THE COGNITIVE STRUCTURE OF EMOTIONS

SECOND EDITION

ANDREW ORTONY, GERALD L. CLORE
AND ALLAN COLLINS

THE COGNITIVE STRUCTURE OF EMOTIONS

More than 30 years after its initial publication, this new edition of *The Cognitive Structure of Emotions* refines and updates Ortony, Clore, and Collins's OCC model of emotions. Starting from a three-way classification of construals of the world – events, the attribution of responsibility for events, and objects – the authors propose a systematic account of emotion differentiation. Rejecting the oft-favored features of bodily feelings, emotion-related behaviors, and facial expressions as too intensity dependent and insufficiently diagnostic, they provide a detailed analysis of emotion differentiation in terms of the cognitive underpinnings of emotion types. Using numerous examples, they explain how different variables influence emotion intensity, and show how emotions can be formalized for computational purposes. Now with a contributed chapter describing the influence of the OCC model, this book will interest a wide audience in cognitive, clinical, and social psychology, as well as in artificial intelligence and affective computing, and other cognitive science disciplines.

ANDREW ORTONY is Professor Emeritus of Psychology, Education, and Computer Science at Northwestern University. He is well known for his work on emotions and for his research on knowledge representation and metaphor, and edited two editions of the landmark book *Metaphor and Thought*. His awards include a James McKeen Cattell Sabbatical award, and a Fulbright Senior Specialist award.

GERALD L. CLORE is Professor Emeritus at the University of Virginia. He is the recipient of numerous awards including the APS William James Award and the SPSP Career Contribution Award. He is a member of the American Academy of Arts and Sciences.

ALLAN COLLINS is Professor Emeritus of Learning Sciences at Northwestern University. He is a member of the National Academy of Education, and a founding editor of the journal *Cognitive Science*. His book with Richard Halverson, *Rethinking Education in the Age of Technology* is in its second edition.

THE COGNITIVE STRUCTURE OF EMOTIONS

Second Edition

ANDREW ORTONY

Northwestern University, Illinois

GERALD L. CLORE

University of Virginia

ALLAN COLLINS

Northwestern University, Illinois

To our wives,
Jelena, Judy, and Anne

Contents

Figures

Tables

Preface to the Second Edition

We started working on the first edition of *The Cognitive Structure of Emotions* in the early 1980s, inspired by our belief that the emerging discipline of cognitive science was so taken with questions of knowledge representation and natural language processing that it was ignoring what for us appeared to be the obvious and important relation between affect and cognition. For us it was clear not only that affect influenced cognition, as we had been reminded in an influential paper by Gordon Bower in 1981 that discussed the effects of mood on memory and thinking, but also that cognition played a major role in the elicitation of emotions. This fact had already influenced some psychologists around that time, perhaps most notably, Richard Lazarus and colleagues (e.g., Lazarus et al., 1970), but such efforts had not permeated what was becoming "mainstream" cognitive science. Put simply, what the first edition did was to propose a theory of emotion based on a detailed analysis of how the content of our cognitions relate to how we feel. But in the spirit of cognitive science, it did one other thing – it attempted to spell out its proposals with sufficient precision and detail that it might be possible to employ them in computational models of what might be called "emotion information processing." And, so, although a psychological theory, many computer scientists interested in building such computational models adopted it as a starting point for building working computational systems, with the book and the theory it lays out often being referred to as (the) OCC (model) after our names, Ortony, Clore, and Collins.

Now, some thirty-five years after the publication of the first edition, the worlds of technology and of artificial intelligence are almost unrecognizable, and to some extent both influence and have been influenced by developments in emotion research. There are now all manner of computing domains for which emotion models are important. For example, social and domestic robots are beginning to have at least a rudimentary capacity to recognize, understand the significance of, and respond to the emotional states of the humans with whom they interact, and the same is true of virtual characters used in training and educational applications (so-called "serious games"), not to mention the simulation of plausible emotion-driven behaviors by the characters enjoyed by the nigh on three billion individuals worldwide whose gaming activities sustain the multi-billion-

dollar gaming industry. Other areas of computer science to which some aspects of emotion theory are relevant include speech recognition and natural language processing, and the fast growing area of sentiment analysis and opinion mining. With all of these developments, it is perhaps not surprising that there has emerged a new subfield of computer science that focuses on computational methods of "understanding," simulating, and stimulating human emotion – a subfield of "affective computing" (Picard, 1997) with its own journal (*IEEE Transactions on Affective Computing*). Although to a lesser degree, there has also been influence in the other direction, with new technologies influencing emotion research. A clear indication of this is the emergence of the field of affective neuroscience (Panksepp, 1998), much of which uses sophisticated brain-imaging technologies to study the neural mechanisms underlying emotion processing. Furthermore, in addition to the increasing use of computational modeling of affective processes (Broekens et al., 2013; Reisenzein et al., 2013), the use of big data to explore various aspects of emotion is becoming more common (e.g. Cowen & Keltner, 2017; Jackson et al., 2019; Zhang et al., 2019).

At its core, OCC is a theory about the foundations of emotional experience. In this new edition we have updated our account by taking advantage of empirical developments since the book first appeared. In the first edition, we made some specific proposals about an issue that, at the time, was largely neglected by emotion theorists, namely, the question of the determinants of emotion intensity. In this new edition, we delve into the question of emotional intensity in much greater detail. A second important respect in which we have updated our original conceptions of the cognitive foundations of emotions is to integrate them with recent conceptual developments in emotion theory. The last thirty years have seen pressure (albeit often resisted) to abandon traditional, essentialist, ideas about emotion, in favor of a constructivist view (Gendron & L. F. Barrett, 2009; Russell, 2003) on which we believe that our own work has had a material influence. With this in mind, we have rejected our early way of thinking in terms of the *eliciting conditions* of emotions, with its implication that emotions are "things" waiting to be elicited when conditions are right, and replaced it with the notion of *emergence conditions*, which view emotions as being more akin to thoughts – emergent phenomena that are constructed, often unwittingly.

Other major changes, apart from significantly updating the literature that we cite (we have incorporated over 250 new references), include more detailed discussions of the relation between language and emotions, and a

refinement of our proposed formalisms of a number of emotion types, and an increase in the number of emotion types that we characterize from twenty-two to twenty-four (a change that demonstrates that the theory can be readily modified). Finally, we were fortunate enough to be able to entice our renowned, computationally oriented, colleagues Jonathan Gratch and Stacy Marsella to write a final, contributed chapter reviewing the relation between the development of the field of affective computing and our own work. They, together with their students, are responsible for some of the most impressive computational models of emotion and coping processes to have been developed to date. It is hard to imagine two more appropriate people to undertake such a task, and we are most grateful to them.

Finally, we, the authors are often more pedantic than is good for us, a shared trait that surely made Irene Pizzie's task of copy-editing our manuscript quite a challenge. Thus to her, for her dedication, professionalism, and cheerful willingness to tolerate out foibles, we owe a special debt of gratitude. We would also like to thank Stephen Acerra, our editor at Cambridge University Press, for encouraging us to embark on this enterprise in the first place, and then overseeing it to the end. But the greatest contribution to this work has been made by our spouses who quietly and with ne'er a complaint accepted sometimes hours of neglect as we obsessively worked away, morning, noon, and night. Words cannot adequately express our gratitude to them. Nevertheless: Jelena Radulovic, Judy DeLoache, and Anne Collins, we thank you!

Preface to the First Edition

As cognitive psychology established itself in the 1970s, it became increasingly apparent that it was a "cold" approach to cognition, and doubts began to arise as to whether or not it could provide the machinery necessary to account for affect and emotion. In 1981, Donald Norman identified the topic of emotion as one of twelve major challenges to cognitive science (D. Norman, 1981). It was at about this time that the three authors of this volume decided to collaborate in an attempt to explore the extent to which cognitive psychology could provide a viable foundation for the analysis of emotions. Certainly, it was no problem for cognitive psychology, with the help of schema theory, to explain such facts as that the same thing can be perceived from different perspectives. This was already encouraging, because the capacity to view a situation from different perspectives struck us as lying at the heart of the fact that different people often experience different emotions in response to the same objective event.

Many emotion theorists have argued that cognitive appraisal is central to emotion, yet no one has been able to say anything much more detailed than that. This book is an attempt to give at least the outlines of an account of how such appraisals are made. In it we present many detailed observations about specific emotions, their organization, and the specific cognitive processes involved in their elicitation, but we would be satisfied if this effort succeeded in demonstrating only that a systematic and comprehensive account of the cognitive antecedents of the emotions is possible. Our goal is to convince our readers that such an approach is viable rather than that our particular version of such an effort is the correct one. We have, in fact, chosen a somewhat arbitrary stopping point for this enterprise. Further use of the same structural principles that we propose would allow one to continue to specify increasingly differentiated sets of emotional states. The more one does this, however, the more one becomes tied to the emotional system associated with a particular cultural view of the world. This, in turn, increases the risk that one will lose sight of the main agenda, which is to characterize the range of "psychological" possibilities for emotions rather than to describe the emotions and emotion-related processes local to any specific time or cultural group.

We started collaborating on this project in the spring of 1980 when we began talking to each other about various emotions and the conditions of their occurrence. In the context of our common fascination with this problem, the differences in our backgrounds and interests made the end product different than it otherwise would have been. As a group we include a cognitive scientist interested in psychological, linguistic, and computational aspects of the study of mental processes, a social psychologist with interests in personality and in the influences of affect on social judgment, and a cognitive psychologist with interests in the formal modeling of human reasoning processes. As we started working out some of the ideas into a more concrete form it became clear that we were going to end up with a book – a book that we began to think of affectionately as *Principia Pathematica*. We saw ourselves as attempting to characterize some of the key principles governing the cognitive mechanisms underlying human emotions, so that title seemed to us apt and even a little humorous. After all, we tried to persuade ourselves, the *Oxford English Dictionary* contains the following (abbreviated) entry:

> PATHEMATIC, *a. rare* [ad. Gr. *pathematicos* liable to passions or emotions, f. *pathema* what one suffers, suffering emotion, f. stem *path-:* see PATHETIC] Pertaining to the passions or emotions; caused or characterized by emotion.

However, many of our friends and colleagues were skeptical. The title we proposed would be incomprehensible to those lacking a classical education, they argued, and the book might well end up in the medical section of bookstores! Then again, there was the problem of hubris – was it not a little pretentious? We were eventually persuaded that discretion is the better part of valor, and settled on a title that, while maybe lacking something in panache, at least has the virtue of truth in advertising.

Introduction

Emotions have many facets. They involve feelings and experience, they involve physiology and behavior, and they involve cognitions and conceptualizations. There are important questions that can be asked about the expression of emotions (e.g., Fernández-Dols & Russell, 2017; Keltner et al., 2019), and the language of emotion constitutes an interesting research domain in its own right (e.g., Lindquist et al., 2015; Wierzbicka, 1999). But in this book, although we will have quite a bit to say about, especially, the relation between language and emotions, our main concern will be with the contribution that cognition makes to emotion. In particular, we will focus on the role of the system of cognitive representations – the value system – that leads people to appraise the situations in which they find themselves as good or bad, as beneficial or harmful, or, more generally, as positive or negative, that is, valenced. The value system, which we take to comprise three classes of cognitive representations – (the current state of) a person's *goals*, *standards*, and *tastes* – is central to the theory that we propose. Broadly speaking, goals (i.e., representations of desired states of the world) are the representations in terms of which *events* are evaluated. Standards are representations of points of reference or criteria, often but not necessarily moral, in terms of which *actions and behavior* are judged, and tastes are the representations of our dispositional likes and dislikes on the basis of which *objects* (in a very general sense) are evaluated. Taken as a whole, these representations are the considerations in terms of which we appraise what is going on in or around us. They comprise the bedrock of the value system, and everything that we evaluate, we evaluate in terms of one or more of these three fundamental kinds of representation.

The overarching theme of this book concerns the general issue of emotion differentiation – the question of what distinguishes one emotion from another – and our approach will be to articulate the differences between emotions in terms of the different kinds of cognitions that

underlie them. Our decision to focus on cognitions as the most fruitful way to approach this question is not arbitrary. Many other aspects of emotions have attracted the attention of emotion theorists, starting with William James, who in his famous paper "What is an emotion?" argued that the feeling of bodily changes associated with the "perception of the exciting fact" *is* the emotion (James, 1884). Furthermore, much attention has been paid to the role of behaviors, often discussed as actions or action tendencies (e.g., Frijda, 1986), and the so-called "facial expression" of emotions has been the beneficiary of a great deal of research over the last fifty years (see Fernández-Dols and Russell [2017] for a review). There is, however, a problem with relying on behaviors, recognizable facial expressions, or awareness of bodily changes as criteria for distinguishing between different types of emotions: even if there were a one-to-one correspondence between the presence of one such feature and a particular emotion type, which there is not, such features are generally only evident in intense emotions; they are rarely evident in mild emotions. Thus, a reliance on them is not a good way to distinguish different types of emotions in general. Yes, maybe one can distinguish intense anger from intense fear by the presence of an uncomfortable warm feeling in the former but not the latter. But such an uncomfortable warm feeling might also be present in the case of embarrassment and will probably not be present at all in cases of mild irritation. Thus, it is not clear what, if anything, such an uncomfortable warm feeling signifies. On the other hand, we can always distinguish a construal of the world in which we blame someone for harming us from one in which we envision an undesirable event, or from one grounded in the belief that we have violated some sort of social norm. These kinds of construals of one's world are always present and always provide a reliable basis – a cognitive basis – for distinguishing one kind of emotion from another, irrespective of intensity and irrespective of what one does or of what's going on in one's body or in one's face.

Our theoretical starting point is the assumption that emotions are reflections of the ways in which people construe situations of concern to them, our general plan being to try to impose some structure on the limitless number of such construals. We view our undertaking as an exercise in theoretical psychology, even though our proposals are intended to be consistent with existing data, as well as making their own empirically testable predictions. What we are trying to do, therefore, is to specify the "psychological" structure of emotions in terms of personal and interpersonal situation descriptions. Given this goal, we make no attempt to review the massive literature on the diverse aspects of the psychology of emotion.

Such a review would only dilute our own efforts. Furthermore, we think it improbable that we could improve on the excellent coverage that already exists (e.g., L. F. Barrett et al., 2016; Strongman, 2003). Accordingly, we do not devote attention to the physiological, neuropsychological, behavioral, or expressive aspects of emotion. This neglect is not because we think them unimportant, but because we do not consider them central to our main goal of laying out a way of thinking about how emotions can be differentiated from one another in terms of their underlying cognitions.

One hundred years ago, the French writer Georges Polti published a short book in which he attempted to reduce the history of drama to thirty-six plot situations, believing them to represent thirty-six different emotions (Polti, 1921). Although this may seem like a curious enterprise, the view that we espouse treats emotions, as did Polti, as embodiments of the kinds of situations that characterize the human condition. As in the history of drama or opera, so too in people's emotional lives, recurrent themes appear, some of which we attempt to characterize in the pages of this book. It is true that our focus is on the cognitive structure of emotions, even though a common starting point in treatments of emotion is to cast emotion and cognition in opposition. Upon analysis, however, it becomes evident that the major differences among different types of emotions are largely a reflection of people's cognitions about and understanding of the various kinds of situations in which they find themselves. Hence, we characterize different emotions in terms of psychological situations, in the belief that such an approach both simplifies and clarifies what has historically seemed complex and mysterious.

One way to illustrate the utility of taking this perspective is to consider Herbert Simon's (1969) example in which he asks readers to account for the complex path taken by an ant as it moves along a wind- and wave-molded beach, going around pebbles and angling up and down the sides of tiny dunes and other obstacles it encounters. Why is the ant's path not a straight line from where it starts to its goal? Are ants complex? Do they have complicated strategies for foraging that cause them to choose the curious paths that they do? In asking these rhetorical questions, Simon is emphasizing the importance of keeping in mind that the complexity guiding the ant's behavior lies not in the ant, but in the environment it is traversing.

The model of emotions that we propose in this book carries a similar lesson about the complexity of emotions. People think of emotions as complicated, which of course they are, but much of that complication is a reflection of the complexities of the psychological environments that

distinguish different emotions and from which they emerge. Consider, as an analogy, how best to explain what is going on when cat owners appear to understand their cat's meows as "I want to go outside" or "I'm hungry." Research finds that in fact owners depend less on what sounds their cat makes than they do on where their cat is when it meows – near the door or near the food bowl. Of course, for emotions, the key is not the actual environment, but how it is construed by the individual and what aspect becomes the focus of attention. The focus is important because affective reactions are generally interpreted as being about whatever is in focus at the time. Thus, the same positive affect that is experienced as "liking" when one focuses on a person or some other object of judgment, can be experienced as "confidence" in that judgment when one focuses not on the object of judgment but on the process of making the judgment (Clore & Schiller, 2016).

Similar logic applies to the differences among emotions. What, for example, differentiates emotions that are commonly referred to by such terms as "fear," "sadness," and "shame"? All are negative, but fear is negative affect with a focus on the *possibility of a bad event*; sadness is negative affect with a focus on (a certain kind) of *bad event that is believed to have occurred*; and shame is negative affect with a focus on *one's own blameworthy action*. The point is that, in the final analysis, different kinds of emotions differ from each other because they are reflections of different focal situations. But even simple differences can have complex consequences. For example, the fact that emotional episodes of the same general type can differ with respect to their observable aspects – behavioral, expressive, and physiological – means, as already suggested, that such features cannot be taken as necessary for the existence or nature of any particular emotion. Whether they are evident depends on both the particular context and on the intensity of the reaction. People's hearts don't always beat faster when they're afraid, and sometimes they beat faster when they aren't afraid. People don't always smile when they're happy, but sometimes they smile when they aren't happy, and people don't always blush when they're embarrassed, but sometimes they blush when they aren't embarrassed. In contrast to such equivocal aspects of internal states, as explained above, people who are afraid always envisage something bad happening, people who feel happy are always reacting to something they believe to be good, and people who are embarrassed always believe that others will view them as having violated some sort of social norm. These are the kinds of emotion-related complex-ities concerning people's cognitive representations of the situations in which they find themselves that we seek to elucidate.

In the chapters that follow, we will have much to say about language and emotion. However, the theory we propose is decidedly *not* a theory about emotion *words*. This point is so important that we will emphasize it repeatedly. In fact, our characterizations of emotions are intentionally cast in terms that are as independent of emotion words as possible. This is important partly because the structure of the emotion lexicon is not isomorphic with the structure of emotions themselves, and partly because a theory about emotions must concern the kinds of things to which emotion words refer – emotions and emotion concepts – not words.

The power of everyday emotion words lies, in large part, in the access they provide to the shared schemas or concepts that underlie them. Schemas are structured representations of complex concepts that contain not only the relations between constituent sub-concepts, but also mental slots representing typical kinds of causes, characteristics, and consequences, which shape relevant experiences and memories (Rumelhart & Ortony, 1977). Thus, for example, the grief schema includes the sub-concepts of sadness and weeping, with the former being the usual cause of the latter, and it contains a range of high values for the intensity variable associated with the sadness sub-concept. Common English words, such as "anger," "fear," "love," "sadness," "jealousy," and so on, provide access to widely shared schemas of emotions, and these schemas do a great deal of cognitive work, shaping how people understand and communicate their experience. If, for example, friends describe their distress concerning some situation, we are likely to feel that we understand them much better if we label their apparent feelings as "embarrassment," "hurt feelings," or some other specific emotion. Doing so allows us to draw on the wealth of knowledge and beliefs that such schemas embody. For the same reason, watershed moments sometimes occur in the course of psychotherapy when inchoate feelings are labelled as "anger," "disappointment," "guilt," or some other specific emotion, enabling patients to better understand their situations because of the inferential potential that the associated schemas provide. So powerful are such schemas that we suspect that much of what we appear to have learned in psychology about the nature and influence of emotions may, in reality, be about the nature and influence of emotion schemas, which is what critics have in mind when they complain that much of emotion theory is really only about "folk concepts" (e.g., Russell, 2003). The concern that in our theorizing we may too often be guided by folk beliefs and concepts about emotions rather than by the experiences of emotions themselves is another reason for our insistence on characterizing the twenty-four distinct emotion types that we shall discuss, not in terms

of powerful but over-stuffed everyday labels, but in terms that are closer to their actual emergence conditions – *a positive feeling about a desirable event*, rather than "joy," *a negative feeling about an envisaged undesirable event*, rather than "fear," and so on.

Having emphasized that we are not attempting to define emotion words, we should also emphasize that we think our account is, in principle, capable of accommodating the fact that there are significant individual and cultural differences in the experience of emotions. Our claims about the structure of individual emotions are always along the lines that *if* an individual conceptualizes a situation in a certain kind of way, *then* the potential for a particular type of emotion exists. However, we do not attempt to specify the mechanisms that determine whether some particular situation will be conceptualized in one way or another. The question of how a situation is conceptualized in the first place, which we take to be a big part of individual and cultural differences, is a problem that is general for cognitive, social, and cultural psychology, not one that is specific to the study of emotions.

By attempting to be very specific about what emotions there are or could be, and about the particular factors that influence their intensity, we hope to achieve two goals. First, we seek to bring some semblance of order to what remains a confused and confusing field of study. Second, our characterizations of emotions are intended to lay the foundation for a computationally tractable model of emotion – a model that can form the basis of applications in the general domain of affective computing. Both goals require a level of specificity that is not always evident in comparable work in the field. Finally, a full account – even if limited to a full cognitive account – of how some particular emotional state arises requires an understanding of the processes involved in the generation of emotions as well as a characterization of the representations on which those processes operate. We, however, are restricting ourselves to the latter, and freely acknowledge that even if we are successful in articulating a plausible characterization of the cognitive structure of emotions, the processing mechanisms required to generate emotions still have to be articulated – an issue that is equally important if one is interested in computational modeling, and one that is discussed in more detail in Chapter 10.

The Study of Emotion

Emotion is one of the most central and pervasive aspects of human experience. Most people experience a wide range of emotions, from the quiet satisfaction of completing a relatively mundane task, to the grief at

the death of a loved one. Yet, while emotions color, deepen, and enrich human experience, they can also cause dramatic disruptions in judgment and performance. Such disruptions can have profound and sometimes terrible consequences for individuals and society as, for example, in mental illness, suicides, and crimes of passion (e.g., Ben-Ze'ev & Goussinsky, 2008). This fact is clearly recognized by creators of literature, which thrives on the imagined emotions of its characters. The basic recipe is simple: the writer describes a situation that readers recognize as being *important* to a character in the sense that it has important implications with respect to the goals, standards, or tastes that the character is known or assumed to have. Then, the character is portrayed as correctly or incorrectly construing the situation as good or bad relative to these goals or standards or tastes. Finally, the construal may occasion some sort of change in the character's judgment or *behavior*. Consider, for example, the main plot of *Othello*. We start with the assumption that the maintenance of Desdemona's love and fidelity is important for Othello. He then (incorrectly) construes Cassio's (presumed) actions as a threat to this goal and ends up consumed with anger and jealousy. Dramatic deterioration in judgment and a correspondingly drastic action then ensue in which he kills both Desdemona and himself. A certain suspension of disbelief is required of readers, but only up to a point. The essential ingredients have to be believable. If literature is a microcosm of the real world, it has to be recognizable as such.

It is apparent that writers can reliably produce in readers an awareness of a character's affective states by describing a situation likely to give rise to them. This suggests that writers use an implicit theory that individual emotions can be specified in terms of personal or interpersonal situational descriptions that are sufficient to produce them. Thus, writers do not always need to, and indeed sometimes may be unable to, specify exactly what emotion a character is experiencing, knowing instead that providing an appropriate and compelling situational description will enable readers to infer the emotion. Indeed, Anna Wierzbicka describes how Tolstoy, in writing *Anna Kerinina*, deals with his problem of not being able to find the right emotion word: "[His] solution is to describe some characteristic situation and state: *he felt as one would feel in a situation like that*. Tolstoj trusts that we are able to imagine what one would feel in that situation; and if someone cannot imagine this, Tolstoj shrugs his shoulders: *I am sorry, I cannot describe these feelings in any other way, if you don't understand, I can't do anything about it*" (Wierzbicka, 1973, p. 501). The fact that millions of readers, often over decades or even centuries, generally infer similar emotions from the described situations suggests that they share the

same implicit theory about the relationship between situational construals and experienced emotions.

For an emotion such as distress to emerge, the distressing event must be construed as being somehow undesirable, and because construing the world is a cognitive process, emotions always have a cognitive foundation – the construal. Perhaps one of the most obvious cases of the major contribution that cognition through construals makes to emotion is afforded by the reactions of players and fans at sports events. When one observes the reactions of the players to the outcome of an important game (for example, in soccer, the final of the World Cup) we see that those on (or supporting) the winning team are elated while those on the losing team are devastated. Yet, in a very real sense, both the winners and losers are reacting to the same objective event. It is their *construals* of the event that are different, and the construal of an event – the understanding of it – as desirable or undesirable, laudable or shameful, and so on, is the result of a cognitive appraisal process operating on goals, standards, and tastes. It is this understanding of the external reality, rather than reality itself, that is key, and it is in this sense that we claim that there is an essential and profound cognitive component of emotions in terms of which emotions can be differentiated.

Before we start, it is important to make clear that some emotions (e.g., disgust) involve much less cognitive processing and structure than others (e.g., shame). Interestingly, however, those that involve relatively little cognition usually have metaphorical analogs that involve much more, whereas the converse is not true. Thus, the emotion that one might experience regarding some totally inappropriate and unacceptable social behavior might well be called (moral) "disgust." On the other hand, it is difficult to imagine a cognitively rich emotion such as embarrassment having a cognitively impoverished analog. We should also make clear at the outset that our claim that emotions always involve some degree of cognition (Clore & Ortony, 2000) is not the same as asserting that the contribution of cognition is necessarily *conscious*. To say that emotions involve cognition, then, is to say that their character depends on the structure, content, and organization of knowledge representations and the processes that operate on them. These representations and processes might sometimes be available to consciousness, but there is no reason to suppose that they necessarily are so.

Our claim that some emotions involve more cognition than others has a parallel in William James, who restricted his discussion to emotions having "a distinct bodily expression" such as "surprise, curiosity, rapture, fear,

anger, lust, greed, and the like" (James, 1884, p. 189). Such states, James called the "standard emotions," which he seems to have considered as involving little or no cognition, arguing that "in advance of all experience of elephants no child can but be frightened if he suddenly finds one trumpeting and charging upon him" (p. 191). However, James acknowledged that there can be more complex emotion-inducing perceptions, ones which, in modern terms, would have to be described as involving a relatively high degree of cognition, such as events having to do with the violation of social conventions: "Most occasions of shame and many insults are purely conventional, and vary with the social environment." In considering these as potential counterexamples to his theory, James goes on, rhetorically: "In these cases, at least, it would seem that the ideas of shame, desire, regret, etc., must first have been attached by education and association to these conventional objects before the bodily changes could possibly be awakened" (p. 195). James's answer to this apparent threat to his theory is to assert that the nature of the emotion-inducing perception is not the issue; rather, the issue is that, once triggered, the perception gives rise to the bodily response that is the basis of the emotion. However, like it or not (and James is now in no position to object), James had essentially characterized a range of cognitive content for the emotion-producing perception from low (e.g., a mother's delight at the sight of her beautiful baby) to high (e.g., the delight of receiving a national honor).

The theory we propose is a theory about cognitive structures rather than cognitive processes. Accordingly, we focus on the content of the mental representations involved in different kinds of emotions rather than on the processes involved in the emergence of emotions (for examples of which, see Scherer and Moors [2019]). Furthermore, we need to make clear that when we speak of cognition in this context we are referring to a range of mental content that includes passive, tacit, knowledge-based apperceptions as well as cognitions in the sense of consciously accessible, explicit thoughts. Cognition has to do with both explicit and implicit knowledge. In fact, many, if not most, cases of emotion involve no prior "thinking" at all, but they all involve access to some kind of knowledge.

In 1988 when the first edition of this book appeared, we noted that "Modern theories of cognition have relatively little to say in the way of specific proposals about affect and emotion." At the time, terms such as "affect," "emotion," and "feeling" were rarely if ever to be found in major works on cognition. In fact, a six-volume handbook (Estes, 1975–1978) summarizing then current cognitive research had but a single reference to "affect." Emotion theorists, on the other hand, were quite concerned with

cognition and emotion. The decade had begun with a debate in the pages of the *American Psychologist* between psychologists Robert Zajonc and Richard Lazarus on the role of cognition in emotion. Zajonc (1980) argued for the primacy of affect over cognition, proclaiming that "Preferences need no inferences." He proposed that whereas psychologists generally assumed affect and emotion to be post-cognitive reactions, in fact, feelings often come before cognition. His evidence included the observation that mere exposure to stimuli created liking for them, so that people indicated a preference for the stimuli they had seen more often even when they were unable to remember which stimuli they had seen before and which were new. Such results led him to conclude that cognition and emotion are partially independent systems.

Here we need a brief digression: while many emotion theorists do not make a sharp distinction between affect and emotion, we view the two terms as embodying an important difference. For us, affect pertains to anything involving (psychological) value. An affective state need not be about anything; it can just be a general valenced (i.e., positive or negative) feeling. Emotions, on the other hand, are always about something; they are cognitively elaborated affective states, and their "aboutness" is determined by their cognitive content. Throughout this book, we will often speak of "undifferentiated affect," by which we mean states of feeling that become emotions as they acquire specific cognitive content. This distinction, while of considerable importance for us, did not play an explicit role in the debate between Lazarus and Zajonc. Perhaps it should have!

In a rebuttal of Zajonc's claims, Lazarus (1982) argued that Zajonc had erroneously assumed that cognitive appraisals must occur in fixed stages and that cognition was necessarily deliberate, rational, and conscious. Lazarus countered in support of the essential role of cognition in emotion. Such arguments persisted for years in the literature. In the 1990's, Joseph LeDoux (1996) proposed a "low route to emotion," finding that visual information could go directly from the sensory thalamus to the amygdala (the assumed seat of fear) before the target of fear could be known by the visual cortex. He concluded from his animal research that fearful avoidance could occur before knowledge of what had been seen, thereby reinvigorating the debate about the primacy of affect.

The affect–cognition primacy issue is not of primary importance to the current project, and indeed the field as a whole has moved on from this question. Rather than answering it, subsequent work has reframed the issue. For example, one model accommodates both positions by adding a temporal dimension to show how early unconscious affect is iteratively

reprocessed to become a cognitively elaborated, consciously experienced emotion (Cunningham & Zelazo, 2007). Subsequently, other aspects of the relationship between cognition and emotion have become popular research topics (for reviews, see De Houwer and Hermans [2010], Power and Dalgleish [2016], and Robinson et al. [2013]). However, the lion's share of this research has focused on how emotion influences cognition rather than on the role of cognition in emotion, with little attention being paid to the question of primary interest to us, namely, how different emotions can be distinguished in terms of the different cognitions that underlie them.

Types of Evidence for Theories of Emotion

There are five kinds of information to which one might appeal in attempting to understand emotions. First, there is the *language* of emotions, which comes replete with polysemy and ambiguity, synonymy (or near synonymy), and an abundance of lexical gaps and linguistic traps (A. P. Fiske, 2020). Of course, emotions are not themselves linguistic things, but the most readily available non-phenomenal access we have to them is through language. Thus, in order to specify the domain of a theory of emotion, it is difficult to avoid using natural language words and expressions that refer to emotions. However, a theory of emotion must not be confused with a theory of the language of emotion. Considerable care needs to be taken in the use to which natural language is put in developing a theory of emotions. Not all distinct emotion types necessarily have associated words in any particular language, and not all the emotion words in some particular language necessarily refer to distinct emotions. The absence of a word in one language to designate the particular emotion that might be referred to by a word in another does not mean that people in cultures using the first language cannot and do not experience that emotion (Lomas, 2016; Wierzbicka, 1986, 1999). Such linguistic gaps are sometimes filled through the use of metaphorical descriptions of emotions. Interestingly, this often happens even when the language does provide a word for the emotion in question, especially when a speaker seeks to communicate the particular *quality* of the emotion. For example, a respondent in a study reported in Fainsilber and Ortony (1987) characterized a feeling of increasing anger by saying that it felt like a storm was brewing inside. For some categories of emotions, a language such as English provides a relatively large number of linguistic tokens, thus reducing the need for metaphorical descriptions of emotional quality. In such cases, it

becomes necessary to identify one of the words in the category as the unmarked form or category label. For example, *fear* has lexical realizations that mark special cases such as very strong fear ("terrified"), very weak fear ("worried"), typical fear behaviors ("cowering"), and so on. Thus, it may be helpful to think of the word "fear" as a relatively neutral word for an emotion type, fear (Frijda et al., 1992). In other words, one can view the word "fear" as designating a distinct emotion type, whereas the word "terrified" does not, which is consistent with the fact that one can speak of being extremely or slightly afraid, but not of being extremely or slightly terrified. However, our ultimate goal is not to *define* emotion words such as "fear" but to specify, in as language-neutral a manner as possible, the characteristics of distinct emotions. Language, therefore, is a source of information about emotions that has to be used with considerable care. We will assume that the words in our common language reflect a number of important distinctions, that they reflect a number of not so important distinctions, and that sometimes they altogether fail to reflect important distinctions. Some of these issues will be discussed in Chapter 9. We have also discussed them at length elsewhere (e.g., Clore et al., 1987; Ortony & Clore, 1981; Ortony et al., 1987), as has Mees (1985).

The second source of information about emotions comes from *self-reports* of experienced emotions. There is as yet no known objective measure that can conclusively establish that a person is experiencing some particular emotion, just as there is no known way of establishing that a person is experiencing some particular color. Nevertheless, one might think that, because emotions are subjective experiences, people have direct access to them, so that if a person is experiencing fear, for example, that person cannot be mistaken about the fact that they are experiencing fear. However, the question is not whether people are really experiencing the emotions they believe they are experiencing; one can safely assume that most of the time they are. The question is what we are to infer from their reports about their emotions, and to this the answer is that the validity of such reports depends in large part on the kind of information that is accessed when the report is made. There are at least four kinds of information that can play a role (Robinson & Clore, 2002): experiential knowledge (i.e., direct access to the experience), episodic memory (i.e., specific past experiences of emotions), situation-specific beliefs (generic beliefs about what typically happens under certain kinds of conditions), and identity-related beliefs (the person's beliefs about their emotions in general). Which of these four kinds of information is used in generating a response depends on the conditions under which the person is reporting