Evolution, Emotions, and Emotional Disorders

Randolph M. Nesse and Phoebe C. Ellsworth
University of Michigan

Emotions research is now routinely grounded in evolution, but explicit evolutionary analyses of emotions remain rare. This article considers the implications of natural selection for several classic questions about emotions and emotional disorders. Emotions are special modes of operation shaped by natural selection. They adjust multiple response parameters in ways that have increased fitness in adaptively challenging situations that recurred over the course of evolution. They are valenced because selection shapes special processes for situations that have influenced fitness in the past. In situations that decrease fitness, negative emotions are useful and positive emotions are harmful. Selection has partially differentiated subtypes of emotions from generic precursor states to deal with specialized situations. This has resulted in untidy emotions that blur into each other on dozens of dimensions, rendering the quest for simple categorically distinct emotions futile. Selection has shaped flexible mechanisms that control the expression of emotions on the basis of an individual’s appraisal of the meaning of events for his or her ability to reach personal goals. The prevalence of emotional disorders can be attributed to several evolutionary factors.

Keywords: emotions, natural selection, evolution, adaptation, appraisal

Happenstance events can shift the subsequent history of life. If ancestors of the hippopotamus had not browsed vegetation in ever-deeper water 50 million years ago, there would be no whales today (Gingerich, Raza, Arif, Anwar, & Zhou, 1994). While such events are rarely predictable, they can be reliably explained by Charles Darwin’s and Alfred Russel Wallace’s great idea, natural selection. Four-legged whale ancestors that could swim better underwater had more offspring; over thousands of generations, their descendants gradually became superb aquatic athletes.

Happenstance occurs in intellectual evolution as well. After finishing The Descent of Man, and Selection in Relation to Sex (Darwin, 1871), Darwin realized that materials he had long collected on emotions could be organized to refute Charles Bell’s earlier claim that the elaborate musculature of the human face was evidence of Divine design. He quickly wrote The Expression of the Emotions in Man and Animals (Darwin, 1872/1965), emphasizing the phylogenetic consistency of emotional expressions from animals to humans. The book is, as advertised, about expression, and it says little about the selective forces that produced emotions, leaving a persisting anti-Darwinian legacy for emotions research (Fridlund, 1992).

However, Darwin clearly recognized that evolution shaped not only the physical characteristics of an organism but also its mental processes and behavioral repertoires. The knowledge that natural selection shaped the brain mechanisms that mediate motivation and emotions offers a solid foundation on which a modern theory of emotions is being built.

Although current psychological theories of emotion differ widely in many particulars, almost all now agree that emotions are adaptive responses that arise from mechanisms shaped by selection (Plutchik, 2003). It is now hard to imagine that just four decades ago emotions were generally seen as products of learning unrelated to natural selection. It took Ekman’s, Izard’s and Eibl-Eibesfeldt’s studies of cross-cultural consistency in emotional expression to overthrow that view (Eibl-Eibesfeldt, 1983; Ekman & Davidson, 1994; Izard, 1991). In retrospect, it is obvious that learning cannot be the whole story and that emotions would not exist unless they were useful. Evolution is not an alternative to other theories of emotions; it is the common foundation for all. Many of its contributions are so simple that they are not always recognized. To highlight the continuing importance of Darwin’s theory of natural selection for emotions, we consider its implications for several classic questions.

What Emotions Are

Definitions of emotions typically describe proximate aspects such as physiology, subjective experience, or facial expression, often emphasizing one or another component (Ekman & Davidson, 1994; Izard, 2007). An evolutionary approach defines what emotions are in terms of how they came to exist. Emotions are modes of functioning, shaped by natural selection, that coordinate physiological, cognitive, motivational, behavioral, and subjective responses in patterns that increase the ability to meet the adaptive challenges of situations that have recurred over evolutionary time (Nesse, 1990). They are adaptations that are useful only in certain situations (Underwood, 1954). Like pain and sweating, they remain latent until an evolved mechanism detects cues associated with the situation in which they are advantageous.

Randolph M. Nesse and Phoebe C. Ellsworth, Department of Psychology, University of Michigan.

Correspondence concerning this article should be addressed to Randolph M. Nesse, Department of Psychology, University of Michigan, East Hall, Room 3018, 530 Church Street, Ann Arbor, MI 48109-1043. E-mail: nesse@umich.edu
Unlike simpler adaptations, however, emotions are not unimodal responses to specific situations, like sweating in response to overheating. Instead, emotions adjust multiple component processes to create an organized response to the adaptive challenges of a given situation. For instance, appraisals that indicate a nearby predator arouse an emergency response that adjusts and coordinates many aspects of physiology and behavior. Physiology may influence cognition, and cognition may influence feeling, which may influence behavior and physiology in a complex, recursive sequence.

Neuroscience and psychological investigations of emotions focus almost exclusively on proximate questions about (a) what a trait is like and how it works and (b) how it develops over the course of an individual’s life. However, a proximate explanation is only half the story (Mayr, 1988; Tinbergen, 1963). The other half of a complete explanation requires answers to evolutionary questions about (c) how the trait developed over time in the history of the species and (d) what evolutionary factors shaped the trait. Taken together, these are Tinbergen’s four questions, the undisputed foundation for all research in animal behavior and behavioral ecology (Dewsbury, 1999; Tinbergen, 1963). Pursuing all four together will speed progress in research on emotions (Alessi, 1992; Ketelaar & Clore, 1997; Nesse, 1999).

What Different Emotions Exist?

How many emotions exist, and what are they? This question has been a source of enduring controversy (Ekman, 1992a; Ekman & Davidson, 1994; Oatley, Keltner, & Jenkins, 2006; Plutchik, 2003). Some theories postulate just two basic states—positive and negative; others postulate a small set of “basic” emotions; and still others argue for a potentially infinite number. All theorists agree, however, that valence is a necessary quality of emotions: Emotions are about pleasure and pain, approach and avoidance (Barrett, 2006b; Ekman, 1992b; Rolls, 2005).

Many one-celled organisms can do only two things—keep swimming in the same direction or tumble randomly before setting off again. In combination with a 0.5-s memory, this allows movement toward food (Adler, 1975; Koshland, 1980). The algorithm is simple: If the food concentration is higher than it was a half second ago, move forward; otherwise, tumble. The ability to detect danger, such as excessive heat or acid, shaped the other primal behavior—escape. Many bacteria can swim only at one speed, but in most organisms, valence can also vary in intensity.

Valence and intensity are essential features of almost all theories of emotions (Smith & Ellsworth, 1985). Valence is at the very root of emotion and motivated behavior (Barrett, 2006b; Schlosberg, 1952; Wundt, 1897), defining an opposition that has been described as approach/avoidance, positive/negative affect (Huppert & Whittington, 2003; Schlosberg, 1952; Tellegen, Watson, & Clark, 1999), promotion/prevention (Higgins, 1997), and the behavioral approach system (BAS)/behavioral inhibition system (BIS) (Gray, 1987). This opposition has led many to propose circumplex models that array various emotions on these dimensions (see Figure 1 for an example).

Many theorists, however, believe that these two dimensions are insufficient to describe the universe of emotional experience (Fontaine, Scherer, Roesch, & Ellsworth, 2007). Darwin (1872/1965), and many before him (Sorabji, 1999).

![Figure 1](image_url)

**Figure 1**

A Circumplex Model of Affect

2000), described a small number of qualitatively distinct emotions as innate, universal natural kinds. Categorical theories of emotion remain prominent (Ekman, 1992a, 1992b). Different theorists have different lists of basic emotions, but all include fear and anger, and most include joy and sorrow. Some include additional emotions, such as surprise (Plutchik, 2003), contempt (Ekman, 1992a; Izard, 1991; Tomkins, 2008), interest (Izard, 1991; Panksepp, 1998), shame and guilt (Izard, 1991; Tomkins, 2008), and acceptance (Plutchik, 2003; Tomkins, 2008).

Modern evolutionary approaches explain specific emotions as coordinated states that give fitness advantages in specific situations that recurred over evolutionary time (Nesse, 1990; Tooby & Cosmides, 1990). These views have changed from a strict modular conception to one that increasingly emphasizes emotions as prototypes without sharp boundaries; they share overlapping elicitors, functions, and physiological and cognitive characteristics (Nesse, 1998; Russell & Fehr, 1994). In contrast, Cosmides and Tooby have argued that selection has shaped thousands of discrete domain-specific mental modules to deal with different situations (Cosmides & Tooby, 1994) and that emotions are superordinate programs that coordinate the modules (Cosmides & Tooby, 2000).

Some theorists reject the idea of categorically distinct emotions, arguing for a multidimensional space with a potentially infinite number of emotions (Barrett, 2006a; Frijda, 1994, 2006; Scherer, Schorr, & Johnstone, 2001; Smith & Ellsworth, 1985). The space contains clusters of common and closely related feelings (such as anger, indignation, and annoyance), which sometimes overlap with other feelings that might be classified as separate emotions by a discrete emotions theorist (such as anger, aversion, and contempt), and sparsely populated regions where there are feelings familiar to only a few cultures or individuals. This point of view is compatible with our own evolutionary perspective.

Just as this evolutionary perspective rejects the idea of sharply distinct basic emotions, it also rejects a sharp distinction between emotions and moods (Beedie, Terry, & Lane, 2005). It is useful to distinguish short-lived emotions aroused by specific cues from moods that may last for days or weeks without specific causes. Also, compared with moods, emotions have more prominent facial and physiological changes, and they may increase fitness by somewhat different routes. They are similar, however, in that both are special states aroused in the situations where they have tended to increase fitness.

**The Origins of Different Emotions**

Specific emotions partially differentiated from more primal generic states because they improved ability to cope with specific kinds of threats and opportunities (Ellsworth, 2007; Nesse, 2004). Figure 2 is a hypothetical phylogeny of emotions. Note the lack of sharp differentiation among

---

**Figure 2**

A Possible Phylogeny of Emotions

---

emotions. Individuals better able to recognize and adapt to the challenges of survival-relevant situations tended to live longer and reproduce more than other individuals. The mechanisms that determine when an emotion occurs evolved in conjunction with the components of the emotion. New emotional variations are useless unless they are expressed in situations where they are advantageous. Detection of significant situations is useless without the ability to express the appropriate emotion.

Different emotions are not defined by different functions or mechanisms, or specific stimuli, or brain modules, or even by particular points in dimensional space. To the extent that there are different emotions, they correspond to different situations that have recurred over the course of evolutionary time. They consist of whatever changes tended to increase fitness in the relevant situation. If each situation were sharply distinct, and if the adaptive responses were different, then selection would shape distinct basic emotions to match each situation. However, there is overlap in both the characteristics of situations and the patterns of response that are adaptive responses, so emotions do not have clear boundaries.

For example, confronting a snake and confronting a bear are similar situations, as are the adaptive responses. Finding a fruit tree or a field of grain evokes positive overlapping responses. An evolutionary view of their origins strongly suggests that emotions are not susceptible to clear definitions or crisp taxonomies. The absence of a designer and millions of years of tiny sequential changes have shaped a mind that is not just complex but indescribable by words and concepts simple enough to be satisfying. The emotions are neither discrete entities nor points on a few dimensions; they are overlapping point-clouds in an N-dimensional space. It should be no surprise that observers in different cultures discern some similar patterns or that they recognize and label constellations of points somewhat differently (Wierzbicka, 1999).

All tangible analogies are inadequate, but something is required. Within the usual analogy of mind as computer, the emotions are like software programs that adjust input, output, memory, processing, and display to cope effectively with a particular kind of task (Ekman, 1992b; Nesse, 1994). However, unlike software programs, emotions were not designed for specific functions. They are closer to the programmable states on an electronic keyboard that adjust the pitch, volume, tone, instruments, background rhythm, distortion, and much more to constellations appropriate for playing rock, blues, classical, soul, tango, and overlapping genres.

The conclusion is disquieting—the clear taxonomy of emotions sought for so long by so many may not exist. No precise description of emotions and their subtypes can be accurate. Although frustrating, this conclusion can liberate us from a fruitless quest so we can turn our attention to the somewhat indistinct structure of emotions, their functions, and the mechanisms that regulate when and how intensely they are expressed.

The Functions of Emotions

Although emotions have sometimes been regarded as mal-adaptive, most contemporary researchers assume that they confer selective advantages (Oatley & Jenkins, 1996; Plutchik, 2003). We have described emotions as special processes that enhance fitness in certain situations. Most emotion researchers have been content to leave it at that and devote their research to other more proximate questions. The few who have attempted to develop evolutionary theories of emotion have generally taken a taxonomical, functional approach, proposing that the differentiation of emotions corresponds to the functions they serve. Positive emotions motivate the organism to take advantage of environmental opportunities and to recognize when it has succeeded in doing so. Negative emotions motivate the organism to avoid misfortune by escaping, attacking, or preventing harm or by repairing damage when it has already occurred.

Different emotions have sometimes been defined by their more specific functions: Fear motivates escape from danger; anger motivates attack; joy motivates continuing on the present course or, if the object has been attained, ceasing to strive for it; disgust motivates avoidance, vomiting, and more metaphorical expulsion; interest motivates exploration; lust motivates seduction and sexual intercourse; sorrow motivates calling for help or giving up on fruitless endeavors, and so on (Gross & Keltner, 1999; Plutchik, 2003). It is worthwhile distinguishing benefits that come from communication, arousal, motivation, memory, and action intentions.

It is tempting to offer a specific function as an evolutionary explanation for each emotion. However, just as the components of emotions are best thought of as parts of one complex pattern, the various functions of an emotion are best understood in terms of how they together increase fitness. One emotion has many functions, and any given function is served by many emotions. Different emotions do not correspond to different specific functions; instead, they correspond to the adaptive challenges encountered in different situations.

Regulation of Emotion Elicitation

An evolutionary approach is sometimes thought to emphasize “innate” responses to universal cues such as snakes, smiles, and darkness and to imply that emotions are fixed action patterns rigidly elicited like reflexes in response to fixed cues. In fact, an evolutionary perspective explains why the mechanisms that regulate emotion elicitation are so flexible and varied.

A looming image has been followed by harm often enough to arouse an innate response of fear and flight (Schiff, Caviness, & Gibson, 1962). Rabbits without an innate fear of foxes have an often-fatal anxiety disorder. However, even fear of snakes is not innate in primates but is only a cue especially conducive to fear conditioning (Mineka, Keir, & Price, 1980). Classical conditioning of emotions allows organisms to experience affect that slightly anticipates an event. Fear two seconds before a
danger is far better than two seconds after. Accordingly, fear can be conditioned more easily, and extinguished less easily, to cues such as snakes and spiders (Ohman & Mineka, 2001). The capacity for operant conditioning offers advantages that are even more obvious. A tendency to repeat whatever works is the most general behavioral adaptation imaginable. Emotions aroused by reward-associated cues have obvious utility (Rolls, 2005). If no reward is forthcoming, motivation declines and disengages goal pursuit (Klinger, 1975), a pattern that is important for understanding the utility of low mood.

Although conditioning adjusts emotions to situations better than fixed responses can, simple learning cannot come close to the effectiveness of human cognition (Goodson, 2002). Our cognitive capacities allow inference about the future, providing a huge advantage. Internal representations of external objects combine with causal schemas to create expectations about the future and about the likely consequences of alternative courses of action. These expectations have predictably powerful influences on emotions. The capacity to anticipate the future also makes it possible to conceive of a goal and pursue it with flexible strategies over many days or weeks. Most human behavior involves goal pursuit, and specific kinds of goal-relevant situations arise repeatedly (Diener & Fujita, 1995; Nesse, 1990, 2004; Oatley & Jenkins, 1996). Opportunities arouse desire and excitement. With steady progress toward the goal, optimism and effort are worthwhile. Frustration is useful to test the scope of an obstacle and ways to overcome it (Oatley & Duncan, 1994). In situations where progress is impossible, low mood disengages effort (Carver & Scheier, 1990; Klinger, 1975). Failure causes disappointment, success causes pleasure. Individuals whose behavior is adjusted by appropriate emotions in these situations have a selective advantage. Figure 3 summarizes some emotions that arise in the pursuit of goals. These are not discrete categories of emotion; they are central tendencies, and because situations overlap, so do emotions.

Of course, seeking a rabbit for dinner is different from seeking admiration from one’s group or affection from a possible mate. So, the kinds of emotions associated with goal pursuit became specialized to deal with different goals in different domains. For instance, signs that a sexual partner is interested in someone else arouse jealousy, a wild and inconsistent mixture of fear, anger, and desire to please (Buss, 2000; Daly, Wilson, & Weghorst, 1982). We recognize and define jealousy because it is aroused by a particular situation. Difficulty in agreeing whether it is a distinct emotion or a combination of other emotions should be no surprise. It is neither. It is a special process that tends to work. Working, in this evolutionary sense, means that jealousy tends to increase reproductive success, even though it may harm an individual’s interests.

Emotions are often elicited in situations where they are useless. This is an inevitable and adaptive outcome. Consider a signal detection analysis of the costs and benefits of panic in a particular situation. If the cost of a false alarm is low, for instance, 200 kcal and 10 minutes, and the cost of not experiencing panic in the presence of a real danger is high, say, 200,000 kcal of damage on average, then a normal system will express many false alarms. In this hypothetical case, the optimal system will express a panic attack whenever a cue indicates a greater than 1 in 1,000 chance that a predator is present. So, 999 out of 1,000 responses will be false alarms that are perfectly normal and useful in the long run. This “smoke detector principle” is crucial for understanding apparently unnecessary anxiety and depression (Nesse, 2005).

Figure 3
Emotions for Situations That Arise in Goal Pursuit

<table>
<thead>
<tr>
<th>Domain</th>
<th>Before</th>
<th>Usual progress</th>
<th>Fast progress</th>
<th>Slow progress</th>
<th>Success</th>
<th>Failure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical</td>
<td>Desire</td>
<td>Engagement</td>
<td>Flow</td>
<td>Frustration</td>
<td>Pleasure</td>
<td>Pain</td>
</tr>
<tr>
<td>Social</td>
<td>Excitement</td>
<td>Friendship</td>
<td>Pride</td>
<td>Anger</td>
<td>Happiness</td>
<td>Sadness</td>
</tr>
<tr>
<td>Physical</td>
<td>Fear</td>
<td>Coping</td>
<td>Confidence</td>
<td>Despair</td>
<td>Relief</td>
<td>Pain</td>
</tr>
<tr>
<td>Social</td>
<td>Anxiety</td>
<td>Defensive arousal</td>
<td>Confidence</td>
<td>Anger</td>
<td>Pride</td>
<td>Shame</td>
</tr>
</tbody>
</table>

Embarrassment
Toward Evolutionary Appraisal Theory

Appraisal theories of emotion (cf. Ellsworth & Scherer, 2003) have a great deal in common with the evolutionary approach we have described. Originally proposed by Magda Arnold (1960), they begin with the assumption that organisms are constantly alert to changes in the situation that might have implications for their well-being. These situational appraisals, along with their associated bodily responses and action tendencies (Frijda, 2006), are experienced as emotions.

Appraisal theorists view situations more abstractly than most evolutionary theorists, identifying certain domain-general characteristics of situations that matter for survival and success in attaining goals. Rather than postulating discrete emotional modules, they see emotions as emerging from appraisals of several important situational characteristics, including the following:

1. Novelty and environmental changes
2. Intrinsic pleasantness/unpleasantness
3. Goal obstacles or facilitators
4. Unpredictability
5. Agency (event caused by self, other, or circumstances)
6. Controllability
7. Compatibility with social norms or personal values

Whenever one of these appraisals changes, the emotional experience changes. Rather than focusing on the concrete features of the situation—a snake, a lightning bolt, a tiger, an unexpected loud noise—appraisal theory proposes that emotions arise from these more abstract appraisals. The snake, the lightning, the tiger, and the noise are all novel (in that they are new elements in the situation), unpleasant, potentially goal-obstructing, and caused by uncontrollable, probably impersonal circumstances, and they all elicit fear. If circumstances were slightly different—for example, if a person realized that the loud noise was her teenager’s ear-splitting music—the appraisal of agency would change, and she would feel anger.

Situations that are similarly appraised evoke similar emotions and similar action tendencies. Appraisal theories, unlike categorical theories, afford a means of describing similarities and differences among events and emotions. Any event can produce an emotional response, even if it has never been encountered before, and that response is predictable if we know how the event is appraised along a small number of dimensions. The appraisal process may be conscious, especially in unfamiliar situations, but often it is not. What the person experiences is an emotion, not its component appraisals, just as we experience a color as a color, not as a combination of brightness, hue, and saturation.

Another important consequence of the focus on appraisals rather than objective features of the situation is that individual differences in emotion are now a central part of the story rather than an inconvenience. There is no doubt that different people often see the “same” situation differently and feel different emotions: The same lecture may be inspiring to one, infuriating to another, and boring to a third. The different emotional reactions correspond to differences in appraisal that result from individual differences in personal values, experiences, and goals.

Individual Variations in Emotions

As with most other traits, variation in emotional responses can be partitioned among variation within individuals, variation among individuals in a group, and variation among groups. Although there is some consistency in emotional responses to certain situations, there is also considerable variability among cultures in the prevalence and salience of different kinds of emotional experience, in emotional reactions to particular situations (Mesquita & Ellsworth, 2001; Mesquita & Frijda, 1992), and in the way emotions are described (Wierzbicka, 1999). There is as much or more variability among individuals within the same culture: Some people are austere, others volatile; some fear dogs, others love them; some are enraged by adolescent excesses, others amused. Focusing on the situations that elicit the same emotion among all members of the species distracts us from the overwhelming preponderance of situations that do not.

The sources of such variation are usually allocated among genes, environment, and interactions between them. Much arises from genetic variation. For instance, Kagan, Reznick, and Snidman (1988) found that differences in inhibition that show up in the first year of life are still apparent in adulthood. About half of the variation in personality measures arises from genetic differences and almost all of the rest from environmental sources that are not shared within the family (Bouchard & Loehlin, 2001). Likewise, genetic variations account for about half of the variation in the startle reflex (Anokhin, Golosheykin, & Heath, 2007) and a third of the variation in alexithymia (Jørgensen, Zachariae, Skyttøe, & Kyvik, 2007). How can natural selection have left so much genetic variation? The most likely explanation is that a wide range of emotional tendencies has resulted in individuals with approximately equal Darwinian fitness. Typical emotional responses are well worth studying, but an evolutionary view makes it clear that human nature is not a single essentialized pattern but a set of capacities and tendencies with substantial individual variation from genetic differences.

Human variation in emotions also arises from learning. Some of this learning is simple conditioning and extinction, but much involves more abstract knowledge. People differ in knowledge and expertise. Snakes rarely evoke fear in herpetologists. Novelty, the initial appraisal and the gateway to emotion (Ellsworth, 1994; Kagan, 1991), disappears with experience. Many events that surprise or frighten a child are commonplace to the same person grown up. Over the course of a lifetime, people learn love for things they once detested (oysters; intellectual effort), indifference toward things they once feared (vacuum cleaners, the class bully), and fear of things they never cared about (Alzheimer’s disease, cholesterol).
A person who is competent in a domain sees fewer obstacles and feels a greater sense of control than a person who is not (Bandura, 1977), and a competent person feels very different emotions in that domain than does an incompetent person. The human capacity for learning includes both the ability to generalize and the ability to differentiate, so that two situations may seem indistinguishable for some people, whereas for others they may seem to have little in common. The first group may feel the same emotion in both situations; the second may feel radically different emotions. The extinction of anxiety responses is not a matter of merely reversing conditioning. Instead, new cortical influences exert inhibition on lower brain regions (Quirk, 2002).

Humans differ not only in what they know but in what they want. Some universal needs map well onto the resources that are useful to increase fitness: health, skills, material resources, mates, children, friends and allies, and social status. However, people vary considerably in their desire for influence, love, respect, or money. A person feels intense emotion when she or he experiences unexpected progress or obstacles in the pursuit of goals that she or he cares about, but little emotion may be aroused by progress toward goals that seem imperative to other people. To understand a person’s emotions, it is not enough to understand the objective situation; one must understand how the person understands the situation.

Social Emotions

One domain is so important for humans that it demands separate treatment. Humans live in complex networks of relationships that require close attention to reputation and extraordinary skills in negotiating reciprocal exchanges. Reproductive success depends on the ability to negotiate these complexities. Like other emotions, the social emotions make most sense when examined relative to the situations in which they are useful.

The benefits of reciprocal exchange have been at the center of evolutionary analyses of human social behavior (Cosmides & Tooby, 1992; Fessler & Haley, 2003; Hammerstein, 2003; Trivers, 1971). These are usually modeled using the prisoner’s dilemma, in which each player chooses, on each move, to cooperate or defect (Axelrod & Hamilton, 1981). Steady mutual cooperation produces the greatest net long-term payoff, but on any given move, a big advantage can be gained by defecting if the other cooperates. Hundreds of studies have shown that people are fairly good at maximizing profits from this game, although they tend to be too generous at the start and too punitive in response to defections by the other (Axelrod & Dion, 1988).

The four situations that recur in this game are good candidates for shaping the evolution of social emotions (Ketelaar, 2004; Nesse, 1990, 1998). Mutual cooperation creates friendship and trust; temptations to defect arouse anxiety; defection creates guilt; suspicion is useful when the other might defect; and if she or he does, anger is useful. A number of studies have confirmed and elaborated the role of emotions in reciprocal exchanges (Keltner & Haidt, 1999). See Figure 4.

Much current work strives to explain human behavioral tendencies that are problematic for economists or evolutionists because they do not obviously benefit the individual (Dugatkin, 2006; Krebs, 2000). Many people are more generous and more spiteful than is optimal (Fehr & Fischbacher, 2003). A whole field has developed to explore possible explanations (Hammerstein, 2003). Some possible explanations for this “irrationality” include strong reciprocity (Gintis, 2000), gene–culture co-evolution (Boyd & Richerson, 2002; Henrich & Henrich, 2006), theories based on tags that indicate reputation (Riolo, Cohen, & Axelrod, 2001), commitment (Nesse, 2001), social structures that maintain punishment of defectors (Boyd, Gintis, Bowles, & Richerson, 2003), and the possibility that runaway social selection that is due to the benefits of being preferred as a social partner can shape extremes of altruism and concern about what others think (Nesse, 2007). More work is needed to untangle the relative and interacting contributions of these factors, but each of them may help us to understand emotions that are otherwise hard to explain, such as loyalty and love that maintain human communal relationships (Mills & Clark, 1994).

Combining a game theory approach with evolution, however, offers a fairly straightforward explanation for the apparently irrational unpredictability of certain emotions because predictability imposes big disadvantages (Nesse, 1990). If you know your opponent will not really fight, you can be much more cavalier than if you have to worry about an irrational attack or spiteful retaliation. Convincing others that you might act irrationally usually requires actually doing so on occasion. In economics, this is called a com-

---

**Figure 4**

Emotions for Situations That Arise in Social Exchanges

<table>
<thead>
<tr>
<th>You</th>
<th>Other</th>
<th>Before</th>
<th>After</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cooperate</td>
<td>Cooperate</td>
<td>Trust</td>
<td>Gratitude</td>
</tr>
<tr>
<td>Cooperate</td>
<td>Defect</td>
<td>Suspicion</td>
<td>Anger</td>
</tr>
<tr>
<td>Defect</td>
<td>Cooperate</td>
<td>Anxiety</td>
<td>Guilt</td>
</tr>
<tr>
<td>Defect</td>
<td>Defect</td>
<td>Disgust</td>
<td>Rejection</td>
</tr>
</tbody>
</table>

---

February–March 2009 • American Psychologist

135
mitment strategy; individuals gain an advantage by convincing others that they will act in ways that are not self-interested (Frank, 1988; Haselton & Ketelaar, 2006; Nesse, 2001). Fortunately, benefits also come from irrational positive commitments, such as continuing a marriage “in sickness and in health.”

**Emotional Disorders**

Emotional disorders have been studied almost exclusively from a proximate perspective. This seems sensible: After all, selection does not shape disorders. However, only evolution can account for the vulnerabilities that leave us susceptible to disease, from the appendix to our narrow coronary arteries to emotional disorders (Nesse & Williams, 1994). In the case of emotions, the crucial conclusions are that negative emotions are useful and that the systems that regulate their elicitation seem delicate indeed. Instead of the reductionism of neuroscience, a Darwinian approach encourages analysis of emotional disorders in terms of their normal functions, the same kind of foundation physiology provides for understanding other diseases (Baron-Cohen, 1997; McGuire & Troisi, 1998; Nesse, 1999).

How much negative emotion is too much? Psychiatrists and many researchers routinely use checklists to make diagnoses based on the number, severity, and duration of symptoms. Depression, for instance, is defined as abnormal if five or more of nine symptoms are present for more than two weeks. Data on what is going on in the person’s life are ignored, with the one telling exception of loss of a loved one. This is, of course, scientifically senseless (Horwitz & Wakefield, 2007; Nesse & Jackson, 2006). The normality of a response, whether it is pain, fever, sweating, anxiety, or depression, requires careful consideration of what produced it.

Just as there can be too little sweating and too little pain, there must be disorders characterized by too little anxiety or unhappiness. Few efforts have been made to identify people with anxiety deficits. They do not complain, but the lack of anxiety in hypophobics (Marks & Nesse, 1994) must impose costs analogous to those experienced by people who lack a capacity for pain, who are almost all dead by early adulthood. Sociopaths lack normal moral emotions; interestingly, most also lack the capacity for normal anxiety (Lykken, 1995). Just as there are disorders characterized by deficient negative emotions, there must be disorders characterized by an excess of positive emotions. Together, they constitute “diagonal psychology,” with topics neglected owing to the emphasis on the costs of negative emotions and the benefit of positive emotions.

It is often assumed that all harmful responses arise from abnormal brain mechanisms. This is incorrect. According to the smoke detector principle, many experiences of negative emotion will be useless or harmful even though they are perfectly normal. They are, of course, only useful in the aggregate over the long run. It would be wonderful if we could tell when they are unnecessary and find drugs or other ways to inhibit their expression. After all, most pain is perfectly normal, but blocking it is one of the wonders of modern technology.

The utility of emotions such as sadness and depression is harder to see. Sadness occurs after a loss, too late to prevent it. But if one thinks from an evolutionary perspective, the question is whether experiencing a loss is a situation with fitness consequences. It is, of course (Horwitz & Wakefield, 2007). Sadness can motivate searching for what is lost, and if it is not found, seeking a replacement (Nesse, 2005). Its aversiveness motivates avoidance of actions that preceded the loss, preventing future losses.

Depression is often described as persistent sadness, and the affective state certainly looks and feels similar. However, the situations are different. A specific loss arouses sadness that fades with time, especially if it is possible to replace what was lost. Symptoms of depression, by contrast, are aroused when an important goal seems unattainable (Carver & Scheier, 1990; Klinger, 1975; Nesse, 2000). The initial response is to seek new strategies, but if no route to the goal seems possible, motivation fades away, freeing up effort for other more profitable tasks. If for some reason the goal cannot be abandoned, then ordinary low mood tends to escalate into pathological depression.

This perspective is very different from that of stress models that tend to explain depression as the result of damage that negative events produce in individuals with a diathesis (Monroe & Simons, 1991). Some depression certainly results from primary brain abnormalities, and in some patients with severe depression, stress hormone regulation is awry. However, the capacity for symptoms of depression exists only because there are situations in which they are useful. This biological view of depression is quite different from common perspectives based on reductionistic studies of proximate mechanisms.

**Etiology**

The tendency to dichotomize the causes of emotional disorders between nature or nurture remains pervasive. Many psychiatrists say that “depression is a brain disease,” whereas psychologists tend to emphasize early experience, habits of negative thinking, or social factors. What is missing is a way to integrate these different factors. A framework based on the origins and functions of normal low mood can help. Without this understanding, studies of depression are like studies of coughing that investigate excess activity in the medullary cough center or possible infections without ever considering the interactions between variations in the situation and variations in the regulation mechanisms.

Neuroscience stands especially to benefit from such a framework (Nesse, in press). For instance, if the capacity for low mood is an adaptation, then it should be unsurprising that no single genetic polymorphism accounts for more than a small percentage of the variance: The mood regulation system is created by products from thousands of genes, each of which may influence vulnerability. Furthermore, they may influence vulnerability via a multitude of routes, from direct actions on transmitter receptors to personality...
tendencies that increase the likelihood of the depressogenic pursuit of an unreachable goal. This is also relevant to the search for more effective medications. Just as pain can be relieved by multiple different pharmacological actions, from decreasing inflammation to the subjective effects of opiates, there is every reason to think that depression can be relieved by disrupting the mediating systems at multiple loci. Being cognizant of the evolutionary origins and functions of mood and anxiety offers a far more biological foundation for studies of etiology and treatment.

Treatment

An evolutionary view of the emotions does not encourage one kind of treatment above others, but it does suggest a simple framework for considering what can be changed to relieve an emotional disorder. Emotions arise from situations the appraised meaning of which influences brain mechanisms. Four things can change: the objective situation, the person’s motivational structure, the perception of the situation, and the brain.

The first target is the situation itself. The best treatment for anxiety and depression from a pending house foreclosure is renegotiating the mortgage. However, most pathological emotions arise from brains predisposed to negative emotions interacting with problems that have no ready solution.

The second target for intervention is the individual’s motivational structure. If the external environment cannot be changed, a change of strategy may help, and if that does not work, giving up the goal may be necessary. Of course, the emotions already are, as it were, on the case.

The third target is the person’s subjective assessment of the situation. More positive beliefs about the self and the future can help, and cognitive therapy is reliably effective (Beck, 1976). And dynamic defenses such as denial and reaction formation are almost as valuable as a sense of humor.

Finally, there are direct influences on the brain, from drugs to electroconvulsive therapy. They may sometimes restore a “chemical imbalance,” but they more often block brain mechanisms that mediate aversive emotions. Interestingly, antidepressants generally do not cause pleasure but merely relieve the pain, much as aspirin does not lower body temperature but merely eliminates fever.

Such a taxonomy is terribly simple. However, many who seek treatment find themselves with a therapist who offers only one treatment, usually with a corresponding ideology about etiology. This is unfortunate. The etiology of aversive emotions is different for different people. An evolutionary analysis offers the beginnings of a foundation for matching treatments to the needs of individuals.

Conclusions

Darwin’s theory of natural selection offers much more for understanding the emotions than does his theory of emotions. Asking evolutionary questions about how selection shaped the capacities for emotions leads to conclusions that address fundamental questions, such as what emotions are. For other questions, such as how to define different basic emotions, an evolutionary perspective suggests that the kinds of answers we have been seeking do not exist. Far from being reductionist and rigid, evolutionary biology offers a framework for understanding the functional significance of emotions in individual lives and the prevalence of emotional disorders.

REFERENCES

Diener, E., & Fujita, F. (1995). Resources, personal strivings, and sub-


