Emotions

An Essay in Aid of Moral Psychology

ROBERT C. ROBERTS
Baylor University

CAMBRIDGE
UNIVERSITY PRESS

PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE
The Pitt Building, Trumpington Street, Cambridge, United Kingdom

CAMBRIDGE UNIVERSITY PRESS
The Edinburgh Building, Cambridge CB2 2RU, UK
40 West 20th Street, New York, NY 10011-4211, USA
477 Williamstown Road, Port Melbourne, VIC 3207, Australia
Ruiz de Alarcón 13, 28014 Madrid, Spain
Dock House, The Waterfront, Cape Town 8001, South Africa

http://www.cambridge.org

First published 2003

Printed in the United Kingdom at the University Press, Cambridge

Typeface ITC New Baskerville 10/12 pt. *System* LaTeX 2ε [TB]

A catalog record for this book is available from the British Library.

Library of Congress Cataloging in Publication data

Roberts, Robert C., 1942–
Emotions : an essay in aid of moral psychology / Robert C. Roberts.
p. cm.
Includes bibliographical references and index.
ISBN 0-521-81978-4 – ISBN 0-521-52584-5 (pbk.)
1. Emotions (Philosophy) 1. Title.
B105.E46 B76 2002
128′.37–dc21 2002025620

ISBN 0 521 81978 4 hardback
ISBN 0 521 52584 5 paperback

Contents

1

Studying Emotions

1.1. WHAT THIS BOOK IS ABOUT

Anthony Trollope comments about an unsavory character who looms large in his novel *The Prime Minister* (Chapter 58):

> The abuse which was now publicly heaped on the name of Ferdinand Lopez hit the man very hard; but not so hard perhaps as his rejection by Lady Eustace. That was an episode in his life of which even he felt ashamed, and of which he was unable to shake the disgrace from his memory. He had no inner appreciation whatsoever of what was really good or what was really bad in a man's conduct. . . . In a sense he was what is called a gentleman. He knew how to speak, and how to look, how to use a knife and fork, how to dress himself, and how to walk. But he had not the faintest notion of the feelings of a gentleman. He had, however, a very keen conception of the evil of being generally ill spoken of.

Without directly mentioning any of Lopez's actions, Trollope here unmistakably sketches a man of momentous moral defects, just by indicating his patterns of emotional responsiveness – that he is more ashamed of being rejected by a classy female adventurer than of being the object of public moral opprobrium, but not at all ashamed of his shameful deeds. His lack of appreciation for good and bad action, suggests Trollope, is due to his emotional unresponsiveness to actions in moral terms (notice how Trollope mixes descriptions of Lopez's emotional dispositions with cognitive ascriptions like "no inner appreciation," "not the faintest notion," "a very keen conception"). The structure of his emotions explains why he does so much evil, why he has so little moral understanding, and why his life and the lives of those he touches closely are so miserable.

The involvement of emotions in what may be broadly termed the "moral" character of our lives is pervasive and deep. Because emotions are often impulses to act, their quality strongly affects the quality of what we do. Those who are prone to strong and inappropriate fear and anger tend to act and behave in a certain set of familiar ways, while compassion and the emotions

of friendship incline people to actions of another kind. These two sorts of emotional tendencies, and many others, may coexist in a single person, thus making people complex and morally puzzling. But emotions are not *just* "causes" of actions; they may also determine the identity of our actions. The very "same" action of shoving a person into a ditch may be done from anger at the shoved person or fear for her life; in the first case the agent is *getting revenge* (let us say), and in the second he is *protecting against danger.* What I have said has already suggested that our character or personality is in large part a disposition to be affected in one set of ways or another: One who is regularly angered by trivial offenses to his private person but seldom or never by significant offenses against others or against the public good is a mean-spirited person. Someone who rejoices in the flourishing of family and friends, for their sake, has a nobler character than one who is unaffected by their weal, or who is affected by it, not for their sake but, say, for the sake of his own convenience. To be emotionally unsusceptible to another's well- or ill-being for the other's sake is to be incapable of friendship with that other, on at least one conception of friendship; the most important relationships of our lives are constituted, in large part, by our dispositions to react with specific emotions to the other and his vicissitudes. Besides these connections to action, character, and relationships, emotions are a kind of eye for value and the import of situations, a mode of spiritual perception that may be deep and wise, or shallow and foolish. Because of these and other types of importance, certain regular patterns of emotional response are characteristic of the flourishing, mature, and "happy" human life, while alternative patterns constitute ill-function and immaturity and tend to misery.

This volume and its projected companion aim to contribute to our understanding of moral personality conceived in a broad sense of "moral," with a particular focus on the place of emotions and emotional formation in that personality. The conception of *moral* to which I refer includes not only our responses to duties and permissions, but also our happiness (which certainly does not imply always feeling good) – what kind of life, and in particular what formation of personality, and thus of relationships with others, constitutes human well-being all around. The work is divided into two parts. The present volume is on the nature of emotions and feelings and, in Chapter 4, begins to treat their connections to the moral life. The projected second volume begins with a general account of the relation of emotions to morality in my broad sense of the word, and then it offers accounts of a number of particular traits of the flourishing personality with special reference to the emotions and emotionlike states that exemplify or interact with them.

The project of understanding the good life in terms of the virtues, and the virtues in terms of their relationships to the emotions, is nothing new. Aristotle says that moral virtue is concerned with passions and actions, and in his accounts of particular virtues the passions often figure even more

prominently than the actions. The association of the virtues with the passions (many of which we would call emotions in modern English) recurs almost wherever the virtues are carefully reflected on in the history of philosophy. Thus Thomas Aquinas devotes Questions 22–48 of the first part of the second part of his *Summa Theologiæ* to a study of the passions, preparatory to his general discussion of virtue in Questions 55–67 and his detailed discussions of the virtues and vices in the second part of the second part, Questions 1–170, many of which themselves involve discussions of passions such as hope, fear, despair, joy, love, hatred, and envy. Book II of David Hume's *A Treatise of Human Nature* – "Of the Passions" – prepares the reader for Book III, in which he presents his ethics of virtue. Adam Smith's ethics, as presented in *The Theory of Moral Sentiments*, is likewise an ethics of virtue that focuses strongly on the passions. In our own period, when John Rawls turns to address justice as a trait of persons (rather than a structural feature of institutions), he finds it necessary to speak not just of dispositions to act, but of moral sentiments such as anger and guilt.[1]

This book is not a historical work, but I intend it as a contribution to this long discussion. I hope that it is in some ways a refinement of its predecessors. At any rate, it is dependent on them for direction and inspiration, as well as for the proposals that have fueled my thought, even when I disagree with them. As befits its historical location at the beginning of the 21st century, this book is more sensitive than its forebears to the possibility that neither emotions nor virtues are the same in every cultural setting, but instead vary to some extent with systems of custom, interest, and belief. While attempting to credit the diversity or potential diversity of human emotions and virtues, my discussions are also more resolutely particularistic. It seems to me that the way to study virtue is to study the virtues, and to do so rather in depth. "Virtue theory," especially in our time but also earlier, has often been long on generalizing accounts and short on careful exploration of particular virtues. Particular virtues are treated as illustrations of general theory, rather than as a fund of insight out of which any generalizations that are possible may emerge.

Accordingly, much of the second volume will study particular virtues, with special emphasis on their dynamic and internal connections with emotions and emotion dispositions. I comment on more general theoretical questions because they seem naturally to arise out of the particular discussions of virtues. Thus the method is "empirical" in the broad sense that it follows Wittgenstein's dictum, "Don't think. Look!", though lots of the looking in this sort of case is a kind of thinking – thinking about examples, as Wittgenstein's also is. Similarly, in the present volume, Chapter 3 is devoted to an extensive detailing of particular emotion types as well as of

[1] *A Theory of Justice* (Oxford: Oxford University Press, 1972), Sections 66–67 and 73–74.

emotionlike phenomena that are sometimes treated as emotions, such as surprise, amusement (at what is comical), and vanity.

1.2. THE SUPPOSED POVERTY OF CONCEPTUAL ANALYSIS

How shall we conduct an inquiry into the emotions that will serve well the study of the virtues? The methods of many disciplines have been used to study the emotions. Philosophers, from Aristotle[2] to the present, have used an approach that today would be called conceptual or philosophical analysis, one that I want to examine closely in this opening chapter because recently it has been under attack and I will argue that it is still the central approach for our purposes. But in the 19th and 20th centuries a number of other approaches have been developed. Emotions have been examined by the methods of evolutionary biology, experimental psychology, brain science, psychoanalysis and other clinical approaches, cultural anthropology, and cultural history and the history of ideas. In each case, one or another of a variety of theories forms a more or less definite background of the examination and shapes its results. For example, evolutionary biologists tend to think of emotions as behavioral response mechanisms that (at least in our evolutionary past, and in some cases also now) promote physical survival, while many anthropologists think of emotions as culturally determined patterns of experience and behavior that serve various social functions (though some anthropologists are psychoanalytic, and so stress less the determinations of culture). Brain scientists tell a rather different story about emotions, one in terms of brain circuitry and neurotransmitters, but typically lean on the evolutionary conception of emotion, while historians of the emotions may exploit psychoanalytic theory or perhaps a more cognitive-behavioral framework. In addition to these disciplines, fiction writing should be mentioned, though it is not theoretical or academic in the way the other disciplines are. Nevertheless, writers such as Jane Austen, Fyodor Dostoevsky, Leo Tolstoy, Charles Dickens, and George Eliot are very astute observers of emotions in the context of the narrative flow of human life, and are especially important for our purposes since they so often depict the emotions as expressing traits of persons' character. Most of the other disciplines focus much less on emotions that differentiate persons of one moral formation from persons of another, and seldom are emotions set in as rich a narrative context as they are in literature. A possible exception is psychoanalysis.

Conceptual analysis is an approach to the investigation of emotions that takes major clues about them from the ways people talk about the emotions in the contexts of their life. As I understand the practice and as the word "clue" suggests, it is not a purely lexicographical or syntactical/semantic

[2] See Aristotle's *Rhetoric*, especially Book II, Chapters 1–11 (1377b15–1388b30).

approach. It is not as though we could expect to find out what emotions are by looking up "emotion" and/or words like "anger" and "nostalgia" in the dictionary, nor could we expect to "analyze" such concepts merely by summarizing the various conditions under which the most proficient speakers of English ascribe emotion or anger to themselves and others. Such information about how the best speakers of English use the emotion words is an important part of conceptual analysis, but the analyst is very much in the business of *interpreting* these facts of usage. For one thing, even the best English speakers use vocabulary loosely and shiftingly, so conceptual analysis will involve normative decisions about what is the right and central usage. But beyond this, the conceptual analyst typically offers some general schema by which he or she proposes to make sense of the "data" of linguistic usage. (Consider the various schemata that have been offered by such philosophers as Robert Solomon,[3] Patricia Greenspan,[4] and Robert Gordon.[5]) Furthermore, as a person who not only speaks about emotions, but also experiences them and experiences their connections with actions, perceptions, desires, sensations, and the like, the analyst is also very concerned to make sense of his or her experience and the experiences of other human beings. Thus as I conceive conceptual analysis, it is particularly based on collection of and reflection about examples from everyday human life, many of which can be understood only in the light of a fairly rich narrative background. This preoccupation represents an overlap with literary and psychoanalytic examinations of emotion and a rather strong contrast with biological and neuroscientific examinations. The conceptual analyst, as I understand his *métier*, will look for formulations regarding emotion and particular emotion types, and will be particularly interested in potential counterexamples, also from everyday life, to his formulations.

Conceptual analysis has been criticized as an inadequate approach to the emotions along two different lines by Amélie O. Rorty and Paul E. Griffiths. The two lines of criticism have in common the suggestion that the conceptual scheme provided by our ordinary language about the emotions is a deeply misleading, and perhaps even internally incoherent, indicator of the nature of emotions. Thus any analysis that takes that scheme at face value and as a point of departure is doomed to deep error. Each of these authors promotes an alternative approach. Rorty proposes that we study the history of the philosophies of the emotions because in her view those variegated philosophies have *constituted* the incoherent or apparently incoherent concept of emotion that analysts try vainly to make sense of. Griffiths thinks that the best scientific accounts of the phenomena that we call "emotions" – essentially, accounts from evolutionary biology and its

[3] *The Passions* (Garden City, New York: Doubleday, 1976).
[4] *Emotions and Reasons* (New York: Routledge, 1988).
[5] *The Structure of Emotions* (Cambridge, England: Cambridge University Press, 1987).

auxiliary experimental psychology – show that "emotions" form such a qualitatively diverse set of phenomena that the concept *emotion* and the concepts of particular types of emotion are useless for the purpose of genuine knowledge. I shall examine the arguments and proposals of Rorty and Griffiths, bringing into my critique of Griffiths some observations about the best recent work on the neuroscience of emotions. I shall then end this Introduction with a sketch of a kind of conceptual analysis that avoids the legitimate criticisms that have been leveled against conceptual analysis of the emotions as it was practiced in the 20th century.

1.3. DECONSTRUCTING *EMOTION* VIA THE HISTORY OF PHILOSOPHY

Amélie Rorty begins her paper, "Aristotle on the Metaphysical Status of *Pathe*,"[6] by commenting on the deplorable state of present-day philosophical theorizing about the passions and emotions. The discussions are "arbitrary and factitious" and "puzzlingly pulled in what appear to be opposing directions" (p. 521); these "persistent and unresolvable contemporary polemical debates carry an air of a chimaeral construction" (p. 545). The reason for this apparent impasse is that the concept under discussion itself contains these "opposing directions"; the discussions only reflect tensions internal to the concept:

> We sometimes hold people responsible for their emotions and the actions they perform from them. Yet normal behavior is often explained and excused by the person 'suffering' an emotional condition. We treat emotions as interruptions or deflections of normal behavior, and yet also consider a person pathological if he fails to act or react from a standard range of emotions. Sometimes emotions are classified as a species of evaluative judgments whose analysis will be given in an adequate theory of cognition. But sometimes the cognitive or intentional character of an emotion is treated as dependent on, and ultimately explained by, a physical condition (p. 521).

We can easily think of a few more "opposing directions" that the concept of emotion can pull us in: Some emotions bond people together, others sunder them; some emotions are recognizable via facial expressions, others are not; some emotions disappear as soon as contrary information is heard and believed, others persist in the face of such information; some emotions have an identifiable propositional content, others have none; some emotions (like shame) are intrisically reflexive or self-referring, others (like joy) are not; some emotions are based in the most excellent, others in the most cock-eyed reasoning, while still others are based in no reasoning at all; some emotions are disruptive episodes, relatively unintegrated

[6] *Review of Metaphysics* **38** (1984): 521–546.

into the characteristic concerns and purposes and intentional actions of a person's life, while others are continuous with those leading concerns and express them; some emotions involve discernible bodily arousal, others do not; some emotions are conscious states, others are not; some emotions are pleasurable, others are painful, and perhaps still others are neither the one nor the other; some emotion types are pancultural, others are culture-specific or culture-determined; some emotions are intentional, brought on by the subject for some purpose of her own, while others are not intentional; some emotions are motivations, while others are not. Rorty points out that these "opposing" divisions within the concept of emotion do not tend to be marked by our lexicalized emotion categories ("anger," "nostalgia," "solicitude," "joy"). For example, there might be instances of anger that fall on each of the sides of most of these divisions. Perhaps this fact helps to hide from us the rampant disorder internal to the concept of emotion.

Rorty's thesis that the seeming unresolvability of the debates about the nature of emotions somehow stems from the extraordinary variety and oppositions among the phenomena that we call emotions seems plausible to me, if we add the further premise that the debating theorists base their positions on hasty generalizations from their favored ranges of cases. For example, one kind of theorist fixes on cases of emotion that have highly definite conceptual content, that respond flexibly to changes of information and reasoning, and that are highly integrated into the individual's conscious purposes and explicit worldview. Another kind of theorist fixes on cases of emotion that respond poorly or not at all to information and reasoning, have a strong component of bodily arousal, and have close analogues in beasts and babies. Both theorists then ignore the "opposite" kinds of cases as long as they can, or they authorize their theories by finding clever ways to explain away the counterexamples or assimilate them to their own paradigm, or they just deny that those are "really" emotions. Without the hasty generalizations, followed by digging in of theoretical heels, we would presumably get descriptively richer, less theoretical, monolithic or reductive accounts, ones that would be less controversial because the generalizations would be spare and cautious, always keeping a welcoming lookout for the instructive counterexample. Among people who practiced this more descriptive philosophy of emotion, there would presumably be far less of the unyielding disagreement that Rorty deplores. In making this proposal I am supposing that the concept of emotion is not internally incoherent, and that its apparent incoherence comes from the hasty generalizations of theorists.

But this is not Rorty's proposed resolution of the difficulty. Although she does not go quite so far as to say that the concept itself is incoherent, she does blame the concept at least as much as its analysts. She thinks that our current concept of an emotion is a contraption whose ill-assorted parts are accretions traceable to diverse periods of the history of philosophy in which very divergent agendas shaped the claims that were made about the emotions. If

we lack a clear view of that history, we are doomed to a conceptual muddle, because we take the concept of emotion at "face value"; that is, we treat it as though it is a single, coherent concept.

> The history of discussions of the passions does not form a smooth continuous history, which expands or narrows the class of *pathe* by following a single line of thought. Sometimes the transformations (say from Aristotelian *pathe* to Stoic *passiones*) arise from moral preoccupations concerning voluntary control; sometimes the transformations (say from Renaissance *amor* to Hobbesian passions and desires) are impelled by metaphysical and scientific preoccupations; sometimes the transformations (say from Hobbesian passions and desires to Humean and Rousseauean sentiments) have a political direction. If nothing else, this should show that *pathe*, *passiones*, affects, emotions, and sentiments do not form a natural class. Additions to that class were made on quite distinctive grounds. Before we can evaluate the competing claims of current polemical debates, before we can understand the force of their various claims, we must first trace the philosophic preoccupations in which they originated (p. 545).

Again, it is not entirely clear whether Rorty is claiming that, for example, the Stoics merely noticed and emphasized that some emotions are subject to voluntary control and had a theory about it and built further theory on it, perhaps overgeneralizing from it, or whether the Stoics *invented* voluntary control of emotions and then passed that trait of emotions (or at least of the concept of emotion) on to us. If the former is so, then it might be *interesting* to know what the Stoics said about voluntary control, but it would not be *necessary* for a contemporary conceptual analyst, as Rorty seems to suggest it is. The analyst would be looking at an emotion like anger and noticing the same feature that the Stoics exploited, namely that people can often control their anger if they have a modicum of understanding of their emotion and make some effort. Since the conceptual analyst would be doing essentially the same kind of thing the Stoic was doing, the analyst would be under no necessity to advert to what the earlier theorist had said.

If the present-day theorist really needs to know the Stoic discussion, the latter must be somehow constitutive of the very subject matter of the present discussion. In that case when Rorty says, "Before we can evaluate the competing claims of current polemical debates, . . . we must first trace the philosophic preoccupations in which they originated," she must be saying that the "opposing" features that set the parameters of our debate actually *originated* in the earlier philosophical discussions. For example, if we can't appreciate the notion that emotions are subject to voluntary control without knowing the Stoic contribution to the subject, then the fact that emotions are subject to voluntary control is not just *noticed* by the Stoics but *created* by them. Even this would not be enough, strictly speaking, to make acquaintance with historical Stoicism a necessary condition for understanding current debates because the voluntariness of emotions might take on a life of its

own after having been socially constructed in terms of Stoic theory. On this interpretation, Rorty's claim that we cannot understand emotions without history of philosophy implies that this history not only created such features of emotions as their voluntariness, reliance on judgments, power to deflect normal behavior, grounding in physiological conditions, and so on, but created these features in such a way that they are *internally tied to the originating theories.*

What kind of understanding of the concept of emotion would emerge from a study of the history of the philosophy of emotion, on the second interpretation of Rorty's thesis? Since by hypothesis our concept of emotion is socially constructed in such a way as to make conceptual–analytical accounts of it chimaeral, the result of the historical studies that Rorty envisages would be our understanding of an incoherent "concept" *as incoherent.* If we wrote the history of the concept of emotion, we would understand *emotion* to be a philosophically constructed chimaera (my dictionary says a chimera is "an imaginary monster compounded of incongruous parts"). We would see that the concept of emotion has no real referent, but only this constructed, chimaeral one. This history would explode a myth, exposing a *purported* concept for the monstrous contraption that it is.

We might wonder why, on this interpretation, the unmasking of the "concept" of emotion could not proceed ahistorically, just by showing the internal contradictions in the concept. Perhaps the idea is that this procedure would never decisively show the concept to be incoherent since a conservative could always fall back on the hope of a future account that will show the concept's coherence. The genealogy of *emotion* might be thought capable of laying this hope finally to rest, by showing once and for all where the contradictory strands in the "concept" came from.

It is not clear to me that Rorty endorses the rather implausible view that I have just sketched. Perhaps she thinks that the influence of philosophical theories on our concept of emotion is of some looser variety, and that phrases like "must first trace the philosophical preoccupations" and "necessary to trace the history" should be taken more weakly than I have done. She does make one remark that seems to make the history of philosophy less crucial:

> Officially we are preoccupied with determining whether emotions can be evaluated for their rationality; or whether they are voluntary; or whether they can be "reduced" to cognitions; or whether they are interruptions of behavior that is normally purposeful. But in fact we know better: when we are really thinking, rather than making pronouncements, we know that we evaluate the appropriateness of emotions by criteria that are much richer than those of logical consistency: we are interested in determining whether they are inadequate or excessive, crude or subtle; whether they are harmoniously balanced with one another; whether we admire the character traits they reveal and the motives that usually accompany them. And when we are careful, we usually also distinguish passions, emotions, affects, sentiments (pp. 521–522).

While I would not describe in just Rorty's terms the kind of conceptual analysis I commend, I agree with the direction of her thought in this quotation. She is saying, in effect, that if we stop crudely theorizing and look carefully at the human emotions and our modes of describing and evaluating them, if we stop thinking in terms of simplistic questions about emotions and *look* to see how they actually and richly function in the course of our lives, then the seeming incoherence in the concept of emotion begins to disappear and we see not incoherence and in principle irresolvable debates, but subtle and rich variety linked by family resemblances. So perhaps Rorty is admitting that we may not strictly *need* the history of philosophy after all, but just a more astute and careful and "empirical" and less theoretically hidebound application of philosophical analysis. But because philosophers have historically picked up on some features of emotions to the exclusion of others, the history of philosophy might help in our analysis by alerting us to features that need accommodating and abstractions we need to avoid. On this interpretation, which we might call the "history of philosophy as aid to conceptual analysis" view, Rorty would not be saying that the concept of emotion is an imaginary monster, nor that the history of philosophy is strictly necessary to its analysis. The history of the philosophy of emotions is *a useful but non-necessary adjunct* to philosophical analysis (along with several other adjunct disciplines), in heading off theoretical dead-ends, raising interesting questions, and making interesting proposals.

My purpose is not to adjudicate the interpretation of Rorty's provocative paper, but to defend a kind of conceptual analysis of the emotions. Since the second interpretation allows for conceptual analysis with a recommendation of aid from the history of philosophy, I have no quarrel with it. And I am interested in the first interpretation, not because I ascribe it with confidence to Rorty, but because it is a challenge to my project.

Let us try out an argument, which we might call the realist common sense objection, against the historically constructed chimaera theory (HCCT). As a proposal for examination why not say the following:

Proposal:
We can explain the "opposing" features of emotions much more straightforwardly. We needn't posit that the history of philosophy has created these features, because we can observe them in our everyday experience. For example, we can explain why people have thought that emotions are strongly connected with judgments by noting that people, in any historical period, including our own, can be roused to anger or fear or nostalgia by narratives, and that their anger or fear can often be dispelled instantaneously by telling them something. We needn't resort to the history of philosophy to explain why people think emotions are grounded in physical conditions such as fatigue or the influence of drugs; appeal to their experience is enough. We do not need the history of philosophy to explain why people are sometimes held responsible for their emotions and sometimes exonerated because of them. Nor do we need it to show us why people think that both normal

and abnormal human functioning depend on emotional states and dispositions. These judgments about emotion can be nicely attributed to the human experience of living. And clearly, the philosophers who built their theories on one or another of these features did so by observing them, just as we do. HCCT reverses the order of priority: the philosophers' theories came from the features, not the features from the theories. And if our attributions of these seemingly opposed features to emotions are results of observation rather than of theory construction, then we may have some confidence that they only seem opposed – that the concept of emotion is not a chimaera but a consistent body of attributions. After all, reality, even psychological reality, is not likely to be incoherent.

The weakness of this response is that on HCCT, the fact that we can observe the opposing features is not evidence that they were not created (in the strong sense required by HCCT) by the history of philosophy. As Rorty says, "All these views are embedded in our common speech and common sense, as well as in the literary works that form our understanding of ourselves" (p. 545). So the position is insulated against the common-sense realist objection. But HCCT needs to have more going for it, if we are to abandon common-sense realism for it, than that it is insulated against objections from common-sense realism. We need some positive reason to accept it, since common-sense realism is common sense. If philosophical reference to each of the features of passions that Rorty finds identified and exploited in the history of philosophy can be as well accounted for on the hypothesis that the philosopher in question identifies a previously existent feature as on the hypothesis that the philosopher invents the feature and then passes it down to us in the form of common sense, then the history of philosophy gives us no reason to accept HCCT rather than common-sense realism about emotion features. In that case we just have an evidential stand-off; and since common sense takes natural precedence, we have no reason to abandon it.

But other considerations seem to weaken further the appeal of HCCT. We might wonder where philosophers got the idea of the feature – say, that emotions are dependent on judgments or that emotions disrupt normal behavior or that emotions are necessary to fully normal behavior – if they did not get it from observation. Philosophers are typically pretty creative people, and so we might think there's no mystery here, but my guess is that if we looked at the contexts in their writings in which philosophers identify the features that have come to play roles in modern discussions of the emotions, we would see that they often appeal to examples and observations. This is certainly true of Aristotle and Hume, and I would guess that it is true of most of the main players in Rorty's history of the emotions.

Also, we might wonder why these features have had such sticking power in human life and why they are sustainable at pretty much all educational levels and with so little direct influence from the history of philosophy. We might think that where concepts are invented more or less out of whole cloth and without much of an observational basis, they require more direct

and continuous intervention from theorists than the concept of emotion seems to enjoy. Another possible argument might be launched by examining anthropologists' studies of emotions among peoples who cannot have been influenced by the history of philosophy. If such studies show the natives identifying features of emotions like the ones that generate the controversies in recent Western intellectual discussions, that would be evidence that these are observable features antedating philosophical theories that exploit them (see Sections 3.2b, 3.3b, and 3.3c).

Yet another potential argument is that if we expand the list of "opposing" features, as I did at the beginning of this section, we may begin to have a hard time finding plausible originating points for them in the history of philosophy. We may wonder why Rorty selects just four oppositions, and whether all four even of these are plausibly explained in terms of the history of philosophy. In any case, the project of showing that the "opposing" features of emotions were born in the history of philosophical discussions of the emotions has yet to be done. The hypothesis cannot be fully evaluated in the absence of a more or less full, book-length demonstration.

Let us consider the history of philosophy as an aid to conceptual analysis. Rather than think of the history of the philosophy of emotions as *constituting* or creating our concepts and experiences of emotions, we might think of this history as *influencing* them, in the course of responding to the phenomena. Different players in that history respond according to their own particular agendas and theoretical frameworks, so that they highlight different features of the emotions, which, as we have seen, do have many diverse features. The anthropologist Robert Levy has proposed that societies may "hypercognize" or "hypocognize" emotion types. For example, the Tahitians, among whom Levy did field work, hypercognize anger but hypocognize sadness. They have a subtle vocabulary for describing, explaining, evaluating, and prescribing for anger but not even a word that denotes sadness. The Tahitians do become sad, but they are less likely to notice it and do not identify it with the same precision as societies in which it is more "cognized."[7] Something similar might be true of the generic features of emotions: For theorists with differing interests, different features will be salient, and the saliencies will both influence and result from their theories; but this is not to say the features are created by the theories. Perhaps Aristotle, the Stoics, Augustine, and others did not create the features that our concept of an emotion attribute to emotions; instead, they all more or less successfully describe phenomena that have been relatively stable through human history, the same kind of thing that contemporary analytical philosophers, anthropologists, psychologists, neuroscientists, and evolutionary biologists are giving

[7] See "Emotion, Knowing, and Culture," in Richard A. Shweder and Robert A. LeVine (eds.), *Culture Theory: Essays on Mind, Self, and Emotion* (Cambridge, England: Cambridge University Press, 1984), pp. 214–237.

their accounts of. Emotions invite highly perspectival accounts because they are many-sided phenomena. On this picture we may admit possible influences from Aristotle, the Stoics, and so on, on our way of thinking about the emotions, but it would be an exaggeration to talk about the transformation of Aristotelian *pathe* into Stoic *passiones*, as though the subject matter of their discussions is not the rather old familiar facts of anger, fear, joy, and hope. Instead, we could talk about Aristotelian *ideas* about emotions and Stoic *ideas*. In that case the puzzles we experience when we study the emotions as philosophers would be not just products of this history but, more importantly, products of the phenomena – the emotions that we observe in human beings. And the supposed conflicts that we find within the Aristotelian account, or the conflicts between that account and, say, the Stoic account, would be due as much to the actual features of emotions and passions as to theorists' accounts of them. An imperialistically social constructivist account of emotions is as far from the truth about them as a purely neurological account. Each, according to its special interests, "hypercognizes" certain features.

Let us distinguish emotion category concepts from emotion type concepts. Examples of category concepts are ones that have roughly the same degree of generality as *emotion: sentiment,* πάθος*, passio, affect, affectus, passion*, and so on. Type concepts are concepts of subclasses within the categories of emotion, passion, affect, and so on. Examples are *anger, dismay, sorrow, shame,* τὸ νεμεσᾶν,[8] *liget,*[9] and so on. Emotion category concepts encompass a range of type concepts. Thus *emotion* is the class that includes anger, nostalgia, shame, joy, and perhaps (on the periphery) puzzlement, amusement (at humor), surprise, and the startle response. Because *emotion* and *passio* and πάθος encompass partially different ranges of types, they will be different, though largely overlapping, concepts. For example, Thomas Aquinas lists desire (*concupiscentia*) as a type of *passio*, whereas we would probably not regard it as a kind of emotion (though we might include it among the passions); and as Rorty makes abundantly clear in the main body of her paper, the concept of a πάθος in Aristotle's society was much broader than our concept of emotion, encompassing such things as bodily wounds and states of sense perception. I think that the studies in the history of philosophy that Rorty commends can sensitize us to the variability of the category concepts related to that of emotions (passions, sentiments, etc.) and to the relativity of such variation to human interests; they can mitigate a certain platonizing tendency in the study of emotions, a tendency that natural languages

[8] An emotion type discussed by Aristotle, different from envy (φθόνος), characterized by discomfort about someone else's undeserved good fortune. See *Rhetoric*, Book II, Chapter 9 (pp. 1386b10–1387b20).

[9] A dominant emotion in the moral psychology of the Ilongots, a head-hunting group in the Philippines (see Section 3.3b).

seem to engender. The history of philosophy and psychology is full of lists of "basic" emotions, and these lists differ remarkably from one another; the best explanation of this diversity seems to be that the lists reflect different sets of theoretical interests (see Section 3.1c). Also, the history of philosophy, like cultural anthropology, can moderate our naive tendency to think that our emotion type vocabulary divides the world of the emotions in the natural and only possible way (see Section 1.5e).

So the concept of emotion can be thought of as determined by the range of type concepts that it encompasses, but it must be admitted to be somewhat indeterminate because of questionable types on the outer fringes, such as surprise, startle, amusement (e.g., at jokes), interest (e.g., in philosophical ideas), and others. The intuitions of good speakers of English vary as to whether these states are emotions. But the bare question of English usage is not in itself a very interesting one; we want to know *why* type concepts like *anger*, *fear*, and *envy* are solidly in everybody's paradigm of emotion while *surprise* and *startle* are only in some people's. One way to get at an answer to this question will be to take seriously the various "opposing" features of the paradigm cases that Rorty's essay invites us to highlight, as well as others that I have indicated. If we can come up with a broad unifying conception of emotion that accommodates all these opposing features, then we will have a conception that unifies at least the paradigm cases and gives us a plausible account of why English speakers group this range of mental states together under a single class name. That is the main task of Chapter 2.

In Chapter 3 I will then test the conception by analyzing a wide range of type concepts, including not only the paradigm types but pretty much anything that anybody is inclined to call an emotion, including the contested types. My strategy will be not so much to try to decide whether each type belongs to the category of emotion as to try to see in what ways each type is similar, and in what ways dissimilar, to the undisputed paradigm cases. Thus I do not offer my account as a "theory," as implying that all and only what we would properly call an emotion fits the proffered conception. Instead I shall argue that the conception is superior to its competitors in making sense of all the "opposing" features in the paradigm cases. I shall try to show fairly precisely the various ways in which the other cases deviate from the paradigm. But despite the fuzziness on the edges, I think I will have shown that the concept of emotion is not a monster.

1.4. DECONSTRUCTING *EMOTION* VIA THE LIFE SCIENCES

There is a strong movement these days to subsume psychology under biology and related disciplines such as physiology and especially neurophysiology. Paul Griffiths's *What Emotions Really Are: The Problem of Psychological*

Categories,[10] is an especially explicit, philosophically sophisticated, and uncompromising example of this trend. Besides this, the book is focused on emotions and directly attacks conceptual analysis as an approach to understanding emotions. For all these reasons, it is interesting for our purposes, and I hope I will be forgiven for paying so much attention to it.

a. Science Fractures a Concept?

Griffiths's thesis is reminiscent of Rorty's proposal that we deconstruct the concept of emotion using the history of philosophy because, trading on some of the "oppositions" that are present in the ordinary concept of emotion, he proposes that under scientific study the concept of emotion will "fracture" into three radically distinct concepts. Recent science shows that the vernacular concept of emotion covers a range of things that have as little to recommend their assimilation under a single concept as the hodgepodge in Aristotle's class of superlunary objects.

> Emotion is like the category of "superlunary" objects in ancient astronomy. There is a well-defined category of "everything outside the orbit of the moon" but it turns out that superlunary objects do not have something specially in common that distinguishes them from other arbitrary collections of objects. . . . what we know about ["emotions"] suggests that there is no rich collection of generalizations about this range of phenomena that distinguishes them from other psychological phenomena. They do not constitute a single object of knowledge (p. 14).

In particular, he is impressed by the "opposition" between (a) emotions that show a clear physiological syndrome, are reflexlike, pancultural, and phylogenetically ancient, and do not require higher cognitive processing and (b) emotions that do require such processing and may be quite culturally specific and do not show any clear physiology.

> The first group he calls (following Paul Ekman) "affect program responses." . . . the affect program theory deals with a range of emotions corresponding very roughly to the occurrent instances of the English terms "surprise," "fear," "anger," "disgust," "contempt," "sadness," and "joy." The affect programs are short-term, stereotypical responses involving facial expression, autonomic nervous system arousal, and other elements. The same patterns of response occur in all cultures and homologues are found in related species. These patterns are triggered by a cognitive system which is "modular" in the sense that it does not freely exchange information with other cognitive processes (p. 8).

"Higher cognitive emotions" are divided into two discrete categories. Griffiths calls the first category "irruptive motivations," following Robert Frank.[11] These are like the affect program responses in that both kinds of

[10] Chicago: University of Chicago Press, 1997.

[11] *Passions Within Reason: The Strategic Role of the Emotions* (New York: Norton, 1988).

emotion "produce a form of passivity" (p. 245); that is, they are not intentionally produced but *come over* the subject in response to something. These states, which include instances of loyalty, jealousy, and guilt, as well as episodes that vernacular speech would identify with the same names as are used for the affect programs, are "states which interfere with the smooth unfolding of plans designed to secure our long-term goals" (p. 246). Thus they are not only *irruptive* (i.e., passive states) but also *disruptive* of long-term goal-seeking. An example of irruptive motivational anger would be the emotion of a man that drives him to take revenge on people for trespassing his rights even when taking revenge undermines his considered long-term goals (e.g., making money, keeping his friends). Frank argues that such an emotion is evolutionarily adaptive, despite first appearances, because people will be disinclined to trespass the rights of a person who is likely to go ballistic in this way. The irruptive motivations have surface irrationality that hides a deeper function. These "emotions" may occur in the absence of facial expression and autonomic arousal and do involve higher cognitive processing. In our example, the concept of a violated right, which the angry subject deploys in his response to the situation, clearly requires the functioning of "higher" parts of the brain, not just the "informationally encapsulated" ones that operate in the affect programs.

The second kind of higher cognitive "emotions" are "socially constructed." Griffiths distinguishes two kinds of social construction, the "social concept model" and the "social role model," and dismisses the former as trivial. He points out that many social constructionists in emotion theory think that a society constructs emotions by providing categories in terms of which its people respond emotionally to objects and situations. But this "is a model of the emotions themselves only because an emotion is identified with the thought that the eliciting situation is present" (p. 139; here Griffiths refers to a version of the propositional attitude theory that we will discuss in the next subsection). The kind of socially constructed emotions that fill a significant category are the ones he calls *disclaimed actions*. These are essentially fake emotions – behavioral patterns that one produces, under the guidance of cultural rules, for the sake of achieving some goal. Thus, according to Griffiths, they lack the "passivity" that he finds common to the affect programs and the irruptive motivations. Far from disrupting goal-directed behavior, these are stratagems to purpose. Griffiths hastens to point out that the subject of such an "emotion" is not merely pretending: "The subject does not have conscious access to the causes of their [*sic*] behavior and provides an erroneous explanation of their behavior that masquerades as an introspective report" (p. 158). Borrowing from Robert Solomon, he says,

> A good example is the display of anger as an unconsciously implemented "strategic behavior" in a marital quarrel. The agent has reasoned that they can improve their