, 145  
 pride, 167–70  
 primitive emotion, 10–11, 13–14, 55  
 primitivism, 76, 78, 83  
 Prinz, Jesse, 239  
 proposition, 10–11, 19, 70, 77, 80, 87, 128–31, 133, 144, 191–93  
 propositional attitude theory, 77, 80, 127
- Ramachandran, V., 67  
 Rapaport, David, 145  
 rationalist theory, 111, 125–27, 129–33, 143–44, 219  
 Rawls, John, 220, 271  
 regret, 178  
 Reid, T., 138  
*Reisenzein, Meyer Schutzwhol*, 52

- remorse, 178  
 resentment, 170, 172–73  
 Roberts, Robert, 79  
 Rodin, A., 203  
*Romeo and Juliet*, 9  
 Rorty, Amélie, 42, 84  
 Rousseau, J. J., 270  
 Ryle, Gilbert, 117, 171
- Sanskrit aestheticism, 207  
 Sartre, Jean-Paul, 4, 76, 107, 109, 117, 163, 175, 270  
 Schachte, Ernest, 137  
 Schachter, S., 51–52  
 Schelling, T. C., 127  
 schematic mechanism, 260–61, 266  
*Schizoid Phenomena, Object Relations and the Self*, 6  
 screaming, 23  
 second somatosensory cortex, 46  
 self-respect, 169–70  
 Seneca, 159, 187, 270  
 sensory experience, 12, 14. *See also* perceptions, sensory  
 Shakespeare, W., 9, 17, 100  
 shame, 158  
 sin, 167, 169, 170  
 Singer, J. E., 51–52  
 Smyth, Ethel, 210  
 social construction theory, 39, 131  
 Socrates, 173  
 Solomon, Robert, 29, 37–38, 62, 108, 117  
 somatic theory, 45–46, 56  
 spinal cord, 47–48  
 Spinoza, Benedict, 79–80, 82, 163, 173, 176–78, 180, 203–4, 227, 270, 275  
 startle reaction, 243  
 stimuli, 34, 36, 41–42  
 Stocker, Michael, 79, 81–82, 84, 224  
 Stoics, 3, 10, 79, 82, 128, 185–86, 191, 214  
   Neo-Stoics, 186–87, 190–93, 203, 207, 275  
 subjectivity of emotions, 107–20
- Taylor, Charles, 206, 210, 212  
*Tear Is an Intellectual Thing*, A, 1  
 teleology, 66, 117  
 temporal lobe, 42  
 Thalberg, Irving, 138  
*Theaters of the Mind: Illusion and Truth on the Psychoanalytic State*, 6  
 thought, 79, 80, 83  
   active, 177–78  
   encoded, 11  
   passive, 177–78  
 Tomkins, Sylvan, 242  
*Treatise of Human Nature*, 0  
 Trollope, A., 205
- Upanishads*, 222  
 urgency, 153–55. *See also* impatience  
 utilitarianism, 190
- value, 70, 167, 168, 170, 189–90, 192, 210, 212  
   transvaluation of, 167–68, 170  
 Van Gogh, Vincent, 114  
 voluntary/involuntary quality of emotions, 76, 86, 93, 179, 205–6, 260  
 von Neumann-Morgenstern utility function, 160
- Walter, Canon B., 30, 33  
 Walton, Kendall, 79  
 Watson, John, 32, 140–41, 145  
*What Emotions Really Are*, 7  
 Wilde, Oscar, 70  
 Williams, Bernard, 138–39  
 Winnicott, D. W., 166, 171  
*Winter's Tale*, A, 0  
 wishes, 153  
 Wittgenstein, Ludwig, 166, 177  
 Wollheim, Richard, 250–51  
 Woolf, Virginia, 210
- yoga, 221, 224, 226
- Zajonc, Robert, 32, 33, 35