BOSTON COLLEGE O'NEILL LIBRARY
DOCUMENT SERVICES - REQUEST FORM

DEPARTMENT CHARGE

Name Michael Ring Date 1/31/07

E-Mail Address psychoff@bc.edu

Campus Mail Address McGuinn 301

Phone 552-4100 Department to be charged Psychology

NOTICE WARNING CONCERNING COPYRIGHT RESTRICTIONS

Signature Michael Ring

Delivery Methods:

☐ Campus Mail ☐ Pick-up ☐ E-Mail as a: ☒ PDF ☐ JPEG ☐ TIFF

Book/Journal Title Social Psychology and the Unconscious

Chapter/ArticleTitle Chapter 4: On the Automaticity of Emotion

Author Lisa Feldman Barrett

Volume/Date ______ Pages 173 – 217

☒ Hold originals for pick-up by patron Michael Ring

☐ Return originals via campus mail

Special Instructions:

Do not write below this line

4

On the Automaticity of Emotion

LISA FELDMAN BARRETT, KEVIN N. OCHSNER, and JAMES J. GROSS

Any emotion, if it is sincere, is involuntary

Mark Twain

The year is 1846. You are a stagecoach driver. All is quiet and peaceful as the red sun sets beyond the horizon, the coach rumbles contentedly along, and sagebrush cast long shadows across the lonesome prairie. Hidden in the shadows, however, is a rattlesnake. Disturbed from its slumber by the horses, the rattler suddenly strikes out, scaring the horses into a fearful, frenzied sprint. Out of control, the stagecoach careens towards the edge of a sheer cliff. First you try to soothe your steeds, but they cannot hear you. Then you try to forcibly rein them in, but their strength is too great. Life itself hangs in the balance as you grimly struggle to control the careening stagecoach.

The distinction between wild stagecoach steed and wily stagecoach driver in many ways mirrors the distinction between feeling and thinking embedded within Western Culture. Emotions are assumed to be primitive, automatic, animalistic entities dwelling within us that the more developed human part of our minds come to know about and control. The notion that feeling is first, fast, and feral traces back to biblical stories of the First Family and their misbegotten emotional impulses to taste that tempting fruit. These ideas about emotion continue in modern-day stories of lovers driven mad with jealousy, businessmen blinded by greed, and widows overcome with grief. As Plato suggested long ago in the Phaedrus, in each of these cases, our emotions, like wild horses, drive us to emotional places we do not deliberately choose to visit and thus must be harnessed and restrained.

In the first section of this chapter, we outline the commonsense view that emotions automatically play themselves out when we encounter certain situations. We describe how this commonsense view – with varying degrees of elaboration and complexity – forms the basis of a consensual view of emotion that pervades much of the scientific inquiry into emotion. We refer to this consensual view as "the modal model" of emotion. In the second section, we argue that although the "modal model" has much to recommend it, mounting evidence suggests that it ha several important limitations. In the third section, we argue that the field needs t move beyond a search for entities that conform neatly to our intuitions about wha automatic emotions "must" be like. We call for a richer examination of the bottom up and top-down processes that together give rise to emotion, and suggest that constraint satisfaction approach may provide the conceptual framework that i needed in order to move beyond the modal model.

FROM COMMON SENSE TO A "MODAL MODEL" OF EMOTION

The Commonsense Approach to Emotion

As common sense has it, emotions are triggered automatically, overcome us, and cause us to act. We yell because we are angry. We cry because we are sad. We jump because we are afraid. Anger, sadness, and fear cause and explain our behavior, just as lightning causes and explains thunder. As the pent up electrical energy of an emotion is discharged, the result is a largely inescapable set of stereotyped outputs that occur in a rapid, involuntary, and automated fashion. Prior knowledge, expectations, beliefs, or any other cognitive input have little impact on the process. You might know that lightning is about to strike, but you can't stop it from happening, and plugging your ears will not keep you from hearing the thunderous blast. As a consequence, emotions such as anger happen *to* you, and overcome you, rapidly overriding whatever else you might have been doing, thinking, and feeling. From a first-person perspective, the conscious experience of emotion (the feeling) is taken as clear evidence that the causal mechanism – the "emotion" – was triggered. Feeling angry is evidence that the anger mechanism has fired. What is more, anger seems to overtake others in much the same way. When observing others, expressive behavior is seen as evidence that an emotion is triggered. The given quality of our own experience, and the way that emotion seems to control others without their awareness, seems to provide proof for the automaticity of emotional responses.

This folk conception of emotion underlies our everyday construal of emotions in self and others. Because we experience instances of anger (or sadness or fear) in ourselves and in others as having a rapid onset and a more or less ballistic trajectory, we believe anger (or sadness or fear) must exist as a primitive entity lying in wait within the brain or body, ready to spring forth automatically and at a moment's notice once the appropriate triggers are present. The commonsense idea of emotions as automatic eruptions is even evident in the criminal justice code, where "passion" is seen as a justifiable defense for violent crime. In the US, the "sudden heat of passion" constitutes adequate provocation that reduces an act of intentional homicide to an offense of voluntary manslaughter (Dressler, 2001, p. 527).

In this view, our emotions are rarely, if ever, the product of controlled, deliberate, and conscious thought. Although it is possible to "think ourselves" into an emotional state, controlled processes typically serve to control, rather than to elicit, emotional responses. Any number of aphorisms admonish us to regulate our feelings either by deliberately thinking differently – "looking on the bright side," "making a silk purse out of a sow's ear," "finding the silver lining in every dark cloud," – or by acting differently – "putting on a happy face," "putting a lid on it," "showing some restraint," "never letting them see you sweat," "getting a grip," or more generally by keeping ourselves from expressing the anger, sadness, or fear that we might feel inside. The very fact that we have to take control to regulate our emotional responses is further evidence for us that they are automatic, and it is precisely because we experience our emotions interfering with our more reasoned responses (that we identify as more essentially "human" in nature) that we experience our emotions as automatic, animalistic, and foreign.

More generally, our experiences of emotions erupting outside of our control, and our sense of agency and effort in shaping them, strongly supports our intuition that there are two fundamentally distinct forms of processing that characterize the human mind: automatic processing (which we share with other animals) and controlled processing (which is most developed in humans). As we will see, virtually every major scientific account of emotion incorporates our intuitions about dual-processing modes in the mind, and specifically our commonsense distinction between automatic elicitation and controlled regulation of emotion.

Two Major Scientific Approaches to Emotion

Dual-process models pervade contemporary psychology (e.g., Barrett, Tugade, & Engle, 2004; Chaiken & Trope, 1999; Devine, 1989; Gilbert, 1991, 1998; Power & Dalgleish, 1997; Schacter, 1997; Sloman, 1996; Smith & DeCoster, 2000; Trope, 1986). A central tenet of such models is that behavior is determined by the interplay of automatic and controlled processing. Models vary in their specifics and emphasis, but all hold that responses to an event begin with the automatic (sometimes called nonconscious, implicit, or heuristic) processing of information. Such processing is assumed to be a default mode, initiated by the simple registration of sensory inputs, which in turn passively activate knowledge structures (called schemas, scripts, or concepts, or even internal goal states) that shape perception and action. Controlled (sometimes called conscious, explicit, or systematic) processing can determine, to a large extent, the degree to which automatic processing is expressed in thoughts, feelings, and behaviors. Controlled processing requires attentional resources, is volitional, is largely goal-driven, and can be used to modulate automatic processes when the outputs they produce conflict with valued goals.

As in other domains in psychology, emotion has proven hospitable ground for dual-process logic (Smith & Neumann, 2005). Indeed, in our view, a dual-process model lies at the heart of much of the scientific theorizing and research dealing with emotion for the past century. This dual-process framework can be summarized in a very simple way in Figure 4.1. Some event or goal-relevant stimulus (usually external, although it could be internal) triggers an emotion mechanism (or set of mechanisms), which, in turn, automatically produces a complex sequence of coordinated changes in the brain and body that constitute an emotional respons

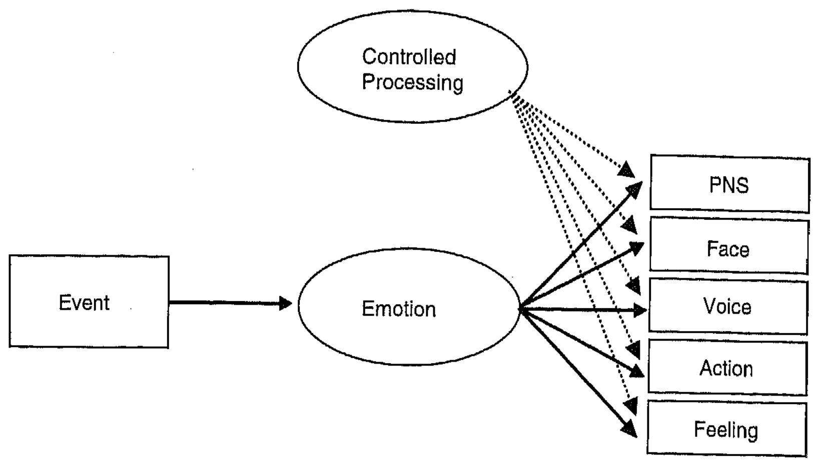

FIGURE 4.1 A dual-process view of emotion. PNS = peripheral nervous system activation.

Substitute for "emotion" any referent for a specific emotion category ("anger", "sadness", "fear", and so on, in English), and you have a model of that emotion. Controlled regulation is thought to occur separately and modulate the extent to which a coordinated emotional response actually manifests in observed behavior.

Within this dual-process framework, two historically distinct (but often complementary) approaches to the study of emotion can be distinguished. One approach has focused on the output side of the emotion-generative process, namely the coordinated expression of complex patterns of behavior that comprise the observable, tangible, and socially impactful component of an emotional response. This has been referred to as the *basic emotion* approach. A second approach has focused on the input side of the emotion-generative process, namely the processing of environmental stimuli that gives rise to the emotional response depicted in Figure 4.1. This has been referred to as the *appraisal* approach. Despite differences in their surface features (for a review, see Scherer et al., 2001), these two approaches share two central assumptions. First, each of these approaches assumes that there are definable kinds of emotion (defined by the brain, or by the deep structure of situations). Second, these approaches assume that emotion generation is dominated by automatic processing (with regulation usually occurring after the fact). Given its ubiquity in the field, we refer to Figure 4.1 as the "modal model" [and in the past have referred to it as the "natural kind model" (Barrett, 2006b) or the "consensual model" of emotional responding (Gross, 1998)]. In the following sections, we describe the role the "modal model" has played in each of these two major approaches.

The Basic Emotion Approach One of the earliest modern examples of the basic emotion approach can be found in Darwin's (1859/1965) "The Expression of the Emotions in Man and Animals." Darwin's ideas about emotion were infused with commonsense beliefs about how mental states seek expression in, and therefore automatically cause, behaviors. Facial and other behaviors (such as baring the teeth in anger) were seen as vestigial expressions of emotion mechanisms that are homologous in human and nonhuman animals. Darwin focused on a small number of emotions (many of which are now referred to as "basic" emotions), stressing the universality as well as the phylogenetic and ontogenetic continuity of their expressions.

William James (1884, 1890, 1894) famously disagreed with Darwin (and commonsense) as to the correct temporal ordering of the elements in Figure 4.1. Instead of the sequence depicted in Figure 4.1, James argued that an individual's emotional response was best characterized by the sequence: stimulus → physiological response → experience. Notwithstanding this notable difference, James clearly retained the core notion that emotion involved an automatic release of a coordinated set of responses to relevant stimuli (with the embodiment of those responses producing the experience of emotion). As William James himself put it:

> The nervous system of every living thing is but a bundle of predispositions to react in particular ways upon the contact of particular features of the environment . . . The neural machinery is but a hyphen between determinate arrangements of matter outside the body and determinate impulses to inhibition or discharge within its organs . . . Every living creature is in fact a sort of lock, whose wards and springs presuppose special forms of key – which keys however are not born attached to the locks, but are sure to be found in the world near by as life goes on. And the locks are indifferent to any but their own keys. (1884, pp. 190–191).

This quote nicely captures the modern idea of cognitive impenetrability – meaning that emotion circuits fire in an obligatory way once triggered by sensory information about a stimulus, and occur regardless of the context. Onset is rapid, involuntary, and requires little or no attention. Although James concentrated on developing the automatic elicitation side of the emotion equation, he also incorporated the commonsense notion that controlled processes may come into play, such as when we regulate emotion after-the-fact, by limiting its expression. As James saw it: "refuse to express a passion, and it dies." (James, 1884, p. 197).

Later models built more directly on this elicitation-regulation distinction, and further developed the dual-process metaphor for emotion processing. Cannon (1927, 1931) and Bard (1928; Bard & Rioch, 1937), who proposed one of the earliest psychological models of emotion localization in the brain, argued that the emotional part of the brain (in their view, the hypothalamic circuit including the thalamus and hypothalamus) produced responses that could be downregulated by evolutionarily more recent neocortical regions. Papez (1937) similarly argued for top-down cortical regulation of subcortical emotional responses, and MacLean (1949) continued this tradition, positing that the newer "mammalian" part of his triune brain architecture exerted top-down regulatory control of the emotional responses that issued from the older and more primitive "reptilian" and "old mammalian" parts of the brain.

Although concepts such as reptilian and mammalian are no longer part of contemporary basic emotion models, these models have nonetheless retained an emphasis on subcortical structures in the generation of emotion (e.g., Panksepp, 1998). Perhaps the best-known example has been offered by LeDoux (1996), who demonstrated that links between sensory systems and the amygdala are necessary and sufficient for the expression of conditioned fear, but who, along with others, has shown that cortical areas (particularly medial prefrontal cortex) are involved in expressing contextual learning that inhibits the conditioned fear response (Milad & Quirk, 2002; Morgan, Romanski & LeDoux, 1993; Quirk & Gehlert, 2003; Quirk, Likhtik, Pelletier, & Paré, 2003).

A dual-process metaphor can also be clearly seen in the family of models that comprise the modern-day "basic emotion" approach to emotion. Like Darwin and James, basic emotion models focus more on the ways in which emotional responses are automatically elicited. Boiled down to their essential ingredients, these views posit that each kind of "basic" emotion issues from a neural program or circuit, hardwired at birth, homologous with circuits found in nonhuman mammals, that is responsible for an automatic syndrome of hormonal, muscular, and autonomic effects that constitutes the distinctive signature of an emotional response. Although the specific set of "basic" emotions varies somewhat across emotion theorists, there is agreement that specific "basic" emotion mechanisms correspond to English emotion categories (e.g., "fear," "sadness," "disgust," "anger," or happiness"). This small set of stereotyped, automatic emotion responses are regulated after the fact, usually by means of controlled attentional processes shaped by epigenetic influences, such as context and learning history.

Perhaps the best-known example is Ekman's neurocultural model (Ekman, 1972), which describes emotions as issuing from "affect programs" (Tomkins, 1962) that, once triggered by an eliciting stimulus, direct a complex pattern of coordinated outputs to produce a stereotyped emotional response. Panksepp's (1998; Panksepp et al., 2000) neurobiological model takes its lead more directly from MacLean's (1949) triune brain concept, but is similar, in principle, to the neurocultural model. Panksepp argues for different "basic" emotion systems (seeking/expectancy, rage/anger, fear/anxiety, lust/sexuality, care/nurturance, panic/separation, and play/joy). Each kind of emotion is a separate, inherited, complex reflex that is hardwired at birth and causes a distinctive syndrome of hormonal, muscular, and autonomic effects.

At the core of both Ekman's (1972) and Panksepp's (1998; Panksepp et al., 2000) models is the idea that there is a hardwired set of emotion-specific mechanisms that fire automatically and thereby generate a suite of emotional responses. Over time, however, there has been softening of the emphasis on fixed, hardwired programs that govern emotions from birth. For example, both Ekman and Panksepp acknowledge that there is a greater range of human emotional responding than can be accounted for by a set of basic emotions. Recent developments in the neurocultural model have attempted to account for the complexity and subtlety of emotional life by arguing for families of emotion response (Ekman, 1992), or by suggesting a distinction between "primordial" and "elaborated" emotions (Keltner & Haidt, 2001), where the former are "basic" emotions that produce a stereotyped response signature, and the latter are responses that are more shaped by the norms and social practices within a culture. Both also allow controlled processing to enter the picture, although primarily as a way of regulating the emotion-generative process. In the neurocultural model, culture not only influences the stimuli that trigger emotion programs, it also specifies display rules and regulatory outcomes, so that observed emotional responses display considerable cultural variation, even as the causal mechanisms are hardwired into the brain. Panksepp (1998) also allows for the environment to modulate emotional outputs in the form of cortical control of the basic emotion systems once they have been triggered.

The Appraisal Approach The dual-process metaphor can also be found in a second family of emotion models, which we refer to collectively as the *appraisal* approach. Appraisal models have been concerned with patterns of cognition that trigger an emotional response. Like the basic emotion approach, many models within the appraisal approach retain the commonsense distinction between automatic elicitation and controlled regulation, although they also incorporate the idea that automatic and controlled cognitive processes (or steps) can interact and give rise to emotional responses. Thus, when appraisal models unpack the input side of Figure 4.1, they typically describe a cognitive logic that involves both automatic and controlled components, although once the emotion is elicited, it is assumed to run automatically to completion.

In these models, emotions are a consequence of how people construe situations. Frijda (1988), one of the best-known and most influential appraisal theorists, calls this the "law of situational meaning." Instead of assuming that a stimulus situation automatically triggers or releases a fixed emotional response (as William James had), appraisal models hold that intervening cognitive processes automatically elicit and determine the quality and intensity of emotional responses. Input an event with a particular meaning, and the result is an emotion of a particular kind. The cognitive processes that compute this meaning, formally known as appraisals, link the external world (an individual's immediate situation) to the internal world (the individual's goals, needs, and concerns). Appraisals diagnose whether the situation in question is relevant to the person's well-being, and if so, identify the nature of that relevance, and trigger an emotion that will maximize the likelihood of producing a functionally effective response consistent with the organism's most central concerns (Kappas, 2001; Smith & Kirby, 2001).

Appraisal models vary in terms of the logic and content of the cognitive judgments that are held to be necessary and sufficient to produce emotional responses. In some models, appraisals describe the way that a situation is experienced, and constitute a descriptive structure of which emotions are felt when; they do not, in and of themselves, indicate the processes by which the meaning is made or arrived at (e.g., Ortony, Clore, & Collins, 1988; Smith & Ellsworth, 1985). In this view, sadness occurs when a situation's meaning involves loss, and fear when the meaning involves danger. Situational analysis, or appraisals, can be determined in any number of ways. A situation's meaning can be determined with associative processing by reinstating an appraisal that derives from a similar situation experienced in the past, or it can be computed on the spot using a rule-based analysis driven by features of the situation and the goals of the person (Clore & Ortony, 2000). Both types of processing can be automatic (Smith & DeCoster, 2000), although on-line computation allows the possibility for more controlled processing.

In other models, appraisals do more than describe the meaning of situations – they are a set of cognitive processes that literally generate an emotional response (e.g., Roseman, Antoniou, & Jose, 1996; Scherer, 1984). Even the appraisal-as-mechanism models differ in which appraisals are seen as necessary and sufficient, which combinations of appraisals elicit particular emotional responses, which categories of emotion are explained, and so on (Roseman & Smith, 2001). They also vary in the fixity of the appraisal process. Some assume that appraisals are made in a specific fixed sequence (e.g., Scherer, 1984, 2001), while others argue for more flexible ordering in appraisal processing (e.g., Frijda, 1986; Roseman et al., 1996), although often it is assumed that particular appraisals (whether a stimulus is pleasant or unpleasant, novel or familiar) come before others that can be more flexibly deployed (e.g., Ellsworth, 1991; Lazarus, 1991). They differ in the extent to which they relax the assumption of cognitive impenetrability (the idea that emotion elicitation is not influenced by factors such as prior knowledge, expectations, beliefs, or any other cognitive input). Appraisal models also differ in whether they consider these rule-based computations to *cause* an emotional response, to *constitute* the response, or to be a *consequence* of emotion processing (cf. Ellsworth & Scherer, 2003; Frijda & Zeelenberg, 2001).

Despite their differences, virtually all appraisal theorists hold that people are continually assessing situations for personal relevance, beginning with an evaluation of whether or not the stimulus is "good for me/bad for me" (Arnold, 1960; Lazarus, 1966; Mandler, 1984; Ortony et al., 1988; Roseman, 1984; Scherer, 1984; Smith & Ellsworth, 1985). The primacy of such an evaluation is consistent with the general idea that we automatically evaluate stimuli (e.g., Bargh, Chaiken, Govender, & Pratto, 1992; Fazio, 2001; Fazio, Sanbonmatsu, Powell, & Kardes, 1986; Ferguson & Bargh, 2004), as well as the specific notion that some aspects of emotional responding – at least those that are related to computing affective valence – are generated automatically (e.g., Berridge & Winkielman, 2003; Cacioppo et al., 1999). Appraisal theorists also hold that different situations that evoke the same appraisal pattern produce the same emotional episode. Each emotion is elicited by a distinctive pattern of appraisals (e.g., Arnold, 1960; Frijda, 1986; Lazarus, 1991; Ortony, Clore, & Collins, 1988; Roseman, 1984; Scherer, 1984; Smith & Ellsworth, 1985), and the pattern of appraisals, rather than a dedicated neural circuit, is responsible for generating the emotional response.

Appraisal models, like basic emotion models, initially retained the automatic elicitation – controlled regulation distinction. Arnold (1960), who was the first contemporary appraisal theorist, assumed that appraisals are an "intuitive and immediate" assessment of the stimulus situation (p. 182). An explicit debate about the importance of the automatic versus controlled processing in emotion generation (Lazarus, 1982; Lazarus & Folkman, 1984; Zajonc, 1980, 1984) refocused attention to the idea that the conceptual processing engendered by appraisal processes can occur automatically, and now most appraisal models incorporate a role for simple, nonconscious appraisals in emotion generation. Even when appraisals occur in response to remembered or imagined events, the cognitive processing involved in the appraisal process is thought to proceed automatically. Particularly in models that propose appraisals as preconditions for activating emotional responses (rather than constituting the responses themselves), appraisals are seen as inputs to the neural circuit view characterized by basic emotion models. Once appraisals have been computed, an emotion is triggered in a way that is very similar to Figure 4.1.

Over time, however, appraisal theorists have shown increasing appreciation for the role of controlled processing in emotion generation, and models have more explicitly relied on the dual-process metaphor that is implicit in basic emotion and early appraisal approaches. For example, Leventhal & Scherer (1987) argued that two kinds of automatic processes generate emotional responses that are modified by a third type of controlled process. An initial sensory-motor level of processing implements a form of biologically prepared perceptual processing driven by innate, unconditioned, hard-wired feature detectors that give rise to reflex-like responses. A second level (thought to mediate the majority of emotional responses) implements schematic processes that automatically match current stimulus patterns to learned stimulus patterns to generate coordinated emotional responses. When these first two types of processes generate a response that is sufficiently intense to enter awareness, consciously guided conceptual-level processes come into play, deploying propositional knowledge to refine emotional responses. Conceptual processing is thought to become ever more automatized (like skill learning) with practice. Many other models similarly suggest that emotions can be generated by some combination of automatic and controlled processing (e.g., Clore & Ortony, 2000; Dalgleish, 2004; Power & Dalgleish, 1997; Smith & Kirby, 2001; Teasdale, 1999; Teasdale & Barnard, 1993; Wells & Matthews, 1994).

In relaxing the assumption of cognitive impenetrability, appraisal models usually allow for the possibility that controlled processing can implement appraisal logic (i.e., the rule-based evaluations that cause the resulting emotional response). The common idea in these models is that various forms of automatic processing (including low-level perceptual processing of stimuli and prior knowledge in the form of schemas) interact with more effortful processing to produce emotional responses (Clark & Isen, 1982; Frijda & Zeelenberg, 2001; Lazarus, 1991). In this way, most appraisal theorists seem to agree that humans play an active role in shaping their information processing, and can exert some control over emotion-generative appraisal processes.

The "Modal Model"

In the past, basic emotion and appraisal approaches have been treated as opposing explanations for emotional responding (Ortony & Turner, 1990; Turner & Ortony, 1992). We believe that despite their differences, both approaches share two intuitively appealing assumptions that can be found in our commonsense ideas about emotion. These common assumptions comprise the core of what we refer to as the "modal model" of emotion. The first of these assumptions is the idea that emotional responses can be characterized as belonging to a small set of discrete categories. The second is the assumption that there is a boundary between the automatic generation of emotion and its controlled regulation after the emotion itself has been triggered.

First, both basic emotion and appraisal accounts focus their attention on explaining a small set of discrete emotions. Although appraisal models acknowledge (at least in principle) the enormous variety in emotional responding and do not assume that particular emotions are basic in any biological way, most models organize emotional responding into the familiar set of discrete categories used by basic emotion theorists. Major research efforts have been directed at identifying the profile of appraisals for a fixed number of discrete emotions (anger, sadness, fear, and so on) that are very similar to the list discussed by basic emotion approaches (e.g., Roseman, 1984, 1991; Smith & Ellsworth, 1985). For example, although Lazarus (1966) initially rejected the idea that there are a limited numbe of categorically distinct "basic" emotions, he came to view that there are a smal number of "relational themes" that correspond to discrete emotions (Lazarus 1991; Smith & Lazarus, 2001). Similarly, Ortony et al. (1988) discussed a larg range of emotional responses, but proposed a hierarchical structure in which som emotions are just more differentiated versions of other emotions. Scherer (2001 suggested that we tend to use basic level categories like "anger" to refer to qualita tively different emotional states, some of which may be "modal" emotions whos appraisal profiles recur with some frequency (Scherer, 1994), whereas other emo tions may follow from fluctuations in appraisal profiles that may yield a larg number of different emotional responses (e.g., Scherer, 2001).

Second, both basic emotion and appraisal accounts rely heavily on the notio that emotions are generated automatically. Thus, both types of models posit tha emotional responses act as an organizing force, "hijacking" the entire system (i.e. disrupting whatever other processes are operative at the time) to deal with th circumstances that elicited the emotion in the first place. Oatley and Johnson Laird (1987), for example, suggested that basic emotions are internal signals tha disrupt ongoing cognitive processing and reset it into specific modes to deal wit basic biosocial challenges. Once an emotion is triggered (whether computed by a emotion program or a set of appraisals), there results an inescapable, involuntar and automated set of synchronized changes in response systems that produce th signature emotional response (like the output side of Figure 4.1).

Both assumptions embodied by the modal model – that there are a small set c different kinds of emotion, and that emotions issue more or less automatically have guided emotion research for the past century. They have guided the que tions that researchers ask, the way that emotional responses are measured, and th interpretation of the data. The resulting research programs have been productiv and important. No model is perfect, however, and as we shall see in the ne section, there are both theoretical and empirical reasons for a course correction i the way that science approaches the study of emotion.

A CRITICAL EVALUATION OF THE MODAL MODEL

The modal model has been a tremendously valuable organizing force in the field of emotion. The basic emotion approach has helped to define emotion as a topic worthy of study in its own right, facilitating the development of empirical methods for examining facial (e.g., Ekman & Friesen, 1978), vocal (e.g., Scherer, 1986), autonomic (e.g., Cacioppo, Klein, Berntson, & Hatfield, 1993), and central (e.g., Davidson & Irwin, 1999) aspects of emotional responding. It has served as the de facto yardstick against which competing accounts of emotion are evaluated. The appraisal approach has helped to establish the importance of personal relevance and meaning in triggering emotional responses, and has attempted to unpack the notion of ballistic, automatic action programs into a more complicated set of both automatic and controlled processes that together contribute to the generation of an emotional response.

Useful as the modal model has been, however, it is limited and limiting in at least two ways. First, as we describe in more detail below, the modal model privileges a relatively limited number of emotions, leaving large gaps in our understanding of the full spectrum of emotional phenomena in need of explanation. Second, as we see it, the available empirical evidence, guided by a research agenda defined by the modal model, does not uniformly support the core tenets of this model. Specifically, the modal view posits automatic processing mechanisms that do not dovetail neatly with our emerging understanding of the behavioral and biological bases of emotion. As a result, a comprehensive functional architecture for emotion that considers a complete scope of emotional phenomena, and specifies a testable set of functional and neural mechanisms, has, to date, failed to coalesce from this model. In the following sections, we critically consider these two important limitations of the modal model.

Are There a Limited Number of Discrete Kinds of Emotion?

An account of emotion, according to Clore and Ortony (2000, p 32), "needs to do justice to the full richness and range of emotions that comprise human emotional life." The modal model, however, focuses attention on just one part of the larger emotional landscape by considering a small number of kinds of emotion about which we can make inductive discoveries, and which conform to the event → automatic response pattern. In this way, the modal model leads us to restrict our scientific inquiry to characterizing only a fraction of our emotional life. This practice is consistent with the viewpoint held by many emotion researchers that emotions should be defined by species-general aspects. As a field, we ask questions like, "How many emotions are there?" "Which specific pattern of antecedent events, neural activity, physiology, and motor behavior defines each emotion?" and "How do we evoke pure instances of emotion, uncontaminated by contextual influences?" Guided by the modal model, we assume that *kinds* of emotions would reveal themselves if only we could find the right eliciting stimulus or measures (cf. Barrett, 2006b).

Although it is clear that people have experiences that they refer to as anger, sadness, fear, and so on, there is also good evidence that they can experience many other varieties of emotional response as well – responses not readily classifiable as fitting one of the canonical emotion kinds or forms. Variability may be the norm, rather than the exception, and according to some evolutionary biologists, variability is the thing to be explained (Lewinson, 2000). Although movies and novels are replete with examples of full-blown canonical emotion responses, emotion scientists have yet to take an empirical tally of how often these occur in everyday life. It is just assumed that they occur often enough to justify an almost exclusive focus on them empirically. Certainly these stereotyped responses are rarely, if ever, seen in all their glory in the laboratory. What is more, as we describe below, behavior, as well as the physiology that supports it, is more context-sensitive, and linked to the requirements of the situation, than the modal model of emotion would lead us to expect (cf. Barrett, 2006b; Bradley, 2000; Cacioppo, Berntson, Larsen, Poehlmann, & Ito, 2000; Davidson, 1994; Lang, Bradley, & Cuthbert, 1990). Functional demands vary with situations, making it likely that instances of the same emotion can be associated with a range of behaviors (e.g., behaviors associated with fear can range from freezing to vigilance to flight). This observation suggests that there is considerable heterogeneity in emotional responses that might be called fearful (or angry, etc.).

Even putting aside the issue of whether important aspects of human emotional life are neglected by the modal model, there remains the question as to the success of the search for definable patterns of coordinated responses that characterize each kind of emotion. Despite a century of effort, and much to everyone's surprise, there has been little accumulation of evidence to support the hypothesis that emotions represent clearly defined kinds (for a review, see Barrett, 2006b). To appreciate this unexpected outcome, consider a key prediction of the modal mode and one of the most compelling ideas in the psychology of emotion: the notion tha emotional states have specific and unique patterns of somatovisceral changes.

Although individual studies have reported distinct autonomic correlates fo different emotion categories (e.g., Christie & Friedman, 2004; Ekman, Levenson & Friesen, 1983; Levenson, Ekman, & Friesen, 1990), meta-analytic summarie generally fail to find distinct patterns of peripheral nervous system responses fo each basic discrete emotion (Cacioppo et al., 2000). Instead, peripheral nervou system responses configure for conditions of threat and challenge (Quigley Barrett, & Weinstein, 2002; Tomaka et al., 1993, Tomaka, Blascovich, Kibler, & Ernst, 1997), and for positive versus negative affect (Cacioppo et al., 2000; Lang Greenwald, Bradley, & Hamm, 1993). What is puzzling here is the gap betwee common sense (of course different emotions affect my body differently!) and th available physiological findings. Studying the physiological bases of emotion i fraught with challenges (Levenson, 1988), and it is certainly possible that method ological problems are responsible for the impoverished empirical evidence fo physiological distinctions among emotions. But it is important to recognize that a equally viable alternative explanation for the lack of consistent findings is tha there are, in fact, no clearly demarcated and discrete kinds of emotion that ar unambiguously distinguished by patterns of peripheral physiological responses.

This sentiment is reinforced by a similar gap between common intuitions th emotions cause distinctive patterns of expressive behaviors and the relevant empirical data. One of the major research efforts in the basic emotion approach has been concerned with detecting and describing universal facial expressions of emotion (Ekman & Friesen, 1978; Ekman, Friesen, & Ellsworth, 1972; Ekman, Sorenson, & Friesen, 1969; Izard, 1971). This research has yielded evidence that posed facial expressions can be judged with considerably greater than chance accuracy by individuals from a wide range of Western and non-Western cultures (cf. Ekman, 1994; but see Elfenbein & Ambady, 2002). Although perceivers can reliably assign posed facial configurations to discrete emotion categories, alternative explanations for these findings have been offered (cf., Russell, 1994; Russell, Bachorowski, & Fernandez-Dols, 2003), including the idea that perceivers are imposing, rather than detecting, true categorical distinctions in the facial configurations that they rate (Barrett, 2006a). More important, however, is the fact that very limited systematic data exist concerning the *production* (as opposed to the perception) of emotion expressions across cultures. Even within a culture, facial electromyography measurements coordinate around positive versus negative affect (Cacioppo et al., 2000) or intensity of affect (Messinger, 2002), rather than discrete emotion categories per se. More generally, it has been suggested that expressive behaviors in mammals rarely broadcast fixed, encoded messages about the sender's internal state (Fridlund, 1994; Seyfarth & Cheney, 2003), suggesting facial movements and vocal signals may not necessarily "display" information about the sender's emotional state (cf. Russell et al., 2003), even though we perceive them as coordinated "expressions."

Neither is it clear that vocal sounds carry specific information about discrete kinds of emotion (for a review, again see Russell et al., 2003). Listeners also do better than chance at classifying acted portrayals of emotion in vocal cues (Banse & Scherer, 1986; Hess et al., 1988; Johnstone & Scherer 2000; Juslin & Laukka, 2003; Wallbott & Scherer, 1986), but these portrayals do not necessarily have the same acoustic characteristics that are observed in naturally produced vocal expressions (for a discussion, see Bachorowski & Owren, 2003). Furthermore, the acoustic properties of produced vocal cues give clues to a speaker's identity (Edmonson, 1987), indicate his or her arousal level (e.g., Bachorowski 1999; Bachorowski & Owren 1995; Kappas et al., 1991), and are thought to elicit affective responses in listeners (Bachorowski & Owren 2001) more than they give evidence about kinds of emotion per se.

In like fashion, a given instrumental behavioral response need not express a specific kind of emotion. For example, although fear may be associated with freezing, fear is associated with a number of other behaviors, ranging from vigilance to attack (for a review, see Bouton, 2005). The threat (or defense) system is organized so that an animal will engage in different behaviors depending on its psychological distance from a predator (Fanselow, 1994; Fanselow & Lester, 1988). Not only are different behaviors associated with the same emotion category, but one type of behavior can be associated with many categories. For example, varieties of aggressive behavior (e.g., defensive, offensive, predatory) are associated with different types of stimulus situations and are caused by different neural circuitry (Blanchard & Blanchard, 2003).

Evidence from reports of subjective experience also calls into question whether or not there are bounded and distinct kinds of experiences. Not only do people vary greatly in whether or not they distinguish between feelings of anger, sadness, and so on, in reports of subjective experience (Barrett, 1998, 2004; Feldman, 1995), but these reports can be decomposed into more elemental psychological properties, such as valence and arousal. In revealing valence, and to a lesser extent arousal properties, self-reports of experienced emotion produce a similar structure to that which is observed for psychophysiological and behavioral measures of emotion.

Perhaps most important, however, is the finding that physiological, behavioral, and experiential outputs for each emotion category are not as highly intercorrelated as one might expect based on the modal model (Mauss, Wilhelm, & Gross, 2004; for a review, see Bradley & Lang, 2000). Psychophysiologists have long observed weak correlations across response systems (e.g., Weinstein, Averill, Opton, & Lazarus, 1968) and even within the same "response system" (e.g., Lacey, 1967). Recent studies similarly have found modest correlations among measures of emotional responding in the context of fear (Hodgson & Rachman, 1974; Lang, 1988; Rachman, 1984), exhilaration (Ruch, 1995), and surprise (Reisenzein, 2000). Although links between emotion experience and facial behavior have tended to be the strongest (e.g., Adelmann & Zajonc, 1989; Blumberg & Izard, 1991; Rosenberg & Ekman, 1994), even these links are often modest (e.g., Fernández-Dols, Sánchez, Carrera, & Ruiz-Belda, 1997; Fridlund, 1994) and inconsistent (e.g., Casey, 1993; Chovil, 1991; Gross, John, & Richards, 2000; for a meta-analytic review, see Cacioppo et al., 2000). Despite much effort, then, there has been surprisingly little evidence generated for the modal model's prediction of strong response coupling in emotional responding.

As this brief review indicates, physiological, behavioral, and experiential data do not strongly support the notion that there are clearly identifiable discrete kinds of emotion. The accumulating neuroscience evidence echoes this observation, thus far failing to yield strong evidence of dedicated neural circuits for basic emotion categories (Barrett, 2006b). Although there is good evidence that specific behaviors – such as freezing, the baring of fangs or claws, or hair standing on end – each may depend upon specific brainstem and subcortical nuclei (e.g., Panksepp, 1998), there is little evidence to suggest that a single brain structure is responsible for the production of the complete range of behavioral changes that should be associated with any single emotion category (Barbas, 1995; Cavada, Company, Tejedor, Cruz-Rizzolo, & Reinoso-Suarez, 2000). Similarly, many different cortical and subcortical brain systems are capable of modulating physiological and behavioral correlates of emotion, such as heart rate and respiration or freezing, and no single response system appears to be linked exclusively to a single specific emotion (Cacioppo & Gardner, 1999).

Lesion studies suggest that the normal experience and perception of some emotions, such as disgust and fear, may depend critically upon the integrity of particular brain structures – the insula and amygdala, respectively – but expressive deficits following amygdala or insula lesions typically are not absolute, and imaging studies suggest that both of these structures also appear to participate in the