BEHAVIORAL AND BRAIN SCIENCES (2005) **28,** 169–245
Printed in the United States of America

Bridging emotion theory and neurobiology through dynamic systems modeling

Marc D. Lewis
Department of Human Development and Applied Psychology, University of Toronto, Toronto, Ontario, M5S 1V6 Canada
mlewis@oise.utoronto.ca **http://home.oise.utoronto.ca/~mlewis**

Abstract: Efforts to bridge emotion theory with neurobiology can be facilitated by dynamic systems (DS) modeling. DS principles stipulate higher-order wholes emerging from lower-order constituents through bidirectional causal processes – offering a common language for psychological and neurobiological models. After identifying some limitations of mainstream emotion theory, I apply DS principles to emotion–cognition relations. I then present a psychological model based on this reconceptualization, identifying trigger, self-amplification, and self-stabilization phases of emotion-appraisal states, leading to consolidating traits. The article goes on to describe neural structures and functions involved in appraisal and emotion, as well as DS mechanisms of integration by which they interact. These mechanisms include nested feedback interactions, global effects of neuromodulation, vertical integration, action-monitoring, and synaptic plasticity, and they are modeled in terms of both functional integration and temporal synchronization. I end by elaborating the psychological model of emotion–appraisal states with reference to neural processes.

Keywords: appraisal; bidirectional causality; cognition; dynamic systems; emotion; neurobiology; part–whole relations; self-organization

1. Introduction

Both emotion theorists and neuroscientists have studied emotions, and their relation to cognition and behavior, for decades. Yet communication between them has remained relatively constricted. Most emotion theorists continue to ignore neuroscience almost entirely. At a recent international conference on emotions, there were 17 symposia and talks on the psychology of emotion, two on the neurobiology of emotion, and only one that attempted to bridge these perspectives. In partial contrast, neurobiologists adopt basic concepts from emotion theory, but they ignore the larger phenomena that are of greatest interest to psychologists. For example, they map isolated appraisal mechanisms onto specific brain regions, but rarely consider an appraisal as a coherent mental model corresponding with an emotional state. Given emotion theorists' goal of moving toward an integrated "affective science," and given the rapid progress in emotional neurobiology that could expedite this move, why do the psychology and neurobiology of emotion remain largely isolated?

One reason for this isolation may be that emotion theorists and neuroscientists view cause–effect relations and part–whole relations in terms that are nearly incommensurable and, thus, find it difficult to talk to each other. Emotion theory relies on causal assumptions that are simple, linear, and often cognitivist in character, and it emphasizes psychological wholes (e.g., appraisal, attention, "emotion") without explaining how they derive from interacting parts. Neural accounts incorporate far greater complexity, including bidirectional causal assumptions, but they focus almost exclusively on interacting parts (i.e., neural structures and subsystems) while ignoring the properties of the whole. Thus, complex causal processes remain elusive for emotion theorists, psychologically meaningful wholes remain elusive for neuroscientists, and there is little common ground for truly integrative modeling.

In this article I suggest a bridge between emotion theory and neurobiology based on dynamic systems (DS) principles. Nonlinear dynamic systems operate through reciprocal, recursive, and multiple causal processes, offering a language of causality consistent with the flow of activation among neural components. Consequently, psychological accounts informed by DS ideas may be more biologically plausible and better able to integrate neural findings. Dynamic systems are also characterized by the emergence of wholes out of interacting parts, through processes of self-organization and Haken's (1977) circular causality. Such multilevel causal processes can relate coherent wholes such as appraisals, emotions, and traits to the interaction of lower-order constituents, integrating levels of description

Marc Lewis is a Professor in Developmental Science at the University of Toronto. He has used dynamic systems methods for studying emotional behavior over development. However, his current research is focused on cortical processes of emotion in children and adults, using dense-array EEG/ERP techniques. He is particularly interested in the contributions of corticolimbic regulatory mechanisms to personality development and developmental psychopathology.

© 2005 Cambridge University Press 0140-525X/03 $12.50

https://doi.org/10.1017/S0140525X0500004X Published online by Cambridge University Press

both for emotion theory and emotional neurobiology. Thus, DS ideas may provide a foundation for building models that incorporate the rich psychological categories of emotion theory with the biological realism of neuroscience, by addressing causal relations and part–whole relations in a manner relevant for both.

I begin with a brief review of approaches to cognition–emotion relations in emotion theory, followed by a closer look at process models proposed by appraisal theorists. I then show how DS principles can help to reconceptualize appraisal–emotion interactions unfolding in time. Here and elsewhere, DS ideas are presented conceptually, not mathematically. There are many possible approaches to formal modeling that are consistent with this presentation. Next, I outline a model of self-organizing appraisal–emotion amalgams, highlighting reciprocal causation, emergence, and other DS principles. I then turn to emotional neurobiology, introduce some of the structures and functions thought to underpin emotion and appraisal, and propose several mechanisms of integration through which they interact. Finally, I extend the psychological model through reference to neural events and show how this translation helps refine the model while integrating neural processes into psychologically meaningful wholes. By taking two passes at DS modeling, first in psychological terms and then with the addition of neural terms, I attempt to show how DS principles bridge the two disciplines and yield a synthesis enriching to both.

This article does not propose a detailed neurobiological theory of emotion or cognition–emotion relations. Rather, it provides a DS-based framework that can serve as a platform for integrating psychological and neural perspectives, and it develops a model that instantiates this framework feasibly, as a working hypothesis. Predictions based on the model are framed in relation to present and future research directions. Yet, the specifics of both the model and the predictions are secondary to the conceptual framework underlying them. There is simply no overarching framework available, to date, for synchronizing psychological and neural perspectives on emotion and its relation with cognition, behavior, and individual traits; and, although dynamic systems ideas are no longer new, they have never been applied to developing such a framework.

2. Cognition–emotion relations: The view from emotion theory

2.1. Definitional issues

Emotion theory is a large amalgam of approaches to the study of emotion, mostly based on cognitive psychology but with contributions from learning theory, physiological psychology, clinical psychology, and other disciplines including philosophy. The central problem for emotion theory is how emotion interacts with other processes that have been studied more extensively: most notably, cognition. In emotion theory, as in other fields, "cognition" includes perception, attention, evaluation, decision-making, memory, and so on. However, these functions are also lumped under the nomenclature of *appraisal*, defined as the evaluation of significance in a situation that can give rise to an emotional response. Despite its lack of precision, the term "appraisal" is important in most emotion theories, because it denotes an evaluative or interpretive function that is critical for eliciting emotion. Thus, following emotion theorists, I will often use *appraisal* to mean, roughly, cognitive processes that are directed toward what is important for the self. The definition of emotion can be imprecise as well; but many theorists view emotions as response systems that coordinate actions, affective feeling states, and physiological support conditions, while narrowing attention to what is important, relevant, or available to act upon.

As can be seen, working definitions of appraisal and emotion are partially overlapping, especially in terms of evaluating "what is important." This can be problematic for emotion theory. In fact, the central assumption that appraisal and emotion are distinct functions is not as solid as one might like, and several theorists do away with it entirely (e.g., Ortony et al. 1988). However, this solution frames emotion as a class of cognition, leaving emotion theorists without a unique phenomenon to theorize *about*. Despite these definitional conundrums, appraisal and emotion are important and useful psychological categories that are central to a large body of theory and research. I will therefore use these terms throughout the article. When a neurobiological perspective is presented in sections 4–6, however, it will be necessary to revisit definitional issues and examine the overlap between "appraisal" and "emotion" more systematically.

2.2. Three approaches

Among the many approaches to cognition–emotion relations taken by emotion theorists, three stand out as subfields in themselves. The first and most coherent is appraisal theory. As noted, an appraisal is an evaluation of a situation in terms of its relevance for oneself, specifically one's goals or well-being (e.g., Lazarus 1968), and it is the means by which we extract meaning from events and make sense of the world (e.g., Frijda 1993b). Appraisal approaches attempt to determine the specific perceptions, evaluations, interpretations, and so forth, that are necessary and sufficient to elicit a particular emotional state. The classic way to do this has been to define and study a set of appraisal dimensions. Theorists disagree on the exact number and content of appraisal dimensions, but goal-conduciveness, novelty, coping potential, and norm-compatibility figure in most theories (Scherer 1999). Hypothetically, a distinctive pattern or profile on these dimensions stipulates a mental model and elicits a unitary emotion, somewhat like the opening of a lock when all the numbers line up (Lazarus 1991; Ortony et al. 1988; Reisenzein 2001; Roseman 1984). Empirically, subjects report on their thoughts and feelings in response to an emotion-eliciting vignette or the recollection of a significant life event. Their responses are scored on various appraisal dimensions, and the pattern of scores is related to a reported or observed emotional state (e.g., Reisenzein & Hofmann 1993; Roseman 1991; Roseman et al. 1996; Smith & Lazarus 1993).

By producing meaning, an appraisal sets the occasion for an emotional response whose purpose is to organize cognition and action relevant to the situation as appraised (Lazarus 1999; Roseman & Smith 2001). Thus, appraisal theorists generally view appraisal as a temporal and causal antecedent to emotion (Scherer 1993b; 1999; Roseman & Smith 2001). This assumption of the causal and temporal precedence of appraisal has been responsible for considerable controversy, epitomized by the famous Lazarus-Zajonc

https://doi.org/10.1017/S0140525X0500004X Published online by Cambridge University Press

debate years ago (Lazarus 1984; Zajonc 1984). The debate was eventually put to rest, partly because it reduced to a war of definitions, but also because many appraisal theorists agree that emotions must influence appraisals in some way (Ellsworth 1991; Keltner et al. 1993; Roseman & Smith 2001; Scherer 1999; Smith & Kirby 2001). To finally resolve the temporal sequencing of appraisal and emotion, a (moment-to-moment) process account is necessary, and this appears to be the current challenge facing appraisal theorists (Scherer 2001).

The second approach to cognition–emotion relations is concerned with the cognitive *function* of emotions. Perhaps the main function of emotions is to direct attention to relevant aspects of the environment in the service of action tendencies for altering that environment (e.g., Ekman 1994; Frijda 1986; Izard 1993). From this perspective, attention is always "motivated" (Derryberry & Tucker 1994), and cognition is generally constrained by the type (e.g., sadness, anger, happiness, fear) of emotional state. A good deal of research demonstrates the effects of emotional states on attention and other cognitive processes (e.g., Bower 1992). Sad versus happy emotions differentially affect attentional style and content (Isen 1990). Anxiety biases or narrows attention (Mathews 1990), perhaps through its effects on filtering or inhibition (Wood et al. 2001). Emotion also biases perception (Mathews & MacLeod 1985; Niedenthal et al. 1994), memory retrieval (Eich & Metcalfe 1989; Isen 1985), social judgement (Forgas & Bower 1987; Keltner et al. 1993), and cognitive organization in general (Oatley & Johnson-Laird 1987; see review by Mathews & MacLeod 1994). According to these approaches, the source of the emotion (e.g., "mood induction") is not as important as its effect on subsequent cognitive events. Thus, importantly, the direction of influence, from emotion to cognition, is generally the opposite of that highlighted by appraisal theories.

A third approach is to examine cognition–emotion relations as influenced by personality traits and clinical disorders. Enduring biases, associations, and emotional habits have been well documented in emotion research (e.g., Bradley et al. 2000; Mathews & McLeod 1994). Personality traits such as optimism, pessimism, and anxiety, and clinical disorders including anxiety and depression, have been shown to increase the probability, endurance, and consistency of appraisals that code for specific aspects of the world (Beck et al. 1987; Gallagher 1990; Jerusalem 1993; Teasdale & Barnard 1993). These appraisals also tend to elicit corresponding emotions (Carver & Scheier 1991; Kuiper & Martin 1989). The effects of emotions on attention are often described in terms of trait-specific biases, as when anxious subjects attend selectively (and unconsciously) to anxiety-related words (Mathews & MacLeod 1985; 1986) or faces (Fox et al. 2002), and depressive subjects attend to negative words (Wenzlaff et al. 2001) and recall negative experiences (Clark & Teasdale 1982) more than normal controls. These approaches are generally not concerned with the etiology of the personality or clinical traits that produce these effects, but only with the effects themselves.

All three approaches to cognition–emotion relations take a particular view of causality. They conceptualize phenomena in terms of simple, linear causal processes – cause–effect relations that go in one direction from antecedents to consequents. Appraisal theorists generally view appraisal as the causal antecedent of emotional states; functionalist researchers view emotions as causing cognitive changes; and trait researchers view traits as causes rather than effects of emotional biases. Yet a bird's-eye view of the field suggests more complex causal relations: in particular, bidirectional effects among emotion-related phenomena. If appraisals give rise to emotions and emotions influence cognitive processing, then bidirectional causation would be important for explaining appraisal–emotion interactions (Lewis 1996). Moreover, if different emotions bias attention differentially, and appraisals are consistently constrained by these differences, then individual traits may result from emotional biases as well as cause them. Thus, ignoring bidirectional causal mechanisms yields incomplete models within the branches of emotion theory and impedes integration across these branches.

It should be noted that a few models of cognition–emotion relations do discuss bidirectional or nonlinear causal processes (e.g., Carver & Scheier 1990; Frijda & Zeelenberg 2001; Mathews 1990; Scherer 2000; Teasdale & Barnard 1993), and at least three draw specifically on DS constructs (Lewis 1996; Scherer 2000; Thayer & Friedman 1997). However, some of these models (e.g., Frijda 1993b) treat nonlinear considerations as secondary and do not develop them (see Lewis 1996), others (e.g., Scherer 2000) pursue them as an interesting diversion from more classical modeling, and still others (e.g., Carver & Scheier 1990) focus on negative feedback and error correction but ignore positive feedback and emergence. Despite such limitations, these contributions are important steps toward revamping the core assumptions of emotion theory. They set the stage for further development of bidirectional models, and they will be discussed and integrated further in Section 3.2.

The predominance of simple, linear causality in emotion theory is matched by a tendency to model psychological phenomena as wholes rather than parts. Global constructs such as emotions, appraisals, and traits, and monolithic systems such as attention and memory, figure strongly in the language of emotion theory. When these constructs and systems are dissected into components, they are done so in information-processing terms: resource allocation, propositional networks, knowledge stores, search procedures, value assignment, and so forth (e.g., Mathews & MacLeod 2002; Reisenzein 2001; Teasdale & Barnard 1993; Wells & Matthews 1994; Williams et al. 1997). The causality connecting these information processes can be characterized as cognitivist or computationalist, whereby a sequence of symbol manipulations computes an output for a predesignated problem (Fodor 1975; Putnam 1975). Examples include dimensional approaches to appraisal and information-processing models of attentional bias (e.g., Roseman 1991; Wood et al. 2001). Thus, part–whole relations, when addressed at all, are addressed mechanistically and linearly (with the exception of Teasdale & Barnard 1993, whose model is mechanistic but nonlinear). Such explanations tend to remain at the level of parts rather than wholes and to lack the psychological realism emotion theorists wish to achieve.

Appraisal theorists have become increasingly concerned with the relations between parts and wholes and have begun to develop *process models* capable of connecting them realistically (Frijda & Zeelenberg 2001; Reisenzein 2001; Roseman & Smith 2001; van Reekum & Scherer 1997). However, they have not yet abandoned the language of cog-

https://doi.org/10.1017/S0140525X0500004X Published online by Cambridge University Press

nitivist causation. Process models of appraisal are reviewed next, with particular attention to the problems that continue to arise from linear causal assumptions.

2.3. Process models of appraisal

Process models examine the relations between appraisal components and emotional outcomes. But what are the components of appraisal? Dimensional approaches, reviewed earlier, focus on combinations of semantic elements (e.g., novelty, coping potential) to represent appraisals. Process models focus instead on psychological elements. A list of such elements might include perception, evaluation, attention, memory, reflection, and planning – essentially a breakdown of "cognition" broadly defined. By identifying the components of appraisal in this way, theorists can examine their interactions with each other and with emotion as they play out over time.

A key premise of process models is that appraisals are not necessarily slow, deliberate, or even conscious, and they can be perceptual as well as strictly "cognitive" (Lazarus 1984; Scherer 1993b). In fact, rapid, pre-attentive stimulus processing has been an important focus of appraisal theory and research (see Lazarus 1995). Leventhal and Scherer (1987) postulated sensorimotor and schematic processing levels that operate prior to cognitive evaluation. Frijda (1993b) has argued that only the shape of an object need be perceived in order to elicit an emotion. A variety of studies have shown that pre-attentive appraisals are sufficient to elicit emotions such as anxiety, fear (Öhman 1988; Öhman & Soares 1994; Robinson 1998), and possibly anger (Berkowitz 1989). Interestingly, this corresponds well with evidence that coarse perceptual processing by the amygdala is sufficient to activate emotion (LeDoux 1987; see also Lazarus 1999; Scherer 1993b) and that animals without cortical controls have intact emotional responses (LeDoux 1995a; Panksepp 1998a).

However, most theorists are primarily interested in the "higher cognitive" levels of appraisal that seem fundamental to normal human emotions. To incorporate both pre-attentive and higher cognitive processes in their models, many appraisal theorists consider appraisal as a sequence composed of two or more phases. Lazarus's (1966) initial formulation specified a rapid primary appraisal followed by a more reflective secondary appraisal. Öhman (1993) proposes an initial pre-attentive monitoring function followed by more controlled processing once significant events are perceived. Frijda (1993b) views secondary appraisals as necessary for fleshing out rudimentary evaluations that elicit primitive emotional responses. Smith and Kirby (2001) propose a rapid associative process followed by a slower, more controlled, and more focused stage of reasoning. Scherer's (1984; 1999) scheme involves a progression of five increasingly comprehensive evaluative checks operating recursively throughout the duration of an emotion-producing event. The earlier checks are rapid, pre-attentive, and concerned with sensory data, whereas the latter constitute higher-order cognitive evaluations that demand conscious attention.

Thus, sequential accounts of appraisal view some appraisal processes as rapid and some as slow; but they nevertheless maintain an assumption of linear causality. Early appraisal processes lead to later appraisal processes, and later appraisal processes lead to emotions. Or, emotions are partially determined by early appraisal processes and more fully determined by later ones. Yet, it is still appraisal processes that determine emotions, and not the other way around. Frijda (1993b) has been the most prominent critic of "linear" appraisal models, and Frijda and Zeelenberg (2001) make the radical (for emotion theory) claim that appraisal can be viewed as an outcome rather than an antecedent. This idea is also at the root of the few DS treatments undertaken so far (Lewis 1996; Lewis & Granic 1999; Scherer 2000).

The consensus view thus denies emotion a causal role in appraisal processes. Emotions are viewed, instead, as consequents of a one-way sequence of information-processing events. As with any cognitive computation, appraisal events are seen to sort and process information, follow a series of steps, mediate between a condition and an action, and lead to a final answer – an emotional response. However, by keeping emotion out of the causal equation, appraisal theorists ignore the demonstrated effects of emotion on attention, evaluation, judgment, recall, and other cognitive processes. Based on these findings, emotions (or emotion components) generated by pre-attentive processes early in an appraisal sequence should have an impact on what happens next. As sketched in Figure 1, they should immediately begin to influence successive perceptual, attentional, and higher-order cognitive processes, thus contributing to the final appraisal configuration via bidirectional causation. One critical barrier to this kind of modeling is that, even when appraisals are broken down into component events, emotions are still construed as monolithic wholes. This makes it impossible to include emotion "parts" in the causal chaining of appraisal events.

Some theorists represent the effects of emotion on appraisal by proposing appraisal–emotion streams. If appraisals are necessary to specify emotion states, yet emotion states influence subsequent cognitive processes, then there may be a continuous flux or stream of evaluative events in which appraisal and emotion are interspersed (Ellsworth 1991; Lazarus 1999; Parkinson 2001; Scherer 2001). However, this compromise abandons the important idea of an emotional episode (Frijda 1993a). Or, it necessitates two models of appraisal: a classical model of emotion-antecedent appraisal for clearly circumscribed events and a loosely defined appraisal–emotion stream for continuous events. Most critically, this approach tends to construe emotions and appraisals as wholes, not parts. Rather than view appraisals and emotions as alternating wholes in a temporal stream, a process account should demonstrate how constituent processes give rise to a whole appraisal in the first place (see Fig. 1).

2.4. Summary and conclusions

Most emotion theorists apply simple, linear causal models to explain the interactions among psychological entities such as appraisals, emotions, and traits. This brand of causality highlights psychological wholes, or information-processing parts, but it does not relate parts and wholes realistically. Process models have the potential for more realistic explanations; but the linear sequencing of appraisal events, and the general disinclination to see emotions as componential, leave no place for emotional effects on the formation of whole appraisals. Nevertheless, evidence from across the branches of emotion theory indicates that many

https://doi.org/10.1017/S0140525X0500004X Published online by Cambridge University Press

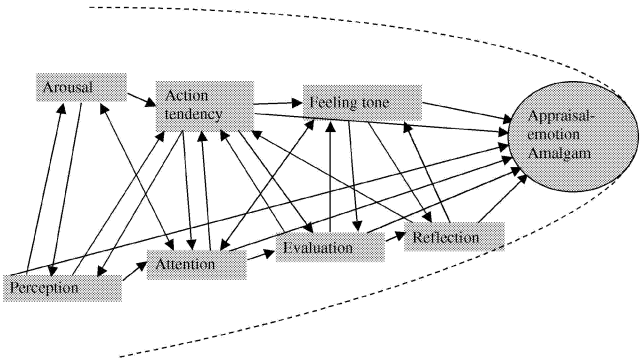


Figure 1. A skeletal model in which bidirectional causal relations between appraisal and emotion constituents lead to a whole appraisal–emotion amalgam. This scheme captures the contribution of emotion components to whole appraisals at the psychological level of description, but the components themselves have to be reconceptualized at the neural level, as discussed later.

cognitive processes influence emotions, emotions influence many cognitive processes, these effects are rapid and spontaneous, and they play out differently for different individuals. These findings suggest complex, bidirectional causal relations that organize parts as well as wholes at both momentary and developmental time scales.

Not only do linear causality and cognitivist assumptions create logical problems for integrating findings from within emotion theory, but they also make it difficult to integrate the burgeoning findings from neurobiology into psychological models. Neurobiology examines processes that are componential, complex, and bidirectional, including reciprocal connections at all levels of neural circuitry from neurons to whole brain systems. In doing so, neurobiology concretely defines constituent parts and subsystems and then relates them via principles of recruitment and synchronization that bear little resemblance to information-processing models. To establish a more comprehensive model of appraisal–emotion relations that can span psychological and neurobiological domains, I next introduce a dynamic systems approach to causality and part–whole relations.

3. Self-organizing appraisal–emotion states

Many characteristics of dynamic systems can be applied to understanding appraisal and emotion. In this section, a basic picture of dynamic cognitive systems is introduced first, with particular emphasis on self-organization, the process by which coherent wholes emerge and consolidate from interacting constituents. Then, a variety of DS principles are applied to appraisal processes and their interactions with emotion, establishing the foundation for a new psychological model.

3.1. Cognition as self-organization

Many cognitive scientists no longer characterize cognition as a linear computational sequence between input and output functions. Rather, they highlight the dynamic, distributed, nonlinear, and emergent aspects of cognition (Clark 1997; Elman et al. 1996; Kelso 1995; Port & Van Gelder 1995; Van Gelder 1998; Varela et al. 1991; Ward 2002). Traditional AI (artificial intelligence) has been overshadowed by connectionist models, and the structure and function of the nervous system have become criteria for plausibility. From this newer perspective, cognition builds on itself, biasing its own outcomes, and moves with only partial predictability from moment to moment. Thus, cognition can be viewed as a process of self-organization in a complex dynamic system. The appeal of viewing cognition as a self-organizing process goes back to the cognitive revolution, when Ashby (1952), McCulloch and Pitts (1943), and von Neumann (1958) explored the surprising properties of processing networks and feedback loops. However, it is not until recent years, with the rapid growth of connectionism, cognitive neuroscience, embodied cognition, and nonlinear dynamic cognitive models, that cognitive self-organization has become recognized as an important alternative.

Self-organization refers to the spontaneous emergence of order from nonlinear interactions among the components of a complex dynamic system. The meaning of "nonlinear" is twofold: (1) cognitive activities are viewed as reciprocal or recursive, as characterized by multiple feedback cycles, such that cause–effect relations are bidirectional or multidirectional; (2) effects in such systems are generally not linear functions of causes. Rather, they may be exponential, subject to threshold effects, and/or sensitive to damping or amplifying effects from other system components. Broadly defined, self-organization refers to the emergence of novel patterns or structures, the appearance of new levels of integration and organization in existing structures, and the spontaneous transition from states of lower order to states of higher order. Examples can be found in ecosystems (Kauffman 1993), social systems (Arrow et al. 2000), cortical systems (Bressler & Kelso 2001; Skarda & Freeman 1987), connectionist networks (Rumelhart et al. 1986a), morphogenesis (Goodwin 1993), and ontogenesis (Lewis 2000b; Thelen & Smith 1994), not to mention tennis, music, and sex. With respect to cognition, self-organization usually refers to the emergence and stabilization of psychological or neural configurations that correspond with (or represent) conditions in the world.

Why emphasize self-organization as a key aspect of dynamic cognitive – and cognitive–emotional – systems? The future states of a dynamic system are a function of its present state, as modified by its own activity (Vallacher & Nowak 1997; Van Gelder & Port 1995). Thus, cognitive systems construed as dynamic systems do not process information transduced from the outside world; they reconfigure themselves in response to an ongoing stream of sensory events (Van Gelder 1998; Varela et al. 1991). Most important, cognitive systems reconfigure themselves by achieving states of higher order through the synchronization of their constituents, both in real time, in response to perturbations, and in development, through experience-dependent change. In real time, coherent, macroscopic unities (e.g., schemas, expectancies, scripts, intentions, beliefs – and, arguably, appraisals) arise through the spontaneous coordination of microscopic constituents (e.g., Kelso 1995; Port & Van Gelder 1995). In development, novel skills and habits emerge, stabilize, and consolidate over months and years (Thelen & Smith 1994). The language of self-organization is well tailored, not only for modeling cognition in general, but also for modeling appraisals, along with the emotions that accompany them. I will now suggest a framework for analyzing appraisals as self-organizing phenomena.

https://doi.org/10.1017/S0140525X0500004X Published online by Cambridge University Press

3.2. Principles of self-organization applied to appraisal and emotion

In this section, I outline several principles of self-organizing systems and demonstrate their relevance for the emergence and consolidation of appraisals (along with accompanying emotional states – as detailed in the next section). I indicate where these or related ideas have been explored by other emotion theorists. Emergent appraisals (or appraisal–emotion amalgams) are construed as globally coherent states arising and stabilizing through nonlinear causal transactions among appraisal and emotion *constituents* (see Fig. 1).

3.2.1. Emergent order. Self-organizing systems show the spontaneous emergence of order out of (relative) disorder. Following some perturbation or trigger, novel forms arise without instruction or programming, based on interactions among the system elements themselves. Self-organizing systems start off with many degrees of freedom, but these are "used up" as orderliness emerges, through macroscopic constraints on the activity of individual elements. Self-organizing appraisals should thus become more orderly (i.e., synchronized, coherent) as they progress from initial to subsequent phases. This proposition is consistent with the idea of an evaluative structure consolidating during the latter phase of appraisal (Frijda 1993b), and with the locking in of meaning once appraisal dimensions are computed (Lazarus 1991; Ortony et al. 1988; Roseman 1984).

3.2.2. Positive feedback and self-amplification. Orderliness emerges in complex systems through a combination of self-augmenting (positive) and self-maintaining (negative) feedback processes. In positive feedback, a change in the activity of a subset of system elements feeds back to those elements and amplifies the change. Positive feedback relations thus "grow" new patterns of activation and recruit additional elements to them. Self-amplifying feedback can be modeled by autocatalysis – the tendency for some biochemical reactions to catalyze themselves recursively (Kauffman 1993; Prigogine & Stengers 1984) – producing exponential gradients of change. Although positive feedback is sometimes notorious for its capacity to break down existing structures, it is also a fundamental vehicle for the emergence of novel structures (Juarrero 1999). Consistent with this profile, appraisals and accompanying emotion states have rapid onsets in which many cognitive and emotional factors change together (Frijda 1993a). Positive feedback has been ignored by most emotion theorists, but Teasdale (1983) and Teasdale and Barnard (1993) have modeled the growth of depressive states based on positive feedback between negative appraisals such as self-blame and negative emotions such as sadness. Similarly, Mathews (1990) has suggested that threat-anxiety states may arise from feedback between attention to danger cues and anxious emotions.

3.2.3. Negative feedback, entrainment, and stability. As the elements of the system interact recursively, negative feedback takes over. A negative feedback loop is one in which the effects of one element on another (e.g., excitation) are compensated by reciprocal effects in the opposite direction (e.g., inhibition). Negative feedback loops attain a steady equilibrium or a continuous oscillation. This is because each opposing causal influence counteracts the other, either simultaneously or via ongoing adjustments. In either case, stabilization ensues. Thus, negative feedback stabilizes the growth and change initiated by positive feedback.

Stability anchored in negative feedback circuits can spread out across the whole system through entrainment (or synchronization) with other circuits. Thus, individual elements or groups of elements lose their independence and become embedded in a larger regime. Deviations are dampened and discrepancies resolved, contributing to the consolidation of global orderliness. Such functional synchronization can (but need not) take the form of temporal synchronization or phase synchrony, a theme addressed in some detail later. In general, self-stabilization is a property of many interactive networks, including cellular automata (Nowak et al. 2000) and constraint satisfaction networks (Thagard & Verbeurgt 1998). This tendency may be universal for adaptive systems, because their components must be coordinated to maintain their functionality (Ford 1987; Karoly 1993). Emotion theorists have not had much to say about negative feedback or synchronization, but social-personality psychologists have described discrepancy-reducing feedback processes as determinants of emotion in the short term and of personality styles in the long term (e.g., Carver & Scheier 1990; Higgins 1987). As well, some connectionist theorists have modeled beliefs, attitudes, impressions, and judgments as states of coherence based on interacting constraints (Kunda & Thagard 1996; Shultz & Lepper 1996), and Nerb and Spada (2001) developed a connectionist model of coherence relations among appraisal and emotion elements based on consistency among beliefs and desires.

3.2.4. Circular causality. Feedback is one form of nonlinear causation. A second form, termed *circular causality* (Haken 1977), describes bidirectional causation between different *levels* of a system. A coherent, higher-order form or function *causes* a particular pattern of coupling among lower-order elements, while this pattern simultaneously *causes* the higher-order form. The top-down flow of causation may be considered an emergent constraint (by the system as a whole) on the actions of the parts. Cognitive scientists have begun to model higher-order mental states such as intentionality and consciousness as emergent forms that constrain the activation of the psychological or neural constituents producing them (Freeman 1995; Juarrero 1999; Thompson & Varela 2001). In behavioral systems, functions such as walking or reaching are seen as maintaining the coordination of their perceptual and muscular constituents (Kelso 1995; Thelen & Smith 1994). Appraisals might also be considered higher-order functional or mental organizations that constrain their constituents. Superordinate appraisal themes, such as Lazarus's core relational themes (e.g., irrevocable loss), might hold subordinate, interacting appraisal processes in place, while those processes simultaneously fuel the superordinate. This idea is consistent with Frijda's (1993b) modeling of appraisal as a cumulative gestalt, and it is explicitly captured by Scherer's (2000) view of appraisal as a higher-order parameter "enslaving" its constituents (also based on Haken).

3.2.5. Increasing complexity. In self-organization, coordination among differentiated components of the system permits complexity as well as coherence. Increased organization allows more elaborate configurations of parts and functions, such as coordinated subsystems arranged in se-

https://doi.org/10.1017/S0140525X0500004X Published online by Cambridge University Press

quence or in parallel. An example from neuroscience is the emergence of perceptual binding, arising from the synchronization of neural firing patterns and allowing for complex perceptual processes such as figure–ground segregation and the sequencing of perceptual events (Engel et al. 2001). Several appraisal theorists describe secondary appraisal processes as enriching or supplementing primary appraisals with more elaborate content and meaning (e.g., Frijda 1993b; Smith & Kirby 2001). For example, an appraisal of someone as untrustworthy may start off relatively vague, but once it consolidates, it permits the formation of an intricate interpretation of the other as sneaky, malicious, and dangerous. Thus, the coordination of concepts and associations in orderly ensembles permits the construction of detailed cognitive models.

3.2.6. Multistability and stochasticity. The outcomes of self-organization are fixed or periodic (or chaotic) organizational patterns, called attractors, that endure for some period of time. The principle of multistability holds that many potential attractors coexist on the state space of an adaptive system (Kelso 1995), and systems rapidly move or "evolve" to one or another of them. Moreover, natural systems are usually thought to be influenced by stochastic forces, such that the attractor to which they gravitate is not fully predictable in advance. Appraisals do seem to fall into one of several identifiable attractors. Despite the many potential combinations of values on various appraisal dimensions, appraisal theorists stipulate a limited number of meaningful combinations, leading to a limited number of specific emotions. The core relational themes proposed by Smith and Lazarus (1993) tap these canonical combinations. Moreover, for an individual with a given emotion trait, only a small subset of these combinations occurs regularly (e.g., Horowitz 1998), and only one of these converges at a time.

3.2.7. Phase transitions. Because orderliness is limited to a small number of stable states, self-organizing systems move across thresholds of instability and jump abruptly to new stabilities. These jumps, sometimes called phase transitions or bifurcations, occur when system orderliness breaks down as the result of some perturbation and new patterns of organization rapidly self-amplify. Sudden shifts between incompatible appraisals may be analyzed in this way. Izard (1993) describes a rapid reappraisal that begins when a man is painfully hit from behind. He turns around angrily only to find that the person who hit him is wheelchair-bound, and he immediately reappraises this person as out of control and faultless. In this case, the reappraisal starts with one new element, leading rapidly to a switch in macroscopic organization.

3.2.8. Time scales and learning. Embedded and interacting time scales are typical of self-organizing systems in nature. In psychological systems, time scales typically include the emergence of behavior in *real time* and the emergence of behavioral traits or habits in *developmental time* (Thelen & Smith 1994). Importantly, processes at each scale influence processes at the other scale (Port & Van Gelder 1995). That is, structural changes in system constituents over development both result from and contribute to the emergence of coherent forms in real time. Time scale relations can be seen in the crystallization and automatization of appraisals across occasions (i.e., social learning) and in the constraints of this learning on real-time appraisals. Appraisal theorists avoid this perspective, but psychologists studying emotional and personality development find it crucial for explaining individual differences. For example, Izard and Malatesta (1987) postulate affective-cognitive structures, arising over a series of emotionally loaded situations, as the units of personality development that constrain future interpretations; and Dodge and Somberg (1987) show how several experiences of abuse lead children to develop a "hostile attributional bias" that filters subsequent experience. The vehicle for this learning is another highly general property of self-organizing systems: the tendency for system elements to change in structure as a result of their activity. This principle is built into learning algorithms for connectionist nets and is the basis of Hebbian synaptic strengthening.

3.3. A psychological model of self-organizing emotional interpretations

In this section, a psychological model of appraisal–emotion relations is grounded in the framework just set out. In a subsequent section, this model is revised and elaborated from a neurobiological perspective. As sketched in Figure 1, the model highlights bidirectional interactions among appraisal and emotion components, but it also traces the evolution of an emotion–appraisal state – or *emotional interpretation* (EI) – through several phases. The sequence begins with a trigger event that is defined by a change in the dynamics of the psychological system. Starting with this event, an initial phase of self-amplification represents the growth phase of an EI. This gives way to a phase of self-stabilization, in which global coherence is established and higher levels of complexity appear. Finally, coherent states that endure for some time give rise to learning, extending the influence of the present appraisal to future episodes.

Before going on, it may be useful to present an example of a self-organizing EI. In a number of Western countries there is growing concern with the phenomenon of "road rage" – intense anger coupled with an appraisal of another driver as one's enemy or tormentor, often leading to shouting, swearing, or physical violence. Self-organizing angry appraisals are not unusual in traffic, but road rage, because it is circumscribed, arbitrary, and extreme, provides a paradigm case. Mr. Smart slams on the brakes when noticing the proximity of the car in front. Anger arises initially from frustration, as Mr. Smart wants to keep driving fast, but also from a sense of violated entitlement: he is in the left lane and should not have to slow down. Fear may also be triggered by the close call, eliciting further anger because of an intermediate evaluation of unmanly helplessness. These emotions arise rapidly, but they are paralleled by a coemerging sense of the other driver as intentionally obstructive (and therefore blameworthy). Mr. Smart's highly focused visual attention, a derivative of anger, takes in the red color of the car ahead, as well as the expensive-looking design, and his anger is amplified by his sense of the unfairness of this show-off blocking his path (based on an implicit memory of some long-forgotten or fantasized rival). A stabilizing angry-anxious state, coupled with ruminative plans for vengeance (perhaps a blast of the horn), anchors attention to the head of the man in front. This lasts for a minute or two while Mr. Smart fashions and modifies plans to pass on the right. However, when the man peers over his shoul-

https://doi.org/10.1017/S0140525X0500004X Published online by Cambridge University Press

der, Mr. Smart evaluates this act as a taunt, generating shame and anger in an elaborated appraisal of humiliation, and calling for extreme action to save his self-image from further subjugation.

3.3.1. Trigger phase. Most appraisal theorists would agree in principle that appraisal–emotion episodes begin with a triggering event. In DS terms, the orderly behavior of the system is interrupted by a perturbation, resulting in a rapid loss of orderliness and an increase in sensitivity to the environment. The perturbing event must have an impact on one or more system elements, or it is invisible to the system. This impact works through a process of nucleation in which the affected element enters into reciprocal, self-enhancing interactions with neighboring elements to create a core proto-organization (Prigogine & Stengers 1984). Thus, a trigger marks a phase transition, characterized by sudden change and temporary disorder as the system switches to a new organization.

Living systems are like taut springs, ready to respond to small perturbations that are biologically meaningful (Kauffman 1995). In the triggering of an EI, the perturbation can be any sensory event, a perceptual or cognitive event (e.g., an image, association, or memory), or a change in arousal or affect (Izard 1993), as long as it induces a self-amplifying interaction among appraisal and emotion elements. It is difficult to predict the conditions for self-amplification, but background psychological and physiological states, both cognitive and emotional, must act as control parameters that adjust sensitivities that determine what is meaningful. A depressive mood makes one insensitive to cheerful events but sensitive to negative events (Teasdale & Barnard 1993), and anxiety biases attention to threat cues (Mathews & MacLeod 1994). Mr. Smart's irritation with traffic served this background function.

A trigger can be used to define the first moment (time-0) of an emotional episode, but it can also occur at any point in an ongoing appraisal–emotion stream. This collapses the distinction between circumscribed emotion episodes and ongoing streams of appraisal–emotion events. Given this assumption, triggers can change existing appraisals by either replacing them or building onto them, depending on the present context and inner state characteristics. For example, an angry EI, in which anger and blame are coupled, can be interrupted by a fear trigger (e.g., cued by awareness of potential retaliation), but it can also be elevated to a state of more intense rage (e.g., if insult is added to injury, as in the case of Mr. Smart). Such cascading triggers can initiate a branching pathway of EI evolution in real time. In fact, the art of fiction writers and movie directors is to deliberately shape the evolution of such pathways.

3.3.2. Self-amplification phase. For an EI to "take over" the psychological system, a phase of self-amplification must follow perturbation and nucleation. Appraisal and emotion constituents may now interact in positive feedback loops, such that the activation of any element leads to the activation of other elements in recursive cycles of increasing magnitude. According to appraisal theorists, appraisal elements (e.g., pre-attentive perceptions of an approaching other) stimulate emotions. As shown by Mathews and others, emotions guide the focus of attention and recall to those features that are emotionally relevant (e.g., facial and postural cues, or the head movements of the driver in front of Mr. Smart). In the present model, this increase in cognitive activities dedicated to emotion-relevant cues strengthens the emergent appraisal and amplifies emotional activation. Thus, perceptual, emotional, and attentional processes amplify one another (positive feedback) but at the same time begin to tune or constrain each other (negative feedback) as the EI grows. Emotional processes may also recruit executive processes in preparation for action (e.g., Mr. Smart's plan to switch lanes), as well as retrieval processes that make sense of the event and guide further emotional activation (e.g., memories of intimidating other drivers). Positive feedback thus "grows" appraisal–emotion processes that capture the psychological system.

While positive feedback predominates, systems are highly sensitive, and small deviations may be rapidly amplified. A raised eyebrow can shift the evolving appraisal from affiliation to humiliation, at least until a coherent interpretation locks into place. During this early phase, appraisals should also be sensitive to the background activation of system elements, including baseline emotion or mood (cf. Frijda 1993a). However, psychological systems, like well-trained neural networks, generally remain in an unstable phase only briefly, then settle to a particular region of the state space and remain there.

The principles of bidirectional and multiple causation in the present model help to resolve some of the problems encountered by mainstream emotion theories. According to the model, appraisal activities and emotional response activities *cause one another*, each activating, propelling, and guiding the other, reciprocally and recursively. Emotional events are granted full causal status, and their effects are felt at or near the beginning of the appraisal sequence. Emotional events are part of the causal chain that contributes to the evolution and consolidation of the appraisal pattern – hence, what evolves is not just an appraisal but an emotion–appraisal amalgam or "emotional interpretation."

3.3.3. Self-stabilization phase. When negative feedback overtakes the system dynamics, change decreases and continuity increases. Appraisal elements become coupled in coherent ensembles or meaningful wholes that are entrained with emotional states. The negative feedback properties responsible for this stabilization can be exemplified by the interaction of anxiety and vigilance. Anxiety initially increases vigilance, but heightened vigilance improves one's grasp of a situation, thus diminishing anxiety. With anxiety decreasing, vigilance also decreases, leading in turn to uncertainty and then increased anxiety once again. This oscillatory pattern may stabilize and endure. However, because such a pattern is ultimately noisy and difficult to measure, one might simply say that a moderate level of vigilance becomes coupled with a moderate level of anxiety. This stabilizing anxious-vigilant mode can now become entrained with other constellations (e.g., a cognitive evaluation of "keeping the lid on"), thus enhancing coherence across multiple subsystems.

Spreading coherence in connectionist networks is modeled by constraint satisfaction solutions involving multiple cognitive elements. To model the coherence of an EI, these constraints must also include emotional elements. The resolution of conceptual discrepancies has been proposed to elicit emotions, but this process is usually modeled as unidirectional. With bidirectional causality, the resolution of discrepancies would not only elicit or stabilize emotions but

https://doi.org/10.1017/S0140525X0500004X Published online by Cambridge University Press

would also be *caused* by emotions that support particular meanings (see Nerb & Spada 2001). Thus, emotions are necessary to cement emerging interpretations (as when anger fuels Mr. Smart's interpretation of blameworthiness), and emotions are maintained by those same interpretations (as when blame perpetuates Mr. Smart's anger), locking cognition and emotion into an enduring resonance.

Despite the ambiguous nature of the world, and even in the presence of conflicting cues, appraisal–emotion configurations resolve into stable patterns. Indeed, appraisals have been found to resist change in paradigms that use perceptual disengagement in the short run (Fox et al. 2002) and rumination in the long run (Martin & Tesser 1996; Nolen-Hoeksema & Morrow 1991). The possibility of several alternative resolutions is consistent with the principle of multistability in dynamic systems, but the consolidation of a particular "solution" from competing ensembles is rapid and nonlinear (cf. Thagard & Verbeurgt 1998; Tononi & Edelman 1998). Circular causality (Haken 1977) may be a key factor in this consolidation. Through circular causality, top-down and bottom-up causal processes become mutually entrained, and a specific EI coheres and stabilizes. Thus, when a superordinate appraisal (e.g., other driver as deliberate foe) begins to cohere, it may "enslave" its constituents (red car, thinks he owns the road, self as subordinate, shame, anger), while the interactions of the constituents that contribute to that particular appraisal are perpetuated and strengthened. This kind of explanation has proven useful for modeling the emergence of global appraisals (Lewis 1996; Scherer 2000).

The process of cognitive-emotional stabilization may also be guided by functions related to action. Emotion theorists emphasize the role of emotion in propelling and constraining action, and *action tendencies* are often considered aspects of emotion. However, the principle of bidirectional causation suggests that action tendencies may feed back to appraisal and, perhaps, to other aspects of emotion as well. Psychological functions that support action include selective attention, planning, and monitoring one's interaction with the environment.

Finally, DS principles suggest that the consolidation of coherent emotion–appraisal states is necessary for complexification, allowing appraisals to become more elaborate and articulated. Before Mr. Smart can develop a full character sketch of the driver blocking his path, and a plan for confronting him, his cognitive processes must become coordinated and integrated with emotion. Elements in attention, working memory, and planning must cohere to produce a detailed mental model or script. However, complexification can also generate new scenarios or contingencies, potentially triggering a new EI at any point. Rehearsing a plan for retaliation can bring fearful images to mind, sometimes evolving into an appraisal of helplessness and self-doubt. Perhaps the most realistic portrait of a self-stabilizing EI includes several well-worn interpretations that fluctuate intermittently, as familiar images and emotions drive the system between available attractors.

3.3.4. Learning. Enduring biases, beliefs, traits, and emotional habits have been well documented in emotion research, but their relation to real-time appraisal processes has not been fully explored. From a developmental perspective, recurrent episodes of emotion–appraisal contribute to the formation of individual characteristics (Izard 1984; Lewis 1995; Malatesta & Wilson 1988). Thus, each appraisal episode provides the occasion for learning cognition–emotion associations that tend to recur on future occasions. According to the present model, the self-stabilization phase of an EI is the necessary precondition for this learning. Once appraisals have stabilized, interpretations, action plans, and expectancies endure for some period of time, as mediated by coupled cognitive and emotional elements. These enduring couplings seem necessary to strengthen the connections responsible for learning. As noted earlier, self-organizing systems, including brains, show increased connectivity among elements that are reciprocally activated (coupled) in real time.

A good deal of organization already exists in the cognitive-emotional system prior to the emergence of a new EI, and much of this organization derives from learning across past emotional episodes. Mr. Smart's road rage must have grown from a lineage of episodes of frustration and anger, although he may well have a low frustration tolerance by temperament. Associative learning is often considered the principal mechanism for laying down this organization over time. According to the present model, learning is a long-term organization produced by enduring EIs, and the consolidation of order within an individual EI is a short-term organization based in part on learning. As is typical of self-organizing systems, the accumulation of order at two very different time scales is considered interdependent.

3.4. Summary and conclusions

In this section, a framework was established using principles of self-organizing dynamic systems applied to appraisal and emotion. Then a model was developed with specific emphasis on emotion–appraisal states, or emotional interpretations (EIs), as emergent forms, and on the self-amplifying and self-stabilizing causal processes that give rise to them. Rather than sequences of appraisal steps producing emotions, appraisal–emotion processes were viewed as bidirectional interactions among cognitive and emotional constituents, giving rise to new organizational regimes, shifting from growth to coherence in real time, and setting the occasion for complexification, action, and learning.

However, this modeling exercise leaves many questions unanswered. When and how do perturbations trigger new appraisals? How does self-amplification and sensitivity give way to self-stabilization and coherence? How does the superordinate EI constrain the component interactions that give rise to it? What mechanisms relate action and learning to appraisal and emotion? These and other questions point to a major gap, or core ambiguity, in the model presented thus far: What exactly are appraisal elements or constituents, what are the "elements" of emotion, and how can one define these parts in relation to the wholes emerging from their interaction? At the level of psychological modeling, one must describe constituents of appraisals with terms such as "perception," "evaluation," and "attention," whereas emotion constituents, when discussed at all, are defined in terms such as "arousal" or "action tendencies." These terms hardly describe constituents, however. They describe whole, global functions or processes in themselves. This problem reflects a general tendency for psychological theory to gravitate to a level of description that is superordinate, global, and functional. This tendency makes it difficult to concretize causal relations be-

https://doi.org/10.1017/S0140525X0500004X Published online by Cambridge University Press

tween parts and wholes, and thus to fashion true explanations.

The principles of self-organizing systems have been useful for modeling the evolution of global states on the basis of interacting cognitive and emotional constituents. However, much of the detail has yet to be filled in, and the psychological level of description has little more to offer. I now turn to emotional neurobiology to reconceptualize appraisal and emotion "parts," identify their mechanisms of interaction, put flesh on the bones of the model developed thus far, and demonstrate a conceptual bridge between emotion theory and neurobiology more generally.

4. The neurobiology of emotion and appraisal

To link emotion theory with emotional neurobiology through the application of DS principles, it seems necessary to identify the anatomy and function of brain regions (or structures) that mediate the various constituents of appraisal and emotion, and then to identify the processes that connect these regions in global activities that correspond with psychologically meaningful wholes. Neurobiology provides specific and detailed accounts of what goes on within circumscribed regions and between closely related regions, but it has a more difficult time specifying part–whole relations that integrate activities over the entire brain. In this section and the next, these part–whole relations are highlighted by visualizing neural processes through a DS lens.

It should be emphasized that much theoretical and empirical work in neuroscience has utilized DS principles. For example, cortical activities have been modeled as self-organizing processes of synchronization that yield coherence and/or complexity (Bressler & Kelso 2001; Érdi & Barna 1984; Meyer-Lindenberg et al. 1998; 2002; Nunez 2000; Skarda & Freeman 1987; Skinner & Molnar 1999; Varela et al. 2001), transitions have been studied as instabilities or phase transitions triggered by perturbations (Freeman 1995; Thatcher 1998), individual development has been modeled as emergent patterning of synaptic networks (Edelman 1987; Edelman & Tononi 1997; Harkness & Tucker 2000; Post & Weiss 1997; Schore 2000), and circular causality has been identified between superordinate mental states (e.g., attention, expectancy) and subordinate neural events (Engel et al. 2001; Freeman 1995; Szentagothai 1993; Thompson & Varela 2001). These and other applications suggest many avenues for exploring what is surely the ultimate complex system.

Most of these accounts model phase synchrony or time-based correlations among neural populations in the cortex, and they provide electrophysiologial data in support of this modeling. However, emotional processes necessarily involve subcortical systems, and there has been much less work demonstrating synchronization across multiple levels of the neuroaxis. Therefore, the present analysis of interacting cortical and subcortical systems is grounded more in anatomical and functional data than electrophysiological data; and neural integration is defined and discussed in functional terms before exploring the intriguing possibilities of temporal synchronization.

4.1. Working definitions

The main argument that has been advanced so far, and that will now be extended through a neurobiological analysis, is that reciprocal and recursive relations among appraisal and emotion components lead to the emergence of appraisal–emotion wholes, and that neither whole (i.e., fully articulated) emotions nor whole appraisals can exist in isolation from each other. In Section 2, I suggested that appraisal components include (1) perception, (2) evaluation, (3) attention, (4) memory, and (5) higher-order executive functions such as planning and reflection. I also noted that these terms are inadequate for any real explanation relating parts to wholes, because they denote global functions or processes themselves. Defining emotion parts meets with the same difficulties, and many emotion theorists avoid this conundrum by considering emotions as monolithic wholes. However, it seems unreasonable to treat emotions as wholes when so much attention is given to appraisal components. There are several features that are usually considered indispensable to emotion, and these are useful designators for emotion components. They include (1) arousal, (2) action tendencies, (3) attentional orientation, and (4) affective feeling. It is also generally agreed that these functions rely on physiological changes throughout the brain and body.

This parsing of appraisal and emotion components, though not without its problems, provides a segue to a neural analysis. The next step is to map these terms onto neural systems whose interactions can be described in detail. However, an additional difficulty proves to be the anatomical and functional overlap among systems that subserve appraisal and emotion constituents. In fact, many brain systems up and down the neuroaxis can be included under definitions of both (Lane et al. 2000; Panksepp 1991). This ambiguity may be responsible for ongoing debates over which functions to assign to which systems. For example, the rapid amygdala (AM) response to information from the thalamus can be viewed as an affective computation (LeDoux 1989), but it can also be viewed as a preliminary appraisal (Lazarus 1999; Scherer 1993b; see Parrott & Schulkin 1993, and LeDoux 1993, for a discussion of this issue). The purpose of this article is not to further this debate, but to show how a variety of neural subsystems interact to produce appraisal and emotion processes that become coupled in a macroscopic emotion–appraisal state. Nevertheless, the distinction between appraisal and emotion, both at the level of parts and the level of wholes, continues to be problematic, and I will address it in more detail later.

The appraisal components listed earlier – perception, evaluation, attention, memory, and planning/reflection – are mediated by a variety of neural systems from the brain stem to the cerebral cortex. Primitive sensory and evaluative processes are rooted in the midbrain, whereas higher cognitive functions of attention, memory, and planning correspond chiefly with limbic and cortical systems (though they require the participation of diencephalic and brainstem structures as well). With respect to emotion, arousal and action tendencies, as well as the control of physiological response systems, are generally assigned to brainstem and hypothalamic structures, attentional orientation involves links between these structures and corticolimbic systems, and affective feeling probably involves brainstem, paralimbic (e.g., cingulate), and prefrontal cortical structures (Damasio 1999; Panksepp 1998a). Thus, both appraisal and emotion components are mediated by systems splayed out along the neuroaxis, tracing an evolutionary path from more primitive to more recent acquisitions. Nevertheless, we sometimes associate appraisal with higher

https://doi.org/10.1017/S0140525X0500004X Published online by Cambridge University Press

(e.g., corticolimbic) systems. Although there is no adequate justification for this, appraisal processes in real time quickly become elaborated through corticolimbic activities that override or control more primitive structures (e.g., the superior and inferior colliculi; Rafal 2002). In contrast, our construals of emotional processes, even as they extend cortically, continue to highlight hypothalamic, brainstem, and bodily activities that mediate feelings and action impulses. This rough parsing of appraisal and emotion by brain region says more about our definitional habits than about fundamental neuroanatomical divisions; but it does agree with the thrust of models proposed by many investigators in affective neuroscience (e.g., Buck 1999; Damasio 1999; Davidson & Irwin 1999; Panksepp 1998a; 2003). It also corresponds with Panksepp's (1998a) metaphor of emotion systems as branching trees, with their roots in the brain stem and their branches extending upward to the cortex. According to this analogy, higher cognitive aspects of appraisal may refine core emotional meanings, extending both spatially (toward the cortex) and temporally (as emotion episodes unfold in time).

I now describe the location, structure, and function of various neural systems that are likely to mediate appraisal and emotion components (see Fig. 2). These are meant to be summary descriptions, providing just enough detail to flesh out the DS-inspired framework and model developed in the previous sections. The selection of systems is also meant to be provisional, supplying a reasonable but not exhaustive list of candidates for each psychological function. A number of neural systems are described in this section, and their bidirectional and emergent interactions are modeled in the next section. Rather than a detailed neurobiological theory, these sections present a framework for thinking differently about neural processes underpinning appraisal and emotion.

4.2. Anatomy and function of systems mediating appraisal

4.2.1. Perception. Sensory information is mediated by the thalamus (among other structures) and then proceeds to various cortical and subcortical systems. Primitive perceptual processes such as motion detection and object direction are mediated by midbrain structures such as the superior and inferior colliculi. More advanced perceptual processes take place in the primary sensory and association areas of the posterior cortex. Unimodal perceptual streams are integrated in temporal and parietal association areas, and higher-order perception, requiring orienting, anticipation, and integration with motor systems, involves bidirectional links between these areas and the prefrontal cortex (PFC). For example, the frontal eye fields control gaze direction, whereas other prefrontal systems mediate the anticipation and organization of information over time. While sensory information is being projected forward to these frontal systems, back-projections control the sensory pickup of significant aspects of the environment (see Fuster 2002, for a review).

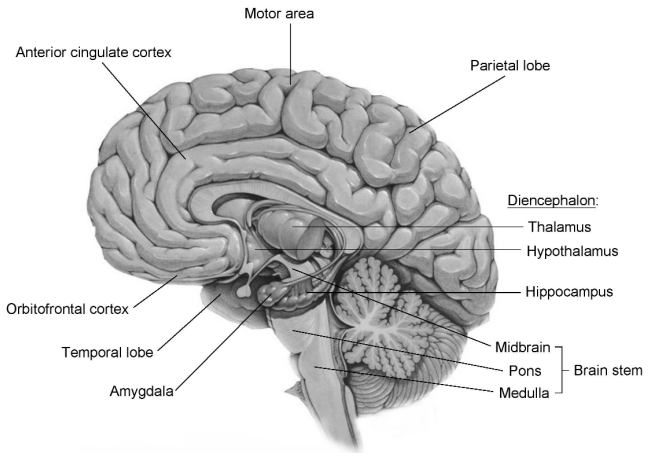


Figure 2. Some of the brain regions and structures mediating emotion and appraisal. The brain image is used with permission of Andrews & Associates/Custom Medical Stock Photo.

4.2.2. Evaluation. Evaluation is a slippery concept covering functions that extend from higher-level attentional systems such as the orbitofrontal cortex (OFC) and anterior cingulate cortex (ACC), to more basic pattern-matching functions of the amygdala (AM) and ventral striatum, to brainstem systems (e.g., the colliculi) that register innately defined events such as motion and rapid approach. The amygdala is a group of nuclei embedded within the temporal lobe of each hemisphere. It is a central limbic structure connected with many other limbic, diencephalic, and cortical areas, and it has long been known to form and maintain associations between neutral stimuli and motivationally tagged stimuli (i.e., emotional conditioning), especially in fear, but in other emotions (including positive ones) as well (LeDoux 1995a; Öhman 1993; Rolls 1999). The striatum, nested around the diencephalon and midbrain, includes various structures that are necessary for switching or selecting motor plans and for incentive motivation necessary for carrying out those plans. The ventral striatum (especially the nucleus accumbens [NAS]) evaluates and responds to conditioned and unconditioned stimuli (Cardinal et al. 2002) and is centrally involved in incentive motivation (Depue & Collins 1999).

4.2.3. Attention. While some evaluative processes are preattentive (Öhman 1988), others rely on attention. Higher-order attentional processes are mediated by prefrontal "executive" cortical systems, including the ventromedial PFC, dorsolateral prefrontal cortex (DLPFC), orbitofrontal cortex (OFC), and anterior cingulate cortex (ACC). More basic attentional processes such as orienting and monitoring are mediated by more posterior (posterior cingulate, parietal, and temporal) orienting systems (e.g., Posner et al. 1988). Most important for now are two "paralimbic" prefrontal systems (adjacent to limbic systems and containing more primitive cell types) considered to be critical for motivated attention (Barbas & Pandya 1989). The first is the OFC on the ventral surface of the PFC. The OFC encodes and holds attention to context-specific, motivationally relevant contingencies, both learned and unlearned (Rolls 1999). The OFC is thought to extend or build onto the more basic conditioning functions of the AM, through reciprocal interactions between them (Cardinal et al. 2002). Specifically, it is far more flexible than the AM, it "attends to" changes in the hedonic valence of anticipated events (Hikosaka & Watanabe 2000; Rolls 1999), and it can be seen as holding and recording "implicit appraisals" of motivationally relevant situations (Schore 1994). Its downstream

https://doi.org/10.1017/S0140525X0500004X Published online by Cambridge University Press

connections are also integral to emotional states, and it has been shown to "regulate" emotion by inhibiting AM activation (Davidson 1998; Davidson et al. 2000). The second attentional system critical for appraisal is the ACC, which is found on the medial surfaces of the PFC. Its functions include monitoring and evaluating potential actions, monitoring and resolving conflicts, error detection, and so forth (Carter et al. 2000; Gehring et al. 1993; van Veen et al. 2001). The executive system mediated by the (dorsal) ACC is characterized by voluntary choice and is central for directed attention and for learning (Frith et al. 1991; Gemba et al. 1986). However, the more ventral regions of this complex system are also involved in emotion, as discussed later. The OFC and ACC are highly interconnected. Both belong to a "paralimbic ring" (including temporal regions connected to the hippocampus) – an ensemble of primary importance for emotion, motivation, and memory (Mesulam 2002). In fact, both the OFC and ACC have been shown to be central loci for the interaction of attention and emotion (Barbas 1995; Lane et al. 1998). These systems also interact with dorsolateral prefrontal regions responsible for planning and working memory (Mesulam 2002).

4.2.4. Memory. There are a number of diverse memory systems in the brain, and their functions and interactions are not fully understood. Explicit, semantic, or episodic memories are mediated by the hippocampus (HPC), nearby regions of temporal cortex, and related diencephalic structures (Aggleton & Brown 1999). However, the coding, retrieval, and organization of this information relies on prefrontal cortical systems (see Shimamura 2002, for a review). Of particular importance is the lateral PFC, which is implicated in working memory. Working memory probably involves bidirectional interactions between lateral PFC and HPC systems, by which cues are retrieved, sorted, and organized temporally (Moscovitch 1992), while associations are activated in other cortical and subcortical regions. Other structures mediate different kinds of memory. These include (1) the OFC, which encodes and maintains changes in reinforcement contingencies (Rolls 1999) and which may function as an independent working memory system in the presence of emotion (Luu et al. 1998); (2) the ACC, involved in learning the demands of novel tasks (Gemba et al. 1986); (3) the amygdala (AM), responsible for emotional conditioning; and connections between the AM and ventral striatum that sensitize responding relative to that conditioning (see Cardinal et al. 2002). The functions of some of these systems (e.g., OFC, AM) are characterized as implicit or associative memory, in contrast to the conscious and explicit memory functions mediated by the lateral PFC and HPC.

4.2.5. Planning. The appraisal functions of planning, strategy formation, and reflection probably involve all or most prefrontal executive systems, but particularly the dorsolateral, ventrolateral, ventromedial, anterior prefrontal, and ACC systems (Shallice 2002). These systems are capable of organizing information temporally in working memory, monitoring and checking this information, relating the information to present goals and motivational concerns, comparing and contrasting alternative action sequences, and consciously selecting among available choices. The ACC is particularly important for intentional monitoring and choosing among potential actions (Luu & Tucker 2002; van Veen & Carter 2002), partly through its connections to the DLPFC and HPC. Planning also requires activation of the supplementary motor area, closely connected to the ACC, where action plans are volitionally executed (Goldberg 1985). Planning and strategizing also make use of striatal structures, such as the nucleus accumbens (NAS), that narrow attention to incentive events and organize action sequences for behaving accordingly (Depue & Collins 1999). Some theorists emphasize left-hemisphere involvement in action planning, because its analytical/sequential processing style facilitates the organization of discrete steps in a complex motor sequence (e.g., Tucker & Williamson 1984).

4.3. Anatomy and function of systems mediating emotion

4.3.1. Arousal. Emotions are sometimes defined as action tendencies, and any form of action requires the selective arousal of various brain and body systems. In fact, differences in arousal have been used to classify emotions and explore their biological underpinnings (Bradley & Lang 2000). Arousal at the brain level is generally construed as activation of neural systems. Where does this activation come from? Neurotransmitters that modulate synaptic processes extrinsic to their site of origin, sometimes called neuromodulators, are released from cell bodies in the brain stem (BS) and basal forebrain (BFB), along "ascending" pathways. Some of these cell bodies are organized in nuclei (e.g., nucleus basalis of Meynert, locus coeruleus) that have been collectively referred to as the reticular activating system, and whose primary function is indeed to activate neural processes. These cells release dopamine, norepinephrine, acetylcholine, and serotonin to terminal sites in all limbic, striatal, and cortical areas, including the AM and HPC (Izquierdo 1997) and prefrontal cortical systems (Fuster 1996). Their action can thus be construed as global or holistic (Panksepp 1998a). However, that does not mean that they act uniformly throughout the brain, or that neuromodulators as a group simply arouse. Rather, they have specific effects differentiated by the type of neurochemical, the nature of the task or situation, the region of brain in question, and the particular receptor type receiving the input (Arnsten & Robbins 2002). Most important for the present argument, these chemicals are released from BS/BFB areas in response to signals from the AM, OFC, ACC, and other structures involved in the appraisal of emotion-relevant events. Thus, while their activity may not be viewed as "emotional" per se, their participation in the activation of emotional responses is unquestionable.

Arousal of bodily systems is also a critical component of emotions, as it is necessary to prepare for and support the behaviors they induce. States of bodily arousal are tapped by psychophysiological measures such as heart rate and skin conductance, and comparisons of such measures with PET and fMRI data suggest that physiological arousal is correlated with increases in brain activity induced by emotional stimuli (see Bradley & Lang 2000, for a review). Bodily arousal systems are controlled in large part through the autonomic nervous system (ANS), with arousal functions mediated by the sympathetic system and compensatory energy-preserving functions mediated by the parasympathetic system. More specific forms of bodily arousal are mediated by the endocrine system, through which hormones

https://doi.org/10.1017/S0140525X0500004X Published online by Cambridge University Press

circulate in the bloodstream. Brainstem and hypothalamic structures serve as controls of the ANS, via direct nerve pathways to organs, muscles, and skin. Hypothalamic nuclei also control – through chemicals released by the pituitary – the action of circulating hormones that act on specific glands as well as many other tissues. Taken together, these systems regulate blood flow, muscle contraction, digestion, temperature, respiration, perspiration, and behavioral tendencies related to sex, nurturance, and fight or flight, partly through direct effects and partly through the coordinated release of hormones by endocrine systems throughout the body. Conversely, many bodily sensations resulting from these changes are fed back to brainstem and hypothalamic systems (Thayer & Lane 2000), resulting in rudimentary emotional feelings, and to higher paralimbic and cortical systems that subserve emotional consciousness, as reviewed later.

4.3.2. Action tendencies. Emotion theorists ascribe specific action tendencies to each emotion. The activation of an action tendency (e.g., escape, approach, nurture, explore) and the urge to act accordingly can be mapped onto brainstem motor control systems as well as chemical support processes mediated by various brain regions. Phylogenetically old systems that control primitive, packaged response patterns (e.g., defensive and attack behavior, sexual behavior, vigilance, facial expressions, freezing) are located in the brain stem and in the basal forebrain (BFB). These systems are linked through the hypothalamus to limbic, striatal, and cortical systems; but they orchestrate emotional behavior even in the absence of higher brain systems, as demonstrated by the "sham rage" seen in decorticate animals. They have therefore been referred to as basic emotion circuits (Panksepp 1998a). For example, different regions of the periaqueductal grey are associated with freezing responses related to fear and aggressive responses related to rage (e.g., Fanselow 1994). These response tendencies are then modified and elaborated by higher brain systems in normal emotional behavior (Panksepp 1998a).

Limbic structures have traditionally been associated with emotion, especially the amygdala (AM), which is critical for emotional conditioning (LeDoux 1995a), as reviewed earlier. Connections from the AM to lower (hypothalamic and BS) structures are credited with activating motivational response systems that trigger action modes as well as neurochemical support, given current stimulus events. Other limbic structures, including the septal and hippocampal systems and the "paralimbic" cingulate gyrus, may be necessary to support emotional behaviors (e.g., play, sex, nurturance) and memories (e.g., Gray 1982; MacLean 1993). The motivational and behavioral tendencies mediated by limbic structures may be qualitatively distinct and fall into basic emotion categories (Buck 1999), or they may be more globally categorized as general-purpose response systems (Gray 1982). However, limbic structures require the participation of lower structures to activate emotional behavior, whereas the converse is not true (Panksepp 1998a). Perhaps most important, limbic structures integrate emotionally relevant perceptual and action orientations through their connections to other structures, thus regulating emotional responses in a coherent fashion.

BS nuclei that control action tendencies are closely related to neuromodulator release, discussed previously. However, much more specialized neurochemicals, called neuropeptides, are released from the hypothalamus and other brain areas in correspondence with these tendencies. Neuropeptides support specific mammalian agendas (e.g., nurturance, predation, defense, play) and suppress others, and their action is global, as to the diversity of regions affected, but specific in the behavioral and physiological states they induce (Panksepp 1998a). The reception of neuropeptides by brain and bodily systems may serve both the urge to act according to these action agendas and the dedication of resources to see them through effectively. Some neuropeptides are released by secretion and diffusion (into the blood and ventricular circulation). Their action tends to be less immediate but enduring (Kandel et al. 2000). Such effects are thought to support lasting states of behavioral regulation that are key aspects of mood (Panksepp 1993; 1998a). Other neuropeptides act rapidly, and may contribute very differently to emotional processes. While there is enormous variety in the types and action modes of neuropeptides, some of the most important peptides, acting either singly or in coordination with others, recruit multiple areas to a unitary action tendency corresponding with a particular emotion, and maintain that tendency at the expense of others.

4.3.3. Attentional orientation. Attentional systems were described in some detail earlier. Although the characteristics of these systems support appraisal processes in diverse ways, the global orientation of attention to particular features of the world, cutting across all these systems, is considered a fundamental constituent of emotion. The term "motivated attention" expresses the idea that the beam of attention is focused on whatever is emotionally compelling (Derryberry & Tucker 1994), often interrupting or replacing existing attentional frames (Oatley & Johnson-Laird 1987). Neural systems that participate in motivated attention are widespread. First, the superior and inferior colliculi of the midbrain are essentially fixated on changes in the location or proximity of objects or animals with emotional relevance. At the limbic level, the AM responds rapidly to the emotional valence of stimuli and accordingly controls attentional processes, such as those required for continuous-performance tasks (Holland et al. 2000). Neurons in the nucleus accumbens (NAS) are selectively activated in response to contextually relevant targets, entraining other systems such as the OFC to attend to emotionally relevant events (Depue & Collins 1999). As discussed later, these attentional control mechanisms require communication with other brain regions to carry out their functions.

Higher structures such as prefrontal and paralimbic executive systems contribute to attentional orientation and action readiness simultaneously. These systems are recruited by the emotional responses mediated by lower systems, and they regulate these responses in part by directing attention toward viable behavioral options (Barbas 1995; Bechara et al. 2000; Davidson & Irwin 1999). The OFC, insular cortex, and ACC are key structures that maintain motivated attention while modulating emotionally relevant responses (Barbas 2000; Hariri et al. 2000; Lane et al. 1998). For example, the ACC can shift attention rapidly to emotional stimuli in preparation for an immediate motor response (Carreti et al. 2001). Finally, switches in the relative activation of prefrontal systems suggest direct emotional control of attention. For example, contrasting changes in the activation of dorsal and ventral prefrontal regions indicate that the dorsal ACC and ventral

https://doi.org/10.1017/S0140525X0500004X Published online by Cambridge University Press

ACC (including nearby regions of the ventral PFC) compete for the regulation of attention, based on the presence or intensity of negative emotion (Bush et al. 2000) – especially anxiety (Drevets & Raichle 1998) and sadness (Mayberg et al. 1999).

4.3.4. Feeling and consciousness. The final component of emotion is the feeling, phenomenological, or conscious apprehension of emotional states and processes. Many investigators believe that the affective feeling of emotion is a critical aspect of its motivational and adaptive properties (Damasio 1999; Izard 1984; Lane 2000; Panksepp 1998a), and Panksepp argues convincingly that this pertains to nonhuman mammals as well. Lane (2000) provides a scheme for mapping increasingly inclusive levels of feeling/consciousness onto brain regions of increasing sophistication. Visceral processes, such as the contraction of muscles in the gut, are apprehended directly by brainstem mechanisms, constituting a feedback circuit that couples the production and perception of autonomic changes (Thayer & Lane 2000). Panksepp (1998a) views this level of integration as being fundamental to a core sense of self, or primary process consciousness, specifically mediated by the colliculi, periaqueductal grey, and nearby areas, and involving motor maps of the body as much as sensory maps. Visceral processes and perception of action tendencies may also be apprehended by diencephalic and limbic structures (Lane 2000).

Higher-order consciousness of emotional processes is attributed to various cortical and paralimbic structures. The ACC is especially implicated in emotional awareness, possibly as an extension of its more general role in executive attention and response selection (Lane et al. 1998). The ACC also registers the emotional quality of physical pain and, according to recent work, of psychic pain as well (Eisenberger et al. 2003). However, the ventral (i.e., rostral, subgenual) ACC and ventromedial PFC may be focal points for conscious representations of emotional state (Lane 2000). This possibility is consistent with Mayberg et al.'s (1999) findings that ventral ACC activation corresponds with sadness and depressed mood, as well as Drevets and Raichle's (1998) findings of increased ventral activation during anxiety. Closely related regions, the OFC and insula, are also thought to process viscerosensory information as well as other physiological signals from the body, and a number of imaging studies have associated these regions with subjective emotional feelings (see Craig 2002, for a review and synthesis). As discussed by Craig, Schore (1994), and others, the right hemisphere seems particularly important for apprehending these feeling states.

4.4. Conclusions

This analysis of neural systems underlying appraisal and emotion components leads to two conclusions. First, as noted earlier, many systems mediate functions that can be assigned either to emotion or to appraisal. Or, to put it differently, many neural systems that become activated in appraisal also take part in emotional functions, and systems that generate emotional responses may also serve appraisal functions. For example, the AM's dual role in appraisal and emotional response has been described, and the ACC's involvement in attention and planning is difficult to differentiate from its role in attentional *orientation* and emotional consciousness. This consideration makes it tempting to put aside the distinction between appraisal and emotion categories, at least for some important neural systems. Second, each appraisal and emotion "component," defined at a psychological level, becomes a distributed system in itself, or even a collection of fairly distinct systems, when analyzed at the neural level. Thus, there are many brain "parts" that mediate emotional feeling, many brain systems that subserve different kinds of attention, and so forth. This consideration makes it even more tempting to abandon psychological definitions, this time of the components themselves, and focus instead on neural entities as the "parts" that interact in self-organizing states. As suggested earlier, neurobiology seems much better than psychology at specifying interacting parts. Nevertheless, neural parts are only interesting insofar as they perform functions that cross the line into psychological description – even if, or especially if, such description is necessarily more global. The coordination of neural parts in the service of global appraisal and emotion functions is examined next.

5. DS mechanisms of neural integration

Earlier I proposed that, at the psychological level of analysis, appraisal and emotion constituents interact in the consolidation of an emotional interpretation (EI). However, at the neural level, emotion and appraisal constituents are more difficult to differentiate and classify. In this section, I will argue that, even if some neural structures are assigned to appraisal versus emotion categories, the interaction of these structures rapidly gives rise to processes that span both categories. Neural structures perform no function at all until they interact with each other, and even interactions among closely related structures, far below the whole-brain level, bind emotion and appraisal processes tightly together. Thus, brain function prohibits any real independence between appraisal and emotion, even at the level of subordinate processes. Why might this be so? In psychological terms, interacting emotion components tend toward coherence in service of an integrated *response* to the world, whereas appraisal components tend toward coherence in service of an integrated *interpretation* of the world. However, each of these functions is groundless without the other. An integrated response must be informed by the features and properties of a situation, provided by appraisal. An integrated interpretation must be constrained by the relevance provided by emotion. The biology of the nervous system seems to have evolved to unite these functions at many system levels, from part-system interactions to whole-system configurations. It also seems to have united these functions at every level of the neuroaxis, from the phylogenetically primitive brain stem (BS) to the recently evolved prefrontal cortex (PFC).

Thus, interactions between appraisal and emotion processes, within and across each level of the neuroaxis, bind together the "what" of appraisal with the "what to do about it" of emotion, in a functional unity. This binding is proposed to rely on several specific mechanisms of integration, without which global brain-behavior states could not cohere. I now propose five mechanisms of integration among neural systems subserving appraisal and emotion processes – mechanisms that yield the self-organization of global neural configurations mediating whole EIs. At the end of

https://doi.org/10.1017/S0140525X0500004X Published online by Cambridge University Press

each subsection, broad predictions are proposed as a guide to future research.

5.1. Nested feedback loops and self-synchronization

The first mechanism of integration is feedback among neural structures. Various sets of loops can serve as vehicles for positive and negative feedback interactions among the brain structures underlying appraisal and emotion processes. Neural systems of every size and function are characterized by feedback relations (e.g., Hebb 1949). Reentrant processes at the level of cells and cell assemblies exemplify feedback on a small scale, but I will focus only on large-scale loops, exemplified by reciprocal projections among limbic, cortical, and striatal structures. I first describe these loops as the biological apparatus by which appraisal and emotion processes become integrated. I then connect this framework to research on the functional integration and temporal synchronization of brain systems related to emotion.

Three points need clarification before describing large-scale loops. First, a great many circuits or loops have been proposed by different theorists to subserve emotional processes. Rolls (1999) and Depue and Collins (1999) emphasize loops involving the OFC, striatum, and AM. Thayer and Lane (2000) add to these the ACC, insula, hypothalamus, and BS. Gray (1987) highlights a dopaminergic system underlying behavioral activation and serotonergic and noradrenergic systems underlying behavioral inhibition. The original Papez circuit included the thalamus, hypothalamus, cingulate cortex, and hippocampus, and MacLean (1949) inserted the amygdala in an expanded version that he dubbed the limbic system. All of these proposed circuits are characterized by feedback relations, so any might suffice for DS modeling. However, I focus my modeling on loops among the neural systems I have already identified with appraisal and emotion constituents. This modeling pulls in contributions from many investigators, but it remains independent of any particular theory. Second, as shown in Figure 3, I nest these loops within three higher-order loops subserving evaluation, monitoring, and action functions. This nesting helps one to visualize levels of integration midway between lower-order components and whole-brain systems. Moreover, it demonstrates that intermediate levels of integration mediate both emotion and appraisal processes, not one or the other. Third, I highlight positive and negative feedback relations among neural structures. This is a highly simplified analysis. When a large number of components are connected through a multiplicity of channels, and these channels vary in type of transmission (e.g., type of neurochemical mediators), then positive and negative feedback interact in very complex ways. For example, the flow of excitation and inhibition between any two structures is mediated by the activity of all the other structures to which they are connected. However, as is often the case, a simplified analysis helps one to see the big picture.

This depiction of neural circuitry has another feature. It defines positive and negative feedback relations as causal processes, bridging neural and psychological explanations of self-amplification and self-stabilization. In general, the reciprocal activation of structures with other structures through parallel excitatory connections contributes to the positive feedback dynamics of self-amplification. When activation in one direction is countered by inhibition in the reciprocal direction, negative feedback yields self-stabilization. Because of the ubiquity of reciprocal connections and the complexity of nested loops, large-scale loops will generally include smaller loops showing both positive and negative feedback. This means that higher-order loops will show some characteristics of change and some of stability at the same time. However, the activation of higher-order loops, in which local loops are nested, contributes to self-stabilization and coherence through multiple constraints.

All three categories of loops presented in Figure 3 include the OFC and AM, both key structures in attentional,

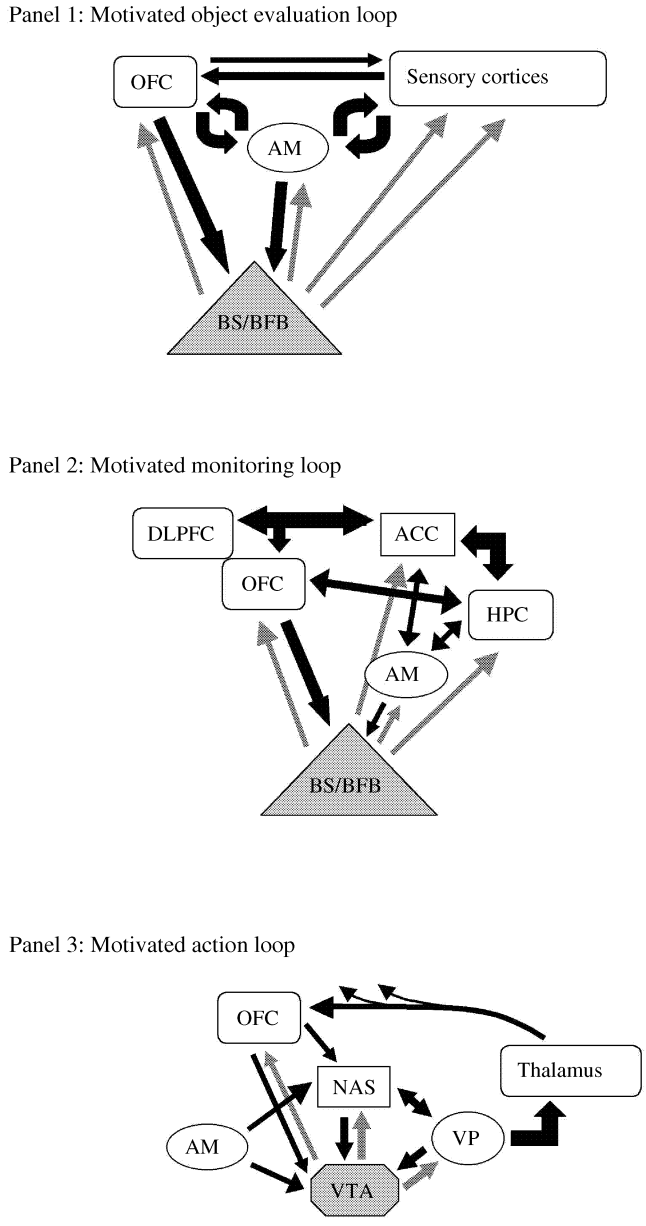


Figure 3. Three higher-order loops comprised of nested feedback circuits, each integrating particular emotion and appraisal processes. Black arrows represent glutamatergic and GABAergic (intrinsic) synaptic pathways. Grey arrows represent neuromodulator (extrinsic) pathways. ACC, anterior cingulate cortex; AM, amygdala; BS/BFB, brain stem/basal forebrain; DLPFC, dorsolateral prefrontal cortex; HPC, hippocampus; NAS, nucleus accumbens; OFC, orbitofrontal cortex; VP, ventral pallidum; VTA, ventral tegmental area.

https://doi.org/10.1017/S0140525X0500004X Published online by Cambridge University Press

motivational, and evaluative processes. They also include a representation of either the brain stem/basal forebrain (BS/BFB), denoting a cluster of what are in fact many discrete neuromodulator systems, or else the ventral tegmental area (VTA), a specific neuromodulator system that projects dopamine to the OFC, NAS, and other areas. I include these neuromodulator systems within the depiction of each loop for parsimony, even though their modes of communication (both chemistry and wiring) are different from those of cortical, limbic, and striatal structures, as discussed later.

The loops shown in the first panel of Figure 3 are centered on the amygdala (AM): (1) reciprocal connections between AM and BS/BFB structures that output neuromodulators back up to the AM and to all other limbic and cortical regions (e.g., Wallace et al. 1992); (2) reciprocal connections between the AM and OFC, and thus between two structures that coordinate evaluation and motivated attention (Rolls 1999); (3) reciprocal connections between the AM and integrative sensory areas such as the inferior temporal lobe, by which emotional associations focus perception (Barbas 2000; LeDoux 1995b); and (4) connections from higher-order sensory areas to the OFC (Barbas 2000; Ongur & Price 2000), permitting prefrontal assessment and regulation of sensory input. All three systems receive neuromodulator projections from BS/BFB nuclei, triggered by AM and OFC descending projections. Taken together, these loops constitute a higher-order "motivated object evaluation" loop, within which particular emotion and appraisal processes become integrated. The *motivational* aspect refers to OFC/AM initiation of ascending BS/BFB connections that enhance and orient evaluative/attentional activity; *object evaluation* refers to the specific stimulus-evaluation functions of the AM and OFC, both of which tune perceptual processing in the sensory cortices.

The next category of loops, shown in the second panel, highlights the ACC and HPC, with the latter including parahippocampal areas through which its connections flow: (1) direct bidirectional pathways between the HPC and OFC/DLPFC (Ongur & Price 2000), which may support the value-related modulation of retrieval and working memory (Petrides & Pandya 2002); (2) bidirectional pathways between the HPC and ACC, extending to the DLPFC, necessary for selective attention and monitoring (Petrides & Pandya 2002); (3) bidirectional pathways between the ACC and AM, through which ACC processes control and are tuned by AM associations (Poremba & Gabriel 1997); (4) bidirectional pathways between the HPC and AM, allowing emotional enhancement of episodic memory (Hamann et al. 1999); and (5) primarily one-way connections from the ACC to the OFC (Barbas 1995), possibly synchronizing voluntary attention with motivational set. Again, all four structures receive neuromodulator projections from BS/BFB nuclei, triggered by the OFC and AM. These loops can be described as nested in a higher-order "motivated monitoring" loop, integrating additional emotion and appraisal processes. The *motivational* aspect refers chiefly to BS/BFB and ACC mediation of emotional feeling and attentional orientation, whereas *monitoring* denotes the modulation of prefrontal and HPC activities by the ACC, permitting voluntary attention to challenging circumstances.

A third category of loops involves the ventral striatum: (1) OFC projections to the NAS, reciprocal connections between the OFC and VTA, and reciprocal connections between the VTA and NAS (Haber et al. 1995; Oades & Halliday 1987), collectively underpinning the motivational control of action selection and initiation; (2) reciprocal connections between the NAS and ventral pallidum (VP, an output structure of the striatum), with both structures reciprocally connected to the VTA, constituting a resonating motivational circuit (Depue & Collins 1999); (3) connections from the ventral pallidum to a nucleus of the thalamus, and thence to the OFC (as well as premotor systems, denoted by upward arrows), completing a larger feedback circuit that continuously moderates motivated behavior; and (4) pathways from the AM to the NAS and VTA, enhancing activation of both systems based on emotional associations (Cardinal et al. 2002). These interconnected loops may be said to constitute a higher-order "motivated action" loop, integrating additional emotion and appraisal processes. Specifically, *motivation* refers to arousal and action tendencies mediated by AM/OFC-initiated dopaminergic innervation of orbitostriatal circuits; and *action* is selected and focused on by the NAS, then sequenced by striatal outputs to the thalamus and motor system (Rolls 1999).

As these neural structures interact in positive feedback relations, each of the three higher-order loops becomes activated. As these interactions settle into negative feedback relations, their activity becomes stabilized. However, because these loops are deeply interconnected with each other, through anatomical structures common to all three (e.g., OFC, AM, BS/BFB), their joint activity can be described in terms of a global macrosystem. Sensory data and motivationally relevant associations are integrated with action tendencies through connections between the object evaluation and action loops. This integration affords an ongoing state of engagement with the world, whereby everything we see and hear with emotional significance engages an urge to act. At the same time, connections between the object evaluation and monitoring loops integrate sensory data and motivationally relevant associations with conscious monitoring and explicit memory. This integration affords an ongoing appreciation of the meaning of events in the world – an emotional context that shapes our attention. Finally, connections between the monitoring and action loops permits us to fashion a continuously updated plan, guiding the leading edge of conscious behavior according to what is most compelling. Through this integration (visited in more detail later), we sense ourselves moving through space and time according to a deliberate agenda. Thus, connections within each loop integrate particular emotion and appraisal processes, but these processes become synchronized with each other in a global macrosystem subserving a unified emotion–appraisal gestalt. The specific configuration of activated circuitry (at the macrosystem level) dictates the unique character of this gestalt as it emerges on each occasion.

What evidence is there to suggest that the co-activation of neural structures within loops mediates emotion and appraisal? A number of neuroimaging studies have looked at the role of emotion in functional integration, revealed by changing activation patterns across some subset of regions. First, several studies have specifically examined functional connectivity between the amygdala and cortical regions in relation to emotion (see Dolan 2002; Dolan & Morris 2000, for reviews). For example, functional connectivity between the AM, visual cortex, and BS is enhanced when subjects

https://doi.org/10.1017/S0140525X0500004X Published online by Cambridge University Press

are exposed to fearful faces, but reduced in the presence of happy faces (Morris et al. 1998). Dolan and Morris (2000) suggest that AM potentiation of sensory processing can be viewed as "effective connectivity" resulting in increased coherence. In the present approach, these findings would suggest self-amplifying feedback and stabilization within the object evaluation loop. Second, various brain regions become activated simultaneously, whereas others may become deactivated, in the presence of emotion. For example, negative emotions such as sadness, anxiety, and guilt are associated with increased activation in the AM, ACC, and OFC and with decreased activation in the dorsolateral PFC and HPC (Liotti et al. 2000; Shin et al. 2000). As well, Dolan and Morris (2000) review evidence that aversive (fear) conditioning produces positive covariation between activation of the hypothalamus and activation of the AM, BFB, OFC, ACC, and HPC. Both sets of findings imply synchronized changes in the object evaluation and monitoring loops, but the specific pattern of changes may vary with the type of emotion. A third kind of functional integration is discussed by Cardinal et al. (2002) as the "upregulation of associability," involving AM activation of BS systems that transmit acetylcholine to cortical sensory association areas following aversive conditioning. Here the motivational effects of fear appear to enhance the receptivity of cortical regions to associations among stimuli. I return to this issue later, but for now it serves as an example of emerging perceptual couplings induced by emotional processes. It also highlights the important point that each "structure" or "system" contains a vast number of synaptic networks, any combination of which can become activated through connections with other structures. Discrete appraisals are discrete because they activate these networks selectively.

These studies provide evidence of *functional* integration among systems subserving appraisal and emotion processes. However, a DS approach might also imply *temporal* synchronization underlying these functional links. As noted earlier, neuroscientists with a dynamical perspective have studied temporal synchronization, mostly across cortical and thalamic sites (e.g., Bressler & Kelso 2001; Engel et al. 2001; Nunez 2000; Skinner & Molnar 1999; Thompson & Varela 2001). These approaches rely on various methods for measuring temporal correlations, such as phase locking or phase synchrony (a fixed temporal relation between the oscillations of independent regions) and coherence analysis (integrating amplitude and phase in one measure), as applied to data from scalp EEG, local field potentials, and single-cell recordings. These techniques are often used to examine synchronization in the gamma band (about 30–80 Hz), sometimes across distal cortical regions, corresponding to attentional states of expectancy or focused perception. A number of studies demonstrate that the degree of synchrony corresponds with the degree of attentional engagement or motor readiness (Lutz et al. 2002; see review by Engel et al. 2001). Thompson and Varela (2001) argue that such synchronization reflects the integration of many dimensions in a unified "cognitive act," rather than the clustering of component perceptual processes. Using dynamical equations, Thelen et al. (2001) demonstrate the emergence of a coherent cognitive act from the coupling of neural components underlying perception, planning, and memory. Such cognitive acts may be analogous to whole appraisals in the present treatment. In any case, this body of research has linked coherent psychological processes to intrinsically generated synchrony, thus demonstrating a key inroad to the study of neural self-organization.

However, the thalamocortical systems studied with this paradigm are concerned with cognition, not emotion, and gamma-band oscillations may not last long enough to tap extended motivational states (see Kopell et al. 2000). The few studies that have investigated temporal synchronization in relation to emotion concentrate instead on the theta band (about 4–8 Hz). Theta waves are thought to signify hippocampal involvement in brain activity, notably in cortico-hippocampal interactions. A "pacemaker" in the septum can generate hippocampal theta, though hippocampal cells fall into theta-frequency oscillations on their own, and other regions including the PFC, ACC, thalamus, hypothalamus, and AM also show theta frequencies (e.g., Vertes & Kocsis 1997). Importantly, theta-band activity corresponds with attentional monitoring, action monitoring, learning, retention, and shifting behavior in situations that demand action (Klimesch 1999; Luu & Tucker 2002). Thus, oscillations in the theta range could be critical for synchronizing appraisals corresponding with emotional demands. Theta-band activity has been found to underpin large-scale synchronization across distant cortical regions (Buzsaki 1996; von Stein & Sarnthein 2000), and possibly throughout the Papez circuit (including the thalamus and hypothalamus) (Vertes et al. 2001). Theta waves have also been described as "carrier waves," in which gamma oscillations representing perceptual details are embedded (van Rullen & Koch 2003), and theta has been proposed as the fundamental rhythm of corticolimbic self-organization (Miller 1991). Thus, self-organizing emotion–appraisal states could potentially be realized by the global synchronizing properties attributed to theta-band oscillations.

Studies of emotion-related theta-band activity are few but revealing. Phase-synchronized theta oscillations were recorded in the AM of animals anticipating a shock, corresponding with blood pressure increases suggesting fear (Paré & Collins 2000). Arousal has also been reported to increase the coherence of theta oscillations between the AM and PFC (see review by Paré et al. 2002). Because theta-band oscillations are recorded in the AM only during emotional states, Paré et al. propose that AM and HPC activities become coupled at theta frequencies in emotion-inducing circumstances, and this coupling harnesses cortical activation related to emotional memories. These ideas find support in recent evidence that HPC-AM synchronization at theta frequencies follows fear conditioning and peaks when animals confront the fear-inducing stimulus (Seidenbecher et al. 2003).

5.1.1. Novel predictions. Theta-band synchronization may mediate emergent couplings among additional neural subsystems when appraisal and emotion processes cohere in a unified gestalt. Specifically, it might be predicted that phase synchrony at theta across the AM, HPC, ACC, OFC, and sensory association areas would emerge when situations or tasks become emotionally relevant, and that this phase synchrony would indeed index their relevance. Thus, theta-band synchrony might correspond with emotional relevance, just as gamma-band synchrony corresponds with perceptual closure. This hypothesis extends Miller's (1991) claim that theta-band synchrony underpins cortico-hippocampal self-organization necessary for processing situa-

https://doi.org/10.1017/S0140525X0500004X Published online by Cambridge University Press

tional context. Emotions may play a critical role in context processing, and they may be needed for cortico-hippocampal loops to cohere in the first place. Specifically, it is predicted that the degree of phase synchrony in the theta range will correlate with measures of emotional relevance and subjective emotional feelings, and that this synchrony will increase as emotional relevance increases over time, just as the degree of gamma- and beta-band synchronization covaries with changing levels of attention and motor readiness (Engel et al. 2001). Suggestions for testing this prediction are advanced in a later section.

5.2. Neuromodulation: Global effects on change and stabilization

A second basic mechanism of neural integration is the action of ascending neuromodulatory pathways from the brain stem (BS) and basal forebrain (BFB) to all regions of the diencephalon, striatum, limbic system, and cortex. As reviewed earlier, these pathways originate in a variety of cell bodies that release dopamine (DA), norepinephrine (NE), acetylcholine (ACh), and serotonin (5-HT). Also, as noted earlier, the function of ascending neuromodulators has often been characterized in terms of arousal or activation, reflecting their role in the motivational enhancement of attention and action readiness. However, these systems can also be differentiated into highly specific effects across diverse brain regions (see review by Gu 2002). Thus, input from a single neuromodulator system can enhance cortical activities in some areas and inhibit them in others. Despite this specificity, ascending neuromodulators must have global effects on the interaction of appraisal- and emotion-mediating systems, something like the global effects of climate change on an ecosystem (see Fig. 4). This is because neuromodulators act diffusely, over a huge variety of neural regions, influencing activity throughout the brain. Consequently, neuromodulator activation may be seen as a mechanism of global change, either augmenting present firing patterns or altering them in favor of competing activities, but also as a mechanism of global stabilization, consolidating particular activities that regulate them in turn.

The feedback among cortical, limbic, and striatal structures, depicted in the last section, relies on transmission from one cell body to the next via neurotransmitters such as glutamate and gamma-aminobutyric acid (GABA). This kind of activation exerts either excitatory or inhibitory effects on the electrical potentials (and ultimately the firing rates) of target cells: glutamate is always excitatory, and GABA is always inhibitory. In contrast, the action of neuromodulators released from BS/BFB structures is partially a function of their type and origin, and partly a function of the properties and locations of their receptors. This complex interaction effect determines whether constellations will be augmented, maintained, or altered. For example, NE can globally enhance alerting, arousal, and vigilance functions, both in higher brainstem regions and in sensory cortex, thus augmenting present perceptual activities. However, it can also switch the control of perception and behavior from frontal to posterior cortical systems, as a result of the interplay of two receptor types (Arnsten & Robbins 2002). Similarly, ACh has often been linked to motivational enhancement of attention through the activation of cortical neurons, thus triggering new appraisals or augmenting existing ones. However, cortical firing can be inhibited by ACh as well (McCormick et al. 1993). DA activates approach and exploratory behavior globally, through its effects on orbitostriatal neurons, yet its prominent role in cortex is to suppress neuronal firing. DA can also have very specific effects on attention and working memory by selectively enhancing and inhibiting the firing rates of different prefrontal neurons (Williams & Goldman-Rakic 1995). The interplay of excitation and inhibition produced by NE and DA may set the OFC into modes of activation that mediate specific appraisals (Schore 1994). Despite their variety of effects, both the diffuse nature of ascending neuromodulators and the persistence of some neuromodulator effects over time work to insure their global influence. Moreover, the structures affected by neuromodulators include neuropeptide systems that activate and maintain particular motivational and behavioral tendencies in brain and body, as previously discussed.

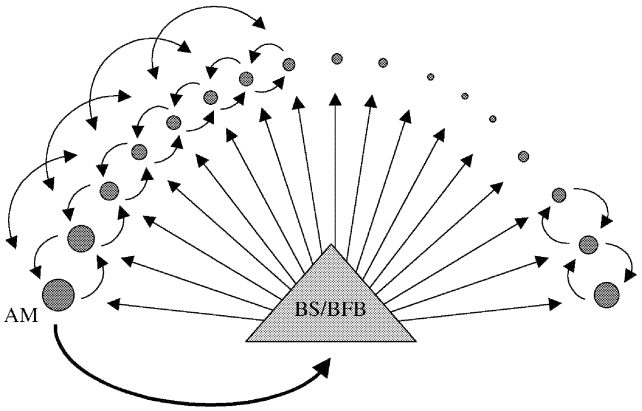


Figure 4. Ascending neuromodulator release from brainstem/basal forebrain (BS/BFB) nuclei visualized as climate change in an ecosystem: some interactions are augmented while others are inhibited. Structures such as the amygdala (AM), which initiate this process, trigger a one-to-many pattern of influence extending to diverse brain regions.

Neuromodulation has often been associated with emotion or motivation, because of its arousal and activation properties, but also because of its participation in circuits that trigger emotional responses. Neuromodulator release can be initiated and regulated by glutamate projections from the AM, OFC, ACC, or NAS (Cardinal et al. 2002; Ongur & Price 2000); it can then enhance activation of these and other structures in turn. Thus, by activating BS/BFB nuclei, the AM and OFC induce neuromodulator release that not only returns to modify their activity but also fans out to many other cortical, limbic, and brainstem regions (Holland & Gallagher 1999; Rolls 1999). For example, the AM activates BS nuclei that release DA, ACh, and NE, which in turn adjust the activation of prefrontal, orbitofrontal, insular, ACC, and other systems involved in various cognitive and emotional functions (Fuster 1996; Kandel et al. 2000; Oades & Halliday 1987). Thus, structures mediating specific appraisal and emotion processes can trigger neuromodulatory release that binds these and other processes together in global configurations. As shown in Figure 4, any system that ignites neuromodulator release triggers a one-to-many, expanding, influence pattern, coupling information-processing activities across diverse sites. Returning to the climate analogy, the activities of one species may induce climate change that reorganizes the entire

https://doi.org/10.1017/S0140525X0500004X Published online by Cambridge University Press

ecosystem, benefiting some organisms while annihilating others.

Panksepp (1998a) reviews evidence that neuromodulators are instrumental in the functional integration of global emotional states as well as the action modes that accompany them. But do neuromodulators play a role in temporal synchronization as well? For many years it has been known that activation of BS reticular systems induces hippocampal theta (e.g., Apostol & Creutzfeldt 1974), and this induced theta is associated with alertness and sensory processing (Bland et al. 1984). More recently, it has been found that various structures within the limbic system, paralimbic cortex (e.g., ACC and parahippocampal areas), and neocortex contain cells that oscillate in the theta band, and that these oscillations are also triggered by ascending neuromodulators (particularly ACh) from BS systems (Bland & Colom 1993). Ascending neuromodulators regulate a variety of electrophysiological rhythms in the cortex as well, corresponding with organized perceptual and motor activities (see review by Vanderwolf 1988). Bland and Oddie (1998) review evidence of a multisynaptic pathway, extending from pontine nuclei in the brainstem, through hypothalamic sites, to the septum, from which inputs are distributed to the HPC, parahippocampal cortex, and cingulate cortex. Both ACh and serotonin (5-HT) stimulate theta activities through this pathway. These authors also report that the amplitude of theta oscillations in these structures corresponds to the intensity of neuromodulator action.

5.2.1. Novel predictions. This literature suggests that ascending neuromodulators (especially ACh) are critical for generating theta-band oscillations in a variety of corticolimbic regions. In the last section, I argued that emerging phase synchrony across corticolimbic structures could underpin global emotion–appraisal states. Integrating these perspectives, it might be predicted that BS-induced corticolimbic synchronization is the vehicle for appraisal–emotion coupling. To test this hypothesis, EEG methods for assessing phase synchrony across diverse cortical areas could be linked with neuroimaging measures of activation in the BS and hypothalamus when emotion-inducing stimuli are being processed. Furthermore, because physiological measures of sympathetic activity (e.g., skin conductance, heart rate) are traditionally used to assess emotional response, and because these measures correspond directly with BS and hypothalamic activation, these measures might also correlate with synchrony coefficients across cortical and paralimbic sites. Measures such as skin conductance are quite sensitive to arousal changes in real time. Therefore, both skin conductance and theta-band synchrony might be expected to increase in tandem during time windows leading up to anticipated anxiety-eliciting events, such as aversive cues, and this effect should be facilitated by state or trait anxiety. To my knowledge, psychophysiological indicators of emotion have never been linked with the study of temporal synchronization across cortical regions, but this may be a fruitful direction for cognition–emotion research.

5.3. Vertical integration

A third mechanism of integration is the superordinate integration of activities across multiple levels of the neuroaxis, termed *vertical integration* by Tucker et al. (2000) and others. Up to this point, I have discussed feedback relations that link structures within nested loops, as well as neuromodulatory activities that exert a global influence on the integration and coupling of these structures. The causal relations described until now include reciprocal and recursive interactions among parts (component systems), or else global influences that flow unidirectionally in an expanding, one-to-many fashion. Although these interactions might indeed yield self-organizing states of coherence, there is as yet no mechanism to relate that coherence back to component interactions. The principle of circular causality fills this gap. As reviewed earlier, Haken's (1977) circular causality describes the relation between a superordinate emergent organization and the subordinate components whose coordination gives rise to it. Specifically, the superordinate maintains the coupling of all the constituents in a top-down fashion, while their coupling maintains the integrity of the whole through bottom-up processes. This perspective is central to neuroscientists with a dynamic orientation. According to one interpretation, neural coherence is manifested globally by the synchronization of a functional whole, but manifested locally by the entrainment of each subsystem to this global pattern (Nunez 2000; Rolls & Treves 1998; Varela et al. 2001). The whole thus grounds the parts, entraining (or "enslaving") them to a particular organizational regime. Importantly, this principle is independent of the negative-feedback relations that couple lower-level components to each other, sometimes in extended networks, through parallel rather than vertical causal interactions.

Vertical integration could instantiate circular causality at the whole-brain level, thus providing an additional mechanism of neural integration. As argued by Tucker et al. (2000), vertical integration joins the levels of the neural hierarchy in a bidirectional stream of influence. These authors characterize upward influences as arousing or recruiting, alerting the cortex to the emotional significance of events, and downward influences as providing detailed information and action plans that regulate lower structures. However, they argue, bidirectional causation, incorporating both recruitment and regulation, coordinates the entire neuroaxis, such that functionally (and phylogenetically) distinctive levels cohere in a unitary mode of action readiness and attentional orientation (see also Freeman 1995; 2000). Vertical integration links the stereotypic functions of the brain stem with the executive, planful, and information-rich character of cortical activity, through the mediation of limbic structures. It is considered necessary for coordinating perception, attention, and planning with primitive action tendencies, so the animal can behave flexibly, skillfully, and intelligently when motivated (Tucker et al. 2000). According to Freeman (2000), the resulting cognitive-emotional orientation to the environment, or the *intentionality* that characterizes this orientation, defines the superordinate in a circular causality.

If vertical integration locks the entire neuroaxis into a coherent mode, it may indeed be the superordinate in a circular causality that entrains the interactions of component subsystems. Yet, the mechanism of this meta-integration is unknown. Phase synchrony has been proposed as the superordinate that entrains cortical neurons through top-down modulation during states of attention or consciousness (e.g., Engel et al. 2001; Thompson & Varela 2001). Is there evidence for an emerging, superordinate phase synchrony across the entire neuroaxis? Kocsis and

https://doi.org/10.1017/S0140525X0500004X Published online by Cambridge University Press

Vertes (1992) reported that not only corticolimbic systems but also BS systems (e.g., raphe nucleus) include cells that oscillate in the theta band. Moreover, these oscillations were found to be synchronized with hippocampal theta, leading these authors to conclude that phasic rather than tonic neuromodulator activity induces theta-band oscillations in corticolimbic regions. However, these results still implied a one-way influence from BS to limbic regions. Two years later, Kocsis and Vertes (1994) reported that two nuclei in the hypothalamus, the supramammillary nucleus and the mammillary body (MB), both showed "theta bursting" synchronized with HPC cells. However, because connections from the MB are descending, not ascending to the HPC, they reasoned that theta-band rhythms in the MB might originate from the HPC – the reverse of the direction previously assumed. This work demonstrated the likelihood of two-way synchronizing pathways between the HPC and lower-brain structures, suggesting temporal synchronization as a possible vehicle for vertical integration. However, these reciprocal influences still did not imply an emergent circular causality. Recently, Kocsis et al. (2001) reported findings that crossed this line. They studied a brainstem structure called the tegmental nucleus of Gudden, known to be connected to the MB in a negative feedback loop. They highlighted the importance of spontaneous theta-band activity found in MB neurons (Alonso & Llinus 1992) and hypothesized that it was generated and maintained within this loop, providing a source of theta oscillations independent of the HPC. They then explored the correlation between these two sources of theta-band activity in relation to stimulus events. They found that theta waves recorded in the tegmental nucleus of Gudden (and thus highly correlated with MB activity) became either moderately or highly synchronized with hippocampal theta, depending on sensory stimulation. With no direct stimulation, synchrony was moderate. However, tail-pinch stimulation ramped up the coherence coefficient to an extremely high value. These results suggest that meta-synchronization between two independent self-synchronizing systems (one spanning the hypothalamus and brain stem, and the other in the limbic system) emerges spontaneously as a function of experience. Thus, it may be that the emergence of a superordinate phase synchrony, linking BS and corticolimbic subsystems, is a good candidate for circular causality, constraining the coordination of the subsystems from which it arises.

5.3.1. Novel predictions. By this analysis, vertical integration along the neuroaxis, mediated by meta-synchronization of independent oscillatory subsystems, could be a key mechanism for generating and maintaining a cognitive-emotional orientation to the world. If this is correct, then the spontaneous synchronization of these subsystems should depend not only on sensory stimulation, but also on the emotional relevance of that stimulation. One might assume that having one's tail pinched is motivationally relevant for a rat, but Kocsis et al. (2001) report on only two levels of relevance: absent and high. Further research could investigate whether gradations in the emotional relevance of stimuli are correlated with the degree, consistency, and especially the extent of synchronization along the neuroaxis. Perhaps different emotions tap different subsets of oscillatory systems, and the discovery of additional oscillators in lower-brain regions may be facilitated through emotion-induction paradigms. Thus, extending the hypotheses suggested earlier, theta-band phase synchrony may covary with other measures of emotion continuously, not just dichotomously, and across the entire neuroaxis, not just corticolimbic systems. This finding would support the proposition that emotion is a necessary feature of global self-organizing processes at the whole-brain level.

5.4. Action monitoring: The focus of integration

All the mechanisms of integration discussed so far have a somewhat solipsistic quality: they act within the brain and body, with little concern for the outside world. However, the convergent wholes to which self-organization gives rise should also be identifiable at a functional level – a level at which superordinates of one system (the organism) can interact with those of other systems (the environment). Emotion theorists generally agree that the function of emotion is to cause or propel action by selecting a particular action tendency or creating a state of action readiness (Frijda 1986; Izard 1991). Thus, a fourth mechanism of neural integration might be found in action monitoring – the preparation, execution, and regulation of action. Because the brain has evolved in mammals to delay impulsive actions, appraisal components of attention, planning, and memory normally interact with the primitive action tendencies associated with emotions. This permits the construction and elaboration of an intelligent response, or plan, over an extended time course (Tucker et al. 2000). A continually updated plan, by which behavior is monitored and controlled, may guide the integration of neural systems mediating emotion and appraisal. This idea links vertical integration within the brain to emerging transactions between the brain and the world.

Vertical integration implies emergent synchronization across all levels of the neuroaxis in service of a coherent response to events in the world (e.g., having one's tail pinched). Thus, an action orientation may be critical for the synchronization and coherence of the brain (Freeman 2000). Indeed, the global macrosystem defined by nested feedback loops and the neuromodulatory processes that support its activity are clearly involved with action. Two loops are particularly important. The OFC–striatal–thalamocortical circuits of the motivated action loop initiate and coordinate the steps of a motor sequence, while the hippocampal–ACC–PFC circuits of the motivated monitoring loop monitor and regulate the flow of action consciously and intentionally. These two functions work together in any intelligent behavioral process.

The motivated action loop is specifically dependent on dopamine (DA) released by the VTA. DA facilitates behavior through incentive motivation, but it also constricts behavior by focusing and maintaining attention to specific targets, via a "closed" prefrontal–striatal–thalamocortical loop (Groenewegen 1997). This "dopamine-induced focusing" narrows the activation of NAS neurons to contextually relevant targets highlighted by projections to the NAS from the PFC, AM, and HPC (Depue & Collins 1999). Focusing is also facilitated by the dendritic architecture of the striatum and pallidum, which funnels widespread inputs from the cortex into unitary motor output patterns (Rolls 1999). Thus, integrative feedback processes harness the striatum, where DA and neuronal architecture work together to narrow the focus for action. This implies that vertical integra-

https://doi.org/10.1017/S0140525X0500004X Published online by Cambridge University Press

tion not only synchronizes brain activity; it also anchors it to a narrow behavioral stream.

The ACC plays a key role in the intentional regulation of action, particularly as influenced by motivational factors (e.g., Holroyd & Coles 2002). Action monitoring by the ACC has been studied using event-related potential (ERP) methods that tap the electrical activity of medial-frontal cortical regions. In particular, the error-related negativity (ERN) and error positivity (Pe) constitute two characteristic ERP deflections, approximately 80 msec and then 200–500 msec following an error response, when the subject is aware of and concerned about an error. These components are thought to index not only attention to errors but action monitoring more generally (Holroyd & Coles 2002; Luu & Tucker 2002; van Veen & Carter 2002). Both components have been localized to the ACC through ERP source analysis techniques as well as fMRI (e.g., Carter et al. 1998; Gehring et al. 2000; van Veen & Carter 2002), though sources in the parietal cortex and rostral ACC appear to contribute to the Pe as well. Frontal ERP deflections following a behavior-relevant cue (e.g., the no-go N2) look similar to those following errors (Luu et al. 2003; van Veen & Carter 2002), suggesting that activation of the ACC mediates the monitoring of behavior even prior to taking action. Finally, in an influential paper, Holroyd and Coles (2002) link DA involvement in ACC activity with its role in reward anticipation (e.g., Schultz 1998). They then propose that the ERN results from a drop in DA innervation from VTA neurons, disinhibiting ACC neurons when errors occur and actions need revision. This mechanism suggests functional coordination between striatal and ACC activities, mediated by DA modulation, when action plans are monitored and updated.

In the preceding sections I developed the argument that phase synchrony in the theta range may underpin the functional integration of systems mediating appraisal–emotion processes. Is there evidence for theta-band synchronization in action monitoring? A number of investigators have proposed that ERPs are produced by the summation of brain waves of the same or related frequencies (e.g., Karakas et al. 2000). This interpretation is particularly congenial with a DS framework in which discrete events are viewed as temporary states of coherence or coupling. Makeig et al. (2002) showed that perceptual ERPs result from the compilation of alpha-band oscillations, synchronized by the triggering effect of a stimulus event. They viewed this synchronization phenomenon as "phase-resetting" across multiple channels. Luu and colleagues have argued that action-monitoring ERPs including the ERN and Pe reflect phase-resetting at the theta band; this hypothesis is particularly interesting in light of evidence for the large-scale synchronizing properties of theta oscillations (Buzsaki 1996; von Stein & Sarnthein 2000). As preliminary evidence, Luu et al. (2003) used an error feedback paradigm and found a rostral ACC source and dorsal-midline source oscillating in phase in the theta range, with their summed deflections constituting the ERN. Luu and Tucker (2002) relate this finding to Asada et al.'s (1999) report of phase-related theta oscillations between generators localized to the ACC and medial PFC during attention-demanding tasks. Luu and Tucker go on to propose a phase-resetting model of action regulation, and they suggest that multiple cortical sources become coupled at theta when ERNs are generated (see Menon et al. 2001, as to where these sources may be located). Thus, medial-frontal ERP components may indeed represent temporary spans of theta-band synchrony subserving action monitoring.

The next question is whether medial-frontal ERPs tap appraisal and emotion processes. Clearly, these ERPs are associated with key ingredients of appraisal: evaluation, attention, and planning. However, a good deal of research indicates that they also tap emotional states or traits. With respect to emotional states, Tucker et al. (1999) found higher-amplitude medial–frontal ERPs when participants received critical feedback on their performance, an effect replicated by Luu et al. (2003) independent of actual recent performance. Tucker et al. (2003) found a similar ERP differentiating between good and bad trait-descriptive words (e.g., "generous" vs. "mean"), potentially tapping appraisals linked to negative emotions such as anxiety. Indeed, ERNs diminish in magnitude when participants are given anxiety-reducing drugs (Johannes et al. 2001). With respect to traits, undersocialized individuals who probably experience little social anxiety show lower-amplitude ERNs (Dikman & Allen 2000). Conversely, obsessive-compulsive individuals show higher-amplitude ERNs than normal individuals (Gehring et al. 2000). Higher-amplitude ERNs have also been associated with lower scores on impulsivity (Pailing et al. 2002). Finally, Luu et al. (2000) found that subjects with higher scores on trait negative affect produced higher-amplitude ERNs. These studies indicate that medial-frontal ERPs correlate with state or trait negative emotionality: Increased negative emotion corresponds to higher mean amplitudes, implying greater theta-band synchrony. In conclusion, the cognitive processes tapped by these ERPs may be recruited to monitor and regulate behavior specifically related to emotion, with greater emotional intensity predicting greater cognitive involvement.

According to this analysis, the evaluative, attentional, and planning aspects of appraisal appear to correspond with action tendencies and emotional feelings, as mediated by functional integration between the ACC and striatum. This integration may be induced by ascending neuromodulator activities, and it may couple the dorsal ACC with other cortical areas, as reflected by the sources of frontal ERPs. Moreover, this integration may be underpinned by the phase-locking of theta-band oscillations that are implicated in large-scale corticolimbic self-organization. In sum, the attentional focus involved in action monitoring may be central for linking the "what" of appraisal with the "what to do about it" of emotion, permitting the modulation of an organized behavioral orientation to a challenging world.

5.4.1. Novel predictions. The relations among action-monitoring ERPs, theta-range synchrony, and neuromodulation suggest a unique approach to the study of appraisal–emotion processes. One implication of this approach is that action-monitoring ERPs actually tap appraisals. While these "mini-appraisals" may not reflect the complexity of everyday cognitive-emotional activities, they nevertheless imply spontaneous synchronization among several discrete systems mediating emotion and appraisal processes. A number of predictions flow from this idea.

First, action-monitoring ERPs suggest a time course for appraisal consolidation, roughly between the appearance of the ERN and the completion of the Pe. Specifically, the Pe, which includes more regions than the ERN, marks the final ERP deflection following an error. The latest stage of

https://doi.org/10.1017/S0140525X0500004X Published online by Cambridge University Press

the Pe is roughly 400–500 msec postresponse, or 600–800 msec following the awareness of the error, which is thought to precede the response itself by 200–300 msec. Thus, an EI may emerge fully within 800 msec of a challenging event. (Note that Miller [1991] reviews evidence that septohippocampal entrainment grows over six cycles of theta – approximately the same period.) This prediction could be explored through concurrent analysis of ERP wave-forms and profiles of change in psychophysiological measures of emotion. Second, differences in both appraisal contents and emotional feeling states should correspond with the type of ERP (e.g., ERN vs. Pe) and with ERP amplitude. Tucker et al. (2003) found amplitude differences that may reflect positive and negative appraisals induced by positive and negative trait-descriptive words. In this kind of study, where appraisal content is directly manipulated, self-report or physiological measures of emotion should constitute an additional predictor of ERP amplitudes. Third, the degree of cortical synchronization, hypothetically contributing to the shape and magnitude of ERP deflections, should correspond with appraisal coherence and emotional measures. Thus, participants' reports of their appraisal and feeling states, and perhaps their success at focusing on particular appraisals linked with particular emotions, should correspond with synchrony coefficients during the time-frame of the ERP wave-form. New methods for analyzing coherence among cortical sources of ERP activity could be devoted to such an analysis (see Scherg et al. 2002 for a lead-up to cortical source coherence analysis).

5.5. Plasticity and learning

The fifth and final mechanism of integration is one that works not within occasions but across occasions. Earlier in this article, I reviewed the DS principle that the flow of activity among the elements of a system changes the elements themselves, enhancing the probability that the same patterns of activity will recur on future occasions. Hebb (1949) applied a similar principle to explain brain plasticity and learning, whereby the co-activation of neurons produced structural changes at synapses between them, increasing their probability of becoming co-activated in the future. By increasing neuronal connectivity across occasions, mechanisms of learning may be seen as facilitating neural integration in real time. Thus, through the growth of synaptic connections among structures mediating appraisal and emotion, learning establishes and reinforces links among the contents of appraisal as well as links between those contents and emotional response patterns. However, learning also establishes neural integration at the scale of developmental time, by providing continuity between appraisal–emotion patterns from one occasion to the next.

An important class of candidate mechanisms for associative learning includes long-term potentiation (LTP). In LTP, particular frequencies or durations of firing of the pre-synaptic neuron produce long-term chemical changes in the post-synaptic neuron, permanently altering the structure of the synapse. For this to occur, the post-synaptic neuron must be activated, glutamate from pre-synaptic terminals must travel to a specific class of receptors in the post-synaptic neuron, and then protein synthesis must take place for a period of time (up to hours) and/or across occasions. As a result, it takes less activation to produce the same response in the post-synaptic neuron on future occasions.

The state of excitability of the receptive neuron and the time course of its activation are crucial determinants of LTP. For many authors this implies that the neurochemical excitation that accompanies emotional states is essential for synaptic modifiability and learning (Freeman 1995; Post et al. 1998; Tucker 2001). Research demonstrates that neuromodulator arousal facilitates LTP (e.g., Centonze et al. 2001; Izquierdo 1997; Izumi & Zorumski 1999) and that neuropeptide action consolidates synaptic change and enhances memory formation (Adamec et al. 1998; Flood et al. 1990). The amygdala (AM) may be critical to memory consolidation in various systems, because of its facilitation of BS/hypothalamic neurochemical release (Packard & Cahill 2001) and its direct projections to the HPC (Hamann et al. 1999). Action-monitoring ACC activation, previously related to motivation, is also found to enhance learning (Gemba et al. 1986). These and related findings suggest that events that are not emotionally significant may not maintain arousal or attention long enough for learning to take place (Gallagher & Holland 1992; Rolls & Treves 1998; Tucker 2001). Finally, LTP has been observed in limbic, paralimbic, striatal, and cortical structures, but lower brain systems do not show plasticity of this kind. As a result, the information that consolidates through LTP must derive from attentional and evaluative cortical systems and their limbic and striatal underpinnings.

Thus, appraisal elements may be linked over occasions through the formation of associations that depend on emotional activation within occasions. This is exemplified by Cardinal et al.'s (2002) notion of the "upregulation of associability." These authors review findings indicating that the AM influences the associability of stimuli through its projections to BS neuromodulatory systems (e.g., Gallagher & Holland 1994). The AM may trigger ascending ACh projections that terminate in the parietal cortex, permitting the learning of novel associations in the presence of motivationally compelling events. Cardinal et al. (2002) link heightened associability in the cortex with the modulation of "affective" motor response patterns subserved by the BS (e.g., orienting). They conclude by suggesting that the learning of new associations, mediated by AM–BS involvement, is accompanied by autonomic changes, motivational arousal, and attentional orientation – all components of emotion in the present treatment.

This analysis suggests that the functional integration of AM–BS–corticolimbic structures within occasions produces learned associations across occasions. However, in previous sections I suggested a role for temporal synchronization in emotion–appraisal consolidation, induced by neuromodulator activity, participating in vertical integration, and tapped by action-monitoring ERPs. If phase synchrony is necessary for integrated appraisal–emotion processing in real time, then it should mediate learning as well. It has already been suggested that theta-band activity plays a role in learning (e.g., Klimesch 1999) and that ACC activation, hypothetically tapping theta-band synchrony, enhances learning as well (Gemba et al. 1986). Theta-band activity has also been shown to induce or facilitate LTP directly (e.g., Natsume & Kometani 1997; Yaniv et al. 2003). However, theta oscillations may specifically mediate the contribution of emotion to learning. Paré and Collins (2000) report an increase in the responsiveness of theta-synchronized AM cells over learning trials when animals are afraid. Paré et al. (2002) review evidence indicating that

https://doi.org/10.1017/S0140525X0500004X Published online by Cambridge University Press

theta-band synchronization between HPC and AM neurons is specific to emotional states. Not only does theta synchrony produce recurring time windows that permit synaptic interactions, but it also amplifies the potency of AM neurons that otherwise fire infrequently. These authors conclude that theta-band synchrony promotes learning by facilitating interactions between cortical and limbic regions in the presence of emotion. This suggestion fits well with the present analysis.

5.5.1. Novel predictions. I have suggested that continuity in our responses to emotionally significant events emerges from the integration of appraisal contents and emotional response patterns during coherent EIs. However, attentional and evaluative processes in corticolimbic regions, and the neurochemical patterns that support them, must remain integrated for some period of time for this learning to take place. Only then can the emotional and attentional contents of whole-brain states become available for memory consolidation. A simple prediction based on this idea is that longer-lasting EIs will be more likely to recur than shorter ones, all other things being equal. For example, lasting emotional states in childhood should be good predictors of future appraisal–emotion constellations, compared with emotional states that resolve quickly (Lewis 2000a). If such states are mediated by corticolimbic synchronization at theta frequencies, then the consistency and duration of theta oscillations should mediate these predictions. Moreover, the amplitude and consistency of action-monitoring ERPs, arguably tapping theta synchrony, might make useful predictor variables. Thus, the amplitude and consistency of medial-frontal ERPs, measured in early childhood in an emotion-induction task, should be capable of predicting personality constellations emerging over development. Or, on a shorter scale, emotion-induced medial-frontal ERPs should be predictive of the outcomes of treatment for various behavior problems. My colleagues and I are presently investigating both these hypotheses (see Lewis & Stieben [2004] for an outline of our approach).

6. Putting the pieces together: A neuropsychological model

In this section, I refer to the anatomy and function of brain regions involved in appraisal and emotion, and the mechanisms of integration connecting them, to flesh out the psychological model of self-organizing emotional interpretations (EIs) presented earlier. This is not intended as a detailed theory but as a plausible direction for integrating psychological and neurobiological insights through DS modeling.

6.1. Trigger and self-amplification phases

Triggers were characterized as phase transitions induced by perturbations that disrupt the orderliness of a baseline state (e.g., an existing appraisal), increase sensitivity to environmental circumstances, and result in novel proto-organizations. Perceptual events often trigger emotional episodes, but so do memories or even physiological events such as pain. Thus, the mechanism for triggering an EI should be viewed as highly general. Two mechanisms of neural integration collude to trigger global change: the dynamic properties of feedback loops and the global character of ascending neuromodulatory effects. Because the neural systems mediating appraisal and emotion processes are interconnected in feedback loops, they are poised for sudden shifts to positive feedback dynamics. However, ascending neuromodulator flow, usually initiated by structures within these systems, must activate multiple components to participate in this feedback. For example, the OFC, NAS, and especially the AM induce neuromodulator release in BS/BFB structures. Neuromodulator pathways back up to these and other cortical and subcortical systems increase the activation of some systems while decreasing it in others. Where activation is increased, recursive downward connections further harness BS/BFB nuclei in a cycle of rapid self-amplification. Thus, the OFC, NAS, and AM are candidate structures that can induce positive feedback among multiple systems selectively activated by neurochemical afferents.

When new EIs are triggered through these activities, one's perception shifts, attention begins to orient to new events, affect alters or intensifies, and action tendencies arise. However, changes in the world or the body must be "meaningful" to cause the AM, OFC, or NAS to initiate new neurochemical patterns. Meaning is specified by the evaluative and motivational functions of these systems, according to innate encoding and learned associations (Rolls 1999). But meaning is also constrained by present appraisals and emotions. Thus, already-activated synaptic pathways within and across levels of the neuroaxis constitute a baseline EI to which new activation contributes. Baseline states not only determine thresholds for what is meaningful; they also constrain the character of any novel proto-organizations that emerge. For example, vigilant attention to strangers coupled with anxiety suggests a specific corticolimbic configuration rooted in fear neurochemistry. Sighting a policeman at such a time induces relief, because the OFC is already tuned to possibilities for rescue. At other times, given other preexisting states, sighting a policeman might induce worry instead ("Did I forget to renew my license?"). In keeping with the principles of cognitive self-organization, the present state of the system is the starting point for whatever happens next.

The emergence of an EI is first marked by rapid changes, in one or another higher-order loop, involving reciprocal and recursive augmentation among its constituent structures. For example, within the object evaluation loop, AM activation enhances the sensitivity of sensory cortex, increasing the processing of relevant environmental events, while reciprocal projections enhance AM activation, increasing the salience of learned associations. EIs then grow in strength and scope through the recruitment of additional higher-order loops. Structures such as the OFC, ACC, and AM (mediating evaluation, attention, and emotional memory) may serve as hubs for this expanding activity due to their widely distributed connections. For example, as Mr. Smart scans the road, ACC-mediated action monitoring ("I'm getting too close") and AM-mediated emotional associations ("illegal–police–powerful") together recruit other brain systems mediating feelings (e.g., shame and fear), action tendencies (e.g., urge to escape), and plans (e.g., passing on the right). The integration of appraisal and emotion processes within each loop now extends across loops with complementary functions, such that evaluation, monitor-

https://doi.org/10.1017/S0140525X0500004X Published online by Cambridge University Press

ing, and action functions resonate with each other and with an emerging emotional state.

6.2. Self-stabilization phase

The complexity of neural connections entails interactions between positive and negative feedback in all brain activity. In a given EI, brain states begin to cohere and stabilize when negative feedback predominates among activated sets of structures. For example, the orbitofrontal cortex (OFC) appears to inhibit amygdala activation in normal processes of emotion regulation (Davidson et al. 2000), but ongoing emotion regulation implies continual recruitment of orbitofrontal evaluation by amygdala associations, thus stabilizing the activities of both structures. Speaking more broadly, interactions among all loops may stabilize as attentional orientation, perceptual entrainment, and action monitoring narrow the focus of perception, cognition, and behavior to an emotionally relevant stance. Thus, as positive feedback recruits more neural components to an emerging constellation, negative feedback couples them in a stabilizing regime. At the level of the macrosystem, negative feedback is manifested by the satisfaction of multiple constraints, such as neuronal activation and receptivity patterns across multiple systems, as well as parameters of phase synchrony with co-activated neurons. This shift to coherence may underlie the consolidation of recognition, meaning, and "sense" as an EI settles into place. As discussed earlier, ERP data suggest a minimal time course of roughly 600 to 800 msec for this to take place.

The activation of structures at different levels of the neuroaxis drives the brain into vertical integration. Reciprocal influences among BS, hypothalamic, and corticolimbic regions become coupled, perhaps through synchronization of independent oscillators, and this coupling may give rise to an emergent meta-synchronization that coordinates all lower-order couplings. Vertical integration assures that EIs have some degree of psychological momentum or imperturbability. This is a result of downward control by the superordinate pattern on interactions among constitutive structures. However, it also restricts the repertoire of viable EIs. At the subordinate level, connective circuitry, pathways of neuromodulation, and parameters of phase synchrony may constrain which cells enter the dance. At the superordinate level, only particular global constellations are sufficiently coherent to entrain their constituents, partly because of the animal's history of learning. Thus, vertical integration not only maintains but also selects viable EIs.

Emotional states demand action, and vertical integration includes links between striatal structures that narrow the focus of action and ACC-mediated action monitoring. This coupling may extend broadly across corticolimbic systems as potential plans are constructed, rehearsed, or discarded. Thus, action monitoring both broadens the scope of appraisal and narrows its focus, and both may be critical for stabilizing an EI. Anticipation of action and feedback concerning action outcomes are intrinsically motivating. Underlying this motivation, continuous activation of BS/BFB structures maintains particular synaptic configurations. Neuropeptide modulation of brain and body systems may also play an important role in the stabilization of appraisal, by recruiting multiple systems to a unified behavioral tendency. Mr. Smart's angry, blameful state can be seen in terms of an ongoing action plan, and this plan maintains the integrity of his appraisals as well as his physiology. However, the writer choosing her words is equally anchored by action monitoring to an enduring appraisal of her communication with her readers. Thus, parallel anatomical and chemical mechanisms may support the selection and maintenance of a behavioral orientation that coordinates appraisal and emotion processes.

Coordination among the components of a system is necessary for coherence; but it also permits the evolution of intricate and complex activities. I have not done justice to the complexity of real appraisals, focusing instead on basic mechanisms and simple examples. But action plans provide a useful springboard for thinking about complexification. Integration among corticolimbic structures permits focal attention, planning, and memory to work together to elaborate detailed models of one's engagement with the world. New EIs can also be triggered while plans are constructed and revised, as cortically mediated images are reorganized by new patterns of activity in the OFC, ACC, AM, or NAS. For example, planning one's slides for a conference presentation can initiate embarrassing images of criticism or grandiose images of praise. Here, AM-mediated associations to anticipated outcomes recruit explicit memories through connections to the HPC, harness BS/BFB systems that alter arousal and affective feeling, and induce a cascade of new associations mediated by cortical regions. When Mr. Smart thinks about slowing down, he may see himself as a submissive weakling with accompanying feelings of helplessness and shame. Oscillations between coherent appraisals – angry empowerment versus shameful submissiveness – may provide a more realistic account of Mr. Smart's state of mind, with the seeds of each appraisal embedded in ruminations stemming from the other.

In sum, the self-stabilizing phase of an emergent EI may involve the integration of multiple feedback loops and neuromodulatory systems in an emergent synchrony entrained to action. Moreover, new appraisals can emerge from the complexification of existing appraisals, producing a cascade of psychological states or an oscillation between attractors. By this analysis, the component systems underlying appraisal and emotion become synchronized and resynchronized, often rapidly and recursively, perhaps en route to an EI that settles more permanently – such as a lasting moodlike state (Lewis 2000a). Consistent with Frijda and Zeelenberg (2001), appraisal elaboration can be seen as an outcome rather than an antecedent of emotional response.

6.3. Learning and development

Through self-synchronizing processes, the brain may achieve periods of relative stability, when the constellation of neural interactions is constrained by a global EI and accompanying action orientation. Such stable emotional states may be critical for learning. I have reviewed evidence for emotional enhancement of memory formation, exemplified by neuromodulator (and neuropeptide) facilitation of LTP. I have also suggested, on the basis of recent findings, that phase synchrony among corticolimbic systems is enhanced and maintained during emotional states, and that theta-band synchrony contributes directly to LTP. Taken together, these arguments suggest that stable EIs facilitate synaptic plasticity and lasting synaptic change. The associations that get altered and the memories that get laid down during these periods would reflect the synaptic configura-

https://doi.org/10.1017/S0140525X0500004X Published online by Cambridge University Press

tions mediating the contents of attention, evaluation, and planning. Hence, what gets learned is the present appraisal, or the tendency to appraise a situation in this particular way and not in some other way, given a particular emotional state. Across several occasions, an accumulation of learning events would then be expected to narrow the degrees of freedom for interpreting any subsequent event of this class (see Tucker 2001, on motivated memory and the stability–plasticity dilemma). This may be the basis for consolidating individual styles of interpretation, feeling, and belief, in a self-organizing process spanning years rather than seconds.

Learning based on appraisal stabilization has unique implications for the relation between emotional episodes and emotional development. On each occasion, the coordination and stabilization of neural interactions forms a gestalt, and this gestalt leaves its trace on some of its component structures. Thus, on subsequent occasions, emerging configurations that include these structures are influenced by their previous history and tend to repeat it. It is only through the synchronization of appraisal–emotion processes within occasions (real-time self-organization) that their influence can be transmitted over longer time scales (developmental self-organization), producing rich and enduring personality patterns on the basis of recurring states. However, this continuity is also mediated by real-time outcomes: in particular, the tendency for vertical integration to consolidate in some configurations and not others. Learning changes tendencies for parts to adjust to each other, but it also changes tendencies for wholes to converge in coherent gestalts. What gets learned gets repeated because the multistability of the brain is not infinite. Only certain constellations work within occasions, and their imprint on synaptic networks makes them all the more likely to work on future occasions.

Accounts that explain personality development on the basis of cognitive-emotional interactions can be made more precise through attention to these neural processes. These accounts include Izard's (1984) model of enduring affective-cognitive structures and Magai and McFadden's (1995) view of emotion traits stemming from recurrent interpretations accompanied by emotions. In addition, clinical syndromes of many kinds (e.g., depression, anxiety disorders) include a learning component. Whereas emotion theorists restrict their analysis to the effects of clinical traits on emotion and appraisal, the present model suggests that causation flows the other way as well, from recurring emotion–appraisal states to enduring clinical patterns. In general, the crystallization of personality or clinical traits is difficult to understand without a principle that allows for real-time sensitivity in brain processes that nevertheless become more entrenched with time. The linked requirements of emotional activation within occasions and synaptic shaping across occasions suggest such a principle.

7. Conclusions

After identifying some of the limitations of conventional approaches in emotion theory, I have emphasized the need for a more comprehensive and realistic account of appraisal–emotion interaction. I have approached this goal by outlining a dynamic systems (DS) framework, developing a psychological model based on that framework, and then shifting the focus from psychological to neurobiological processes. After providing a DS-based analysis of the neural underpinnings of emotion–appraisal states, I revisited the psychological model, integrating constructs and findings from neurobiology with those of psychology. To what extent can this modeling help to enrich and unify approaches within emotion theory, and to what extent can it facilitate further efforts to bridge emotion theory with neurobiology?

In general, appraisal theorists concentrate on the perceptual and cognitive evaluations that give rise to emotion, functionalist approaches concentrate on the cognitive changes produced by emotion, and trait theorists examine individual differences in cognition–emotion interactions with little attention to their origins. The limitations of these approaches and their relative insularity from each other have been shown to derive, in part, from a reliance on simple one-way causal explanations, an emphasis on wholes rather than constituent processes, and a commitment to linear cognitivist assumptions even when process-level analysis is attempted. Exemplifying these issues, process-level accounts within appraisal theory were argued to maintain linear modeling strategies that denied emotion a causal role in the formation of whole appraisals.

Dynamic systems principles highlight reciprocal, multiple, and recursive causality in both the emergence and stabilization of emotion–appraisal states. They also highlight multiple constraints and circular causality in the consolidation and endurance of these states. When the same principles are applied to a neural analysis, it can be demonstrated that neural activities underlying appraisal processes either mediate emotional processes simultaneously or they rapidly interact, evolve, and become integrated with activities mediating emotion processes, even at moderate levels of system architecture, well below the whole-brain scale. Thus, the nonlinear causal interactions and spontaneous synchronization of neural systems that give rise to global emotional states are fundamental to the emergence of whole appraisals evolving in tandem with them.

This reanalysis indicates that coherent appraisals are not antecedents of emotions, but emerging outcomes of interactions among constituent systems underlying appraisal *and* emotion. This conclusion facilitates the integration of appraisal and functionalist approaches in emotion theory: The cognitive consequences of emotions and the cognitive antecedents of emotions are the same, separated only by an arbitrary timeline. In other words, emotional and cognitive processes influence each other continuously during an emotional episode, from the first neural changes induced by a triggering event to the synchronization of the entire nervous system in a coherent mode of thinking, feeling, and acting. Given these considerations, the function of emotion can now be analyzed within appraisal processes themselves. Furthermore, the DS principle relating real-time coherence to structural change provides a means for integrating the third branch of emotion theory, dealing with individual traits, by suggesting how traits might arise from, as well as contribute to, cognition–emotion outcomes. Enduring patterns of synchronization among neural systems mediating emotion and appraisal are necessary for the synaptic changes responsible for learning and development. In all these applications, DS constructs provide a foundation for integrating disparate strands within emotion theory and for achieving compatibility between emotion theory and neurobiology, resulting in greater breadth, realism, and detail.

This analysis can revitalize emotion theory in another

https://doi.org/10.1017/S0140525X0500004X Published online by Cambridge University Press

way, while contributing to neuroscientific thinking at the same time. Questions posed from the perspective of emotion theory concern the relation between emotional and cognitive systems. A DS analysis of the psychology of emotion frames these questions in terms of part–whole relations and nonlinear causation, bridging psychological and neuroscientific descriptions. A DS analysis of the neurobiology of emotion then demonstrates, in precise and concrete terms, that cognition and emotion were never two distinct systems at all. They are not distinct at the level of interacting parts, nor at the level of the wholes to which these interactions give rise. Neuroscience relies on psychological definitions of functions such as appraisal and emotion, in order to ensure that the forest is not lost through the trees. Some conceptualization of higher-order functions is necessary to identify the elusive wholes emerging from interactions among relatively well-known parts. But the conceptualization of these functions – their definitions and boundaries – may get transformed in the process, and this is what neuroscience offers psychology in return. Although it remains useful to differentiate cognition and emotion for many research agendas, a neuroscientific analysis finds them to be different aspects of a unitary phenomenon in which interpretation and relevance emerge together.

ACKNOWLEDGMENT

I gratefully acknowledge the financial support provided by grant No.1 R21 MH67357-01 from the Developmental Psychopathology and Prevention Research branch of the National Institute of Mental Health (NIMH), as well as ongoing support from the Natural Sciences and Engineering Research Council of Canada (NSERC). I would also like to thank Nico Frijda, E-man Leung, Phan Luu, Jaak Panksepp, Brian Parkinson, Klaus Scherer, Evan Thompson, Don Tucker, and Phil Zelazo for their thoughtful comments on earlier drafts of this article.

Open Peer Commentary

Why not emotions as motivated behaviors?

George Ainslie[a] and John Monterosso[b]

[a]*Department of Veterans Affairs Medical Center, Coatesville, PA 19320;* [b]*Department of Psychiatry and Biobehavioral Sciences, University of California, Los Angeles, Los Angeles, CA 90024.*
George.Ainslie@med.va.gov jmont@ucla.edu
http://www.picoeconomics.com

Abstract: Lewis's dynamic systems approach is a refreshing change from the reflexology of most neuroscience, but it could go a step further: It could include the expected rewardingness of an emotion in the recursive feedback loop that determines whether the emotion will occur. Two possible objections to such a model are discussed: that emotions are not deliberate, and that negative emotions should lose out as instrumental choices.

We suggest that a key element, motivation *for* emotion, should be included in Lewis's rubric for bridging emotion theory and neurobiology. Despite his admirable departure from reflexology, he follows conventional assumptions that emotions and the neural processes that subtend them are unmotivated occurrences, elicited by specific patterns of stimuli without regard to their hedonic consequences. He proposes a thorough integration of these occurrences, both a horizontal one across "[complementary] evaluation, monitoring, and action functions" (sect. 5.1) and a "vertical integration [that] not only maintains but also selects viable EIs" (sect. 6.2). He does mention motivation at various points, including a "motivated action loop" of dynamic process (cf. Figure 3 in sect. 5.1 of the target article), but he is apparently referring to behaviors motivated by emotions: "Anticipation of action and feedback concerning action outcomes are intrinsically motivating" (sect. 6.2). It is well understood that emotions motivate; the interesting possibility is that they *are motivated*.

Many common experiences suggest that emotions can be both cultivated and nipped in the bud. For instance, you can "swallow" your anger or "nurse" it, you can learn to inhibit your phobic anxiety (Marks & Tobena 1990), panic (Clum et al. 1993; Kilic et al. 1997), or grief (Ramsay 1997), and you can refrain from rejoicing or "give yourself over to it." The road rage in Lewis's example (sect. 3.3) does not attack a passive Mr. Smart, but lures him in competition with other available activities. As Lewis points out, angry empowerment vies with shameful submissiveness (sect. 6.2), and, we could add, with enjoying his radio program or worrying about his impending sales conference. The choice first of all is whether to entertain the emotion or not, before (or in conjunction with) the choice of whether particular actions would go well with it. Techniques to deliberately foster or inhibit emotions in everyday life have been described (Parrott 1991); and most schools of acting teach an ability to summon emotion (e.g., Strasberg 1988) because even in actors actual emotion is more convincing than feigned emotion (Gosselin et al. 1998). The frequent philosophical assertion that emotions have a moral quality – good or bad (e.g., Hume, as presented by Baier 1991) – implies motivated participation; some philosophers have gone so far as to call the passions voluntary (e.g., Sartre 1939/1948). Thus, there have been many suggestions that emotion is a motivated activity. Lewis characterizes the competition of an emotion with alternatives as "oscillations between coherent appraisals," which might mean just between appraisals of whether triggering elements are present; but all affective appraisals include hedonic value, positive or negative, and thus could be capable of motivating choice. Mr. Smart is lured, not forced, into rage.

Psychological theories have been prevented from acknowledging the motivated quality of emotions by two considerations: that deliberately emitted emotions seem inauthentic (Frank 1988) and that a motive to experience "negative" emotions such as panic and grief seems contradictory. Ainslie (2001) has indeed argued that emotional processes which are under deliberate control do not behave in the same way as spontaneous ones – that they deteriorate into daydream status through familiarity and consequent premature satiation, leaving those which are occasioned involuntarily by surprising events as the authentic kind (Ainslie 2001, pp. 48–70). However, this is not to say that the latter kind are independent of reward. The class of reward-shaped behaviors includes many examples that arise too quickly, are too strongly motivated, or are just too trivial to be screened by the will. Direction of attention, withdrawal from painful stimuli, and numerous mannerisms and facial expressions can all be determined by reward without being deliberate. Conscious choice is a special case, the tip of a great iceberg of reward-seeking processes that include even behaviors during sleep (Granda & Hammack 1961).

The greater problem is how individuals could be lured into aversive emotions. Ainslie (2001) has argued also that reward of very short duration can motivate acceptance of experiences that are avoided from a distance and are reported as negative. The highly bowed (hyperbolic) shape of the discount curve for delayed rewards can produce temporary preferences for addictive substances, and, cycling more rapidly, for excoriating itches and biting fingernails; the same shape could lead to an urge to panic that is almost irresistible despite instant regret if it is obeyed (Ainslie 2001, pp. 48–70). The key word here is *almost*. With adequate

https://doi.org/10.1017/S0140525X0500004X Published online by Cambridge University Press