

[Penultimate version. Final version published in 2009 in *Philosophical Psychology*, 22, pp. 407-425. Please refer to the latter if quoting]

From Affect Programs to Dynamical Discrete Emotions

Giovanna Colombetti
Sociology and Philosophy Department
University of Exeter

Abstract

According to Discrete Emotion Theory, a number of emotions are distinguishable on the basis of neural, physiological, behavioural and expressive features. Critics of this view emphasize the variability and context-sensitivity of emotions. This paper discusses some of these criticisms, and argues that they do not undermine the claim that emotions are discrete. This paper also presents some works in dynamical affective science, and argues that to conceive of discrete emotions as self-organizing and softly assembled patterns of various processes accounts more naturally than traditional Discrete Emotion Theory for the variability and context-sensitivity of emotions.

From Affect Programs to Dynamical Discrete Emotions

1. Introduction

Several emotion theorists claim that there is a limited number of *basic* emotions, thought to be pancultural. Silvan Tomkins (1962, 1963) originally called these emotions *affect programs*, and argued that they are genetically determined products of evolution, distinguishable on the basis of their neural, bodily, behavioural and expressive features. His view influenced the work and ideas of emotion theorists such as Paul Ekman and Carol Izard, whose approach has come to be known as *Discrete Emotion Theory*.

Discrete Emotion Theory has been very influential. Yet it has also been challenged and criticized from various standpoints, and attempts have been made to advance different conceptualizations. In what follows, I shall first present the main tenets of Discrete Emotion Theory (section 2), and then discuss some such challenges and criticisms (section 3), as well as some alternative conceptualizations (section 4). I shall focus, in particular, on the criticisms that emotion psychologists James Russell and Lisa Barrett have raised in recent years, and on their alternative proposals.¹ Russell and Barrett claim that Discrete Emotion Theory cannot account for the rich variability and context-sensitivity of emotions, whereas their alternative models can. Barrett's view is that emotions have no existence independent of our categorization, and some of Russell's claims lend themselves to be interpreted as supporting a similar view. Against Russell and Barrett, I will proceed to argue that Discrete Emotion Theory does indeed have the resources to accommodate variability and context-sensitivity (section 3). In addition, I will argue that Russell's and Barrett's alternative proposals are inadequate (section 4).

This discussion will allow me to move on to the *dynamical systems* view of emotions (sections 5 and 6) which, I shall argue, still portrays emotions as *discrete*, and yet is not committed to the affect-program view; one can hold a discrete emotion view, while maintaining neither that there are basic emotions, nor that discrete emotions are affect programs. As illustrated by recent work in what I shall call "dynamical affective science," one can conceive of discrete emotions instead as self-organizing and softly assembled patterns. This conceptualization has the advantage that it naturally accommodates the variability and context-sensitivity of emotions; in addition, it does not suffer from the shortcomings of the alternative proposals advanced by some of the critics of Discrete Emotion Theory.

This paper, then, has two major aims. The first aim is to clarify the debate between Discrete Emotion Theory and some of its critics, in particular emphasize that abandoning the notion of “affect programs” need not imply abandoning the view that emotions are discrete, i.e. distinguishable on the basis of their neural, physiological, behavioural and expressive features; also, it is possible to do justice to the variability and context-sensitivity of emotions, without accepting the view, advanced as an alternative to Discrete Emotion Theory, that emotions have no existence independent of our categorization. The second aim is to draw attention to the dynamical systems approach to emotion, which is increasingly widespread in the psychology of emotion, but which has not received sufficient consideration in the philosophy of emotion.² There are now several affective scientists who believe that the concepts of dynamical systems theory can help us better understand a variety of affective phenomena. Their work and ideas have been relatively marginal within emotion research; also, they have presented their views in relative isolation from one another. This paper therefore aims to bring some of this recent work together, to make its theoretical assumptions and implications explicit, and to place them within the broader context of (some of) the ongoing debates in affective science.

2. Discrete Emotion Theory

The idea that at least some emotions have distinctive features is traditionally traced back to Darwin (1872), who described several facial, physiological and behavioural processes associated with different emotions in humans as well as animals, and laid out the principles meant to explain the function of such processes and why they evolved. In the same spirit, James (1884) and Dewey (1894) viewed emotions as associated with different neural and physiological processes, different functions and different experiences. We can find similar ideas further back in time. For example Aristotle, in the *De Anima*, observed that bodily changes always accompany changes in emotion (see Aristotle, 1986, pp. 128-129). In *The Passions of the Soul*, Descartes provided a detailed description of the various bodily processes that accompany different emotions in humans, as well as the various feelings that these bodily processes cause in the mind. He also indicated that emotional bodily processes have specific survival-related and self-regulatory functions, and even attempted to explain why specific bodily processes are accompanied by specific experiences.

Tomkins' (1962, 1963) idea that there is a limited number of pancultural basic emotions or “affect programs” was influenced by his reading of Darwin (1872). On the basis of the latter's observations, Tomkins concluded that there are *eight* pancultural affect

programs (surprise, interest, joy, rage, fear, disgust, shame and anguish).³ Some years later, Ekman and Izard independently conducted a series of cross-cultural studies to test Tomkins' hypothesis; they reported, in several works, various similarities in the way people across the world (including "visually isolated cultures") produce and recognize the facial expressions of at least *six* emotions (for overview and references, see e.g. Ekman, 2003; Griffiths, 1997; Prinz, 2004). Since then, the view that there is a limited number of pancultural basic emotions or affect programs has come to be known, in the psychology of emotion, as Discrete Emotion Theory. The adjective "discrete" signifies that basic emotions are separate and distinct, and can be distinguished from one another on the basis of different features. Whereas Ekman and Izard originally focused on differences in facial expressions, more recent work has revealed further differences in patterns of autonomic nervous system activity (Ekman et al., 1983; Levenson, 2002), neural and chemical processes (Panksepp, 1998), and vocal expressions (Scherer et al., 2003).

Yet what exactly are affect programs? Tomkins (1962, 1963) claimed that they are genetically determined products of evolution, but did not say much about the mechanisms supposedly implementing them. Ekman (1980, p. 820) later defined affect programs as follows:

The term *affect program* refers to a mechanism that stores the patterns for these complex organized responses [i.e. the responses that occur with an emotion], and which when set off directs their occurrence. [...] The organization of response systems dictated by the affect program has a genetic basis but is influenced also by experience. The skeletal, facial, vocal, autonomic, and central nervous system changes that occur initially and quickly for one or another emotion, we presume to be in largest part given, not acquired.

This characterization appears also in Ekman's more recent work. Ekman (2003, p. 65) still talks of affect programs as "central information storages," and says that "[s]tored in these central mechanisms there must be sets of instructions guiding what we do, instructions that reflect what has been adaptive in our evolutionary past. [...] program refers to mechanisms that store information written before [...]."

Izard & Malatesta (1987, p. 496) use the term "discrete emotions," and write that "a discrete emotion can be defined as a particular set of neural processes that lead to a specific expression and a corresponding specific feeling." This definition does not say anything about

the structure of the neural mechanisms in question, yet it unambiguously portrays emotions as specific brain systems that produce specific responses.

In sum, then, we can define Discrete Emotion Theory as the view according to which at least some emotions correspond to specific brain systems which contain instructions for the generation of distinctive changes in the organism. These changes affect facial and vocal expressions, autonomic nervous system (ANS) activity, behaviour (including action tendencies), and experience. One important corollary of this view is that basic emotions can be “diagnosed” or “read-out” from how they are manifested in the organism.

3. Emotions, Variability and Context-Sensitivity

Discrete Emotion Theory has been highly influential among affective scientists. Ekman’s and Izard’s framework and methodology have inspired much of subsequent experimental developments, and their work has also informed philosophical accounts of emotion (e.g. de Sousa, 1987; DeLancey, 2002; Goldie, 2000; Griffiths, 1997; Prinz, 2004). Yet the assumptions of Discrete Emotion Theory have also raised criticisms from various camps. I shall here focus in particular on the criticisms raised by psychologists James Russell and Lisa Barrett (see Barrett, 2006a, 2006b, 2006c; Barrett & Russell, 1999; Russell, 1980, 1991, 2003, 2005, 2006; Russell & Barrett, 1999). Figure 1 depicts the view that both Russell and Barrett claim to attack in their work.

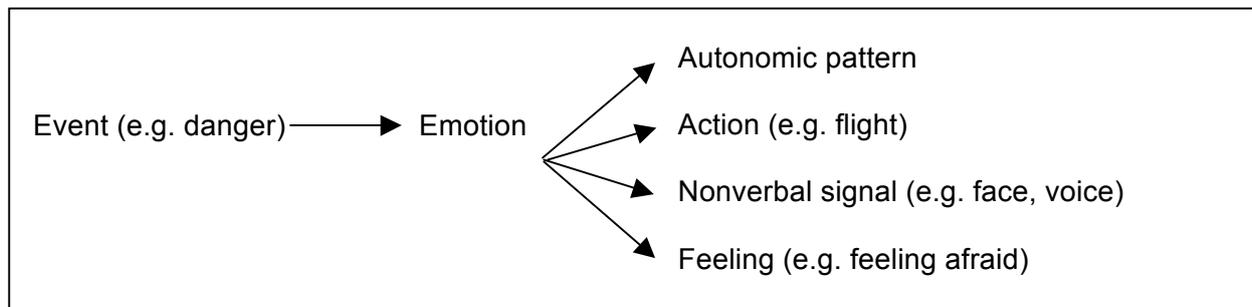


Figure 1: Russell’s representation of Discrete Emotion Theory (adapted from Russell, 2003, p. 151. A very similar figure is reproduced in Barrett, 2006b, p. 21)

Most of their criticisms target specifically the idea that there is a *one-to-one correspondence* between discrete emotions and their distinctive features. First, they point out that attempts to find out whether the same patterns of neural and ANS activity correlate with the same emotions across different studies have failed. Barrett (2006a) remarks that meta-analyses (Murphy et al., 2003; Phan et al., 2002) have failed to reveal reliable one-to-one correspondences between individual emotions and brain systems; different brain areas appear

to be implicated in the same emotion across studies and conditions (lack of consistency), and different emotions often overlap i.e. have some brain areas in common (lack of uniqueness or specificity). In addition, meta-analyses have shown that there is no agreement on the localization of alleged basic discrete emotions, such as e.g. anger and sadness; there is *some* agreement, in particular about fear (which is mostly associated with amygdala activation), but for Barrett the correspondences across studies are not as high as one would expect if the traditional view were right, and fear were a basic discrete emotion. Similarly, both Russell and Barrett note that the claim that different emotions are accompanied by specific patterns of ANS activation is questionable. Although several studies have shown differentiated ANS activation for emotions such as anger, fear, happiness, etc., meta-analyses do not reveal conclusive agreement on this issue (see Cacioppo et al., 2000); what they show, rather, is that individual emotions are not consistently related to the same ANS activity patterns across studies.

Russell and Barrett also emphasize the variability of facial and behavioural manifestations. They remark that faces do not univocally and unambiguously express specific emotions; a smile for instance might indicate happiness, but also embarrassment or desire to please others. In other words, factors other than one's immediate feeling can influence one's facial movements, and so we are not entitled to "read-out" emotions merely by observing someone's face. On their part, behavioural manifestations of emotions vary from context to context. Fear, for example, can involve fleeing, but also freezing or attacking, and anger can involve attacking, yelling, withdrawing, or being exceedingly kind.

Finally, Russell and Barrett note that "full-blown" basic emotions are rare; when we are happy, scared, angry, etc. we rarely seem to undergo *all* the events anticipated by Discrete Emotion Theory. Russell also often remarks that there are many "borderline" emotional episodes that Discrete Emotion Theory is not able to account for.

In sum, these considerations importantly emphasize that emotions and their manifestations are highly variable and context-dependent, and that "reading out" one's emotions is not a straightforward process; indeed, other emotion theorists had already raised similar criticisms (see e.g. Frijda, 1986; Scherer, 1984).

It should be pointed out, however, that Discrete Emotion Theory, in more recent versions at least, has made efforts explicitly to account for the variability and context-sensitivity of emotions. Ekman (2003), as we saw, still talks of affect programs as genetically determined information storages. Yet he also refers to Ernst Mayr's (1974) distinction between "closed" and "open" programs, and claims that affect programs are *open*, namely

that they can be modified by experience; experience can insert new instructions into one's inborn repertoire of affect programs, and thus modify one's emotional responsiveness. Variability and context-sensitivity of emotion are thus accounted for, here, in terms of developmentally modified affect programs.⁴ Izard et al. (2000, p. 17), on their part, define discrete emotions as “systems” of interacting neural, behavioural, expressive and experiential components, which influence one another through feedback loops. They still claim that the “basic structure” of each discrete emotion system is “hard-wired ontogenetically” and “innate” (*ibid.*). Yet they also claim that discrete emotions can interact with one another and “coassemble” (*op. cit.*, p. 15) into patterns that are genetically influenced but also depend on context, and are thus unique to each individual.

Ekman (2003, p. 58) also writes that “[e]ach of these emotion terms—sadness, anger, surprise, fear, disgust, contempt, and happiness—stands for a *family* of related emotions” (my emphasis), and adds that anger, for example, can vary in strength as well as type. He also mentions different enjoyable emotions (i.e. not just happiness) and does not exclude that further research will reveal that all of them have distinctive manifestations. Finally, he mentions that facial expressions can be “blends,” although he adds that his research has shown that blends happen less often than rapid sequences (*op. cit.*, p. 69).⁵

4. The “Categorization Act” Proposal and its Problems

Discrete Emotion Theory thus appears willing and able to adapt its conceptual resources to accommodate variability and context-sensitivity; some of his more recent versions explicitly underscore that emotional behaviour and expressions come in various forms, and allow “affect programs” to be modified by experience.

Yet accommodating variability and context-sensitivity in the affect-program framework comes at the expense of making the affect programs more complicated. I think that there is an alternative view of discrete emotions that lends itself more naturally to accounting for the variability and context-sensitivity of emotions. Before presenting it, however, I want to discuss Russell's and Barrett's own alternatives to Discrete Emotion Theory. Their critique of the idea that emotions are “internal causal entities” that trigger various changes has led them, I think, to go too far in the opposite direction. In particular, in this section I shall resist the suggestion that emotions are somehow individuated via acts of categorization.

This suggestion is explicit in Barrett (see in particular Barrett, 2006b), who writes for example that “emotions are not biologically given, but are constructed via the process of

categorization” (Barrett, 2006b, p. 27). Barrett emphasizes that the various processes that accompany emotional episodes are, at best, only weakly correlated. On her view, there is in fact no coherent and recurrent pattern in the organism that can be said to correspond to an emotion; rather, various processes become an emotion when one classifies one’s own state in accordance with available linguistic categories. It is categorization that parses (what are at best weakly correlated) processes into discrete items; categorization makes us experience our own emotions as discrete feelings,⁶ and it also “shapes” our perception of other people’s states.

Russell also emphasizes that emotional episodes are made up by component processes that are only “loosely related” (Russell, 2005, p. 38), as opposed to tightly connected outcomes of the same affect program.⁷ He writes, for example, that “[t]here is no need to postulate a mechanism that ties all the various components into one package” (Russell, 2003, p. 152). In addition, his claims sometimes lend themselves to being interpreted as supporting a categorization act proposal similar to Barrett’s. His view is that we recognize emotions (in us, as well as in other people) by matching perceived events with *mental scripts*, i.e. implicit folk psychological micro-theories that include information about the temporal unfolding of an emotional episode and about its causes (Russell, 2003, 2005). Russell (2003, p. 152) continues, “[p]rototypical emotional episodes are coherent packages, not because they stem from one mechanism, but because they are just those cases selected by an observer on the basis of their resemblance to a coherent package: the folk concept.” This (rather ambiguous) passage could be interpreted, I think, as meaning that the coherence of an emotional episode is provided by the mental script, and does not correspond to anything within the organism (brain, physiology, expression, etc.) itself.

In any case, my point here is that the “categorization act proposal” is not a viable alternative to Discrete Emotion Theory. The categorization act view does not explain how categories (and/or mental scripts) can ever come to be applied to the various components that eventually make up an emotional episode. The problem, as I see it, is that it is impossible to provide such an explanation if one does not also posit that *there are already* “coherent packages,” or patterns, in the organism. It must be because there are already such packages, that we are able to recognize and identify them by matching them with available categories and/or mental scripts. The latter provide criteria that constrain our interpretation of the perceived episode, but at the same time they can be called upon only if the perceived episode already possesses certain coherent features.⁸

Neither Barrett nor Russell ever explicitly address this point. In some passages,

Russell's view seems to be that, indeed, emotions do arise as coherent packages in the organism *independently* of mental scripts. For example, he claims that one realizes that one is scared when various components "form into a perceptual Gestalt with a specific meaning" (Russell, 2005, p. 34), and one recognizes that these components match the mental script for fear. Russell also believes (personal communication) that it is possible for someone to be in a specific emotional state without knowing it. For example, he thinks that all the components of an episode of anger may occur *save* the recognition, on the subject's part, that what is going on matches the anger script; an outside observer however would still be able to recognize the episode as one of anger.

Importantly, these latter claims do suggest that "coherent packages" must be there already in the organism, thus allowing either the subject or some observer reliably to pick them out. How otherwise could the subject recognize what emotion he or she is undergoing? And how could an observer conclude that someone else is angry?

Acknowledging this point does not commit one to the idea that there are affect programs, in Tomkins' and Ekman's sense. In the rest of this paper I will in fact argue that there is an alternative view of discrete emotions—a "third way"—that *does* posit the existence of "coherent packages" or patterns within the organism, and yet that does not assume that emotions are "internal causal entities" or "stored patterns" that somehow bring about tightly connected outcomes. This third way, as we will see, endorses the claim that emotions are variable and context-dependent; at the same time, it explicitly addresses the question of how coherent emotional episodes "come to be," so to say, and does it by reference to principles of organization that are widespread in nature and that do not appeal to internal causes to explain patterned phenomena.

5. A Third Way: Dynamical Affective Science

This third way is provided by what I shall call *dynamical affective science*, by analogy with "dynamical cognitive science" (see Wheeler, 2005). Like the latter, dynamical affective science makes use of the conceptual and modelling tools of dynamical systems theory for the purposes of description and prediction (see Lewis, 2005; Fogel et al., 1992; Scherer, 1984; see also the papers in Lewis & Granic, 2000).

Dynamical affective science, it must be admitted, has not produced detailed quantitative models of specific affective phenomena (and some may doubt that it ever will). Yet the *conceptual* tools provided by dynamical systems theory, and in particular the notion of *self-organization* (see below), appear to be particularly useful to conceptualize several

features of emotions. Although largely speculative, the dynamical systems approach to emotion appears to provide descriptive and explanatory tools that were previously lacking, and thus help us make some progress in our understanding of affective phenomena.⁹

Dynamical systems theory provides a set of mathematical and conceptual tools that describe how systems unfold over time. For present purposes, it is important that dynamical systems theory can describe the temporal unfolding of *complex* systems, namely systems composed of several interacting parts, and whose behaviour is the result of such interactions. Importantly, these interactions are so dense that it is not possible to track the way in which each individual part contributes to the system's overall behaviour.

In dynamical systems theory, the set of all possible states of a system is represented by a *state space*. The state space can have several dimensions, each of which corresponds to a variable of the system. Each point of the state space represents a possible state of the system; as the system changes its state over time, it traces a *trajectory* which is represented by a curve in the state space. Some sets of points in the state space constitute *attractors*, namely states toward which the system's trajectories tend to converge. There are three types of attractors: *attractor points* are stable states (endpoints) of the system; *limit cycle attractors* are cycles in the state space representing the repetitive behaviour of a system moving between the same states; *chaotic* or *strange attractors* are more intricate sets of points on which the motion is neither stable nor periodic.

An important concept of dynamical systems theory is *self-organization*. Self-organization is a widespread phenomenon in nature. It refers to the capacity of a complex system to maintain structured order within itself by way of mutual interactions among its components. Such interactions may bring about macroscopic changes in the system's behaviour, and these changes may in turn modify the behaviour of the component parts (this process of reciprocal influences is sometimes referred to as *circular causality*). An example often used to illustrate self-organization is the Rayleigh-Bénard instability (see e.g. Kelso, 1995). When liquid is poured into a pan and heat applied from below, rolling motions (*convection rolls*) in the liquid appear which constrain the range of possible behaviours of the liquid's molecules. Importantly, the appearance and unfolding of convection rolls is not dictated by a separate control system; the temperature gradient is called the *control parameter*, but this term is somewhat misleading, for the temperature gradient neither dictates how the molecules should move, nor monitors their behaviour as it unfolds. The initial disposition of the liquid's molecules and the temperature of the heat source applied to it influence the direction of motion and the amplitude of the convection rolls; at the same time,

the appearance of convection rolls influences (more precisely, *entrains*) the motion of the individual molecules. As Kelso (1995) puts it, there is no self in self-organization, and pattern formations occur spontaneously.

It is useful to think of self-organization in terms of *first-* and *second-order constraints*.¹⁰ In the example of the Bénard instability, the reciprocal interactions between molecules of liquid instantiate first-order constraints; the capacity of one molecule to influence the direction of motion of another molecule, for example, constitutes a first-order constraint. Macroscopic circular motions rather instantiate second-order constraints; they embody overarching patterns or *forms* that entrain the behaviour of the liquid's molecules.

In biological systems, self-organization allows complex systems to absorb, and adjust to, changes in the surrounding environment. When a change takes place in the environment which could potentially disintegrate the organism, the self-organizing structure of the latter allows it to reorganize and modify its relation to the environment in a way that is conducive to its survival. The macroscopic changes in behaviour that take place in self-organization provide an overarching pattern or form that becomes a principle of movement of, and yet within, the organism itself; the organism tends toward patterns of behaviour and internal organization that will make its survival and continuation possible. There is increasing evidence that self-organization is a prevalent feature of large-scale activity in the brain (see Le Van Quyen, 2003, for an overview). The brain as a whole undergoes continual change, yet it also organizes itself into transient or *meta-stable* patterns that can endure for some time while they sustain specific cognitive activities (Fingelkurts & Fingelkurts, 2004; Kelso, 1995; Thompson & Varela, 2001).

These concepts are particularly useful, I believe, to account for several of the features of emotion that Russell and Barrett have pointed out, without however committing us to the problematic view that the elements of emotions are only “loosely linked,” and/or that their coherence is provided by acts of categorization. Specifically, dynamical systems concepts appear to be particularly well suited to account for the fact that emotional episodes are highly variable and context-dependent, and yet exhibit recurrent patterns. My suggestion here is to conceptualize an emotional episode as the form or second-order constraint that emerges from the self-organization of various processes (i.e. neural, autonomic, behavioural, etc.), and that entrains such processes into a meta-stable configuration or pattern. Reciprocal interactions among these various processes embody first-order constraints and contribute to the formation of an emotion as an overarching form. Events in the environment—traditional “emotional stimuli”—also influence the emotion-form, and can be seen as control parameters (by analogy

with the temperature gradient in the Bénard instability). An emotional episode thus conceptualized can “vary” in the sense that the processes constituting it can organize themselves in different ways (by analogy, the convection rolls appearing in the Bénard instability can vary in how the liquid’s molecules organize themselves). Yet at the same time the range of its possible variations is constrained by the structure of the state space; areas of stability within the state space pull or entrain (variously constituted) emotional episodes towards them. The presence of areas of stabilities guarantees relative stability-in-spite-of-variations, and the capacity of various processes (neural, autonomic, etc.) to influence and constrain one another allows stability to be achieved in various ways.¹¹

This abstract characterization of an emotional episode can be made more concrete by looking at recent works in affective science that explicitly draw on dynamical systems concepts to make sense of emotional phenomena. One main feature of dynamical affective science, indeed, is that it emphasizes the *patterned variability* of emotion, in various ways.

Developmental psychologist Linda Camras, for example, appeals to dynamical systems concepts to account for the variability and context-sensitivity of facial expressions. In Discrete Emotion Theory, facial configurations are “expressions” in the sense of “outcomes” of an internal set of instructions; in other words, something “internal,” the emotion, is manifested “outside,” on the face (as well as in bodily arousal and behaviour). Camras sees facial configurations rather as “coordinative motor structures” involving muscle groups that are “synergistically related” (Camras, 2000, p. 103). On her view, there is no central system instructing when and how facial action muscles should contract—namely, no pattern-storing mechanism that “directs their occurrence,” as in Ekman’s definition of an affect program. Rather, “contextually sensitive adjustments are made at the lower levels through synergistic relationships among neuromuscular action components” (*ibid.*). In other words, facial expressions self-organize into patterns from the local synergies of various muscles.

As Camras argues, this view can explain what in Discrete Emotion Theory would count as problematic “mismatches” between emotion and expression. Such mismatches are often observed in infants (but not only); infants for example often display prototypical “surprise expressions” in familiar rather than surprising situations. For Camras a possible explanation of this phenomenon is that, in familiar circumstances, the facial configuration of surprise is brought about by the realization of one of its components, notably the brow raise, for the distinct purpose of facilitating desirable inputs. Because of the synergistic relations between muscles, the raising of the brows changes the configurations of other facial muscles and produces the “surprise expression.”

Another consequence of viewing faces as coordinative structures is that the occurrence of one specific facial movement is not necessarily a sign of the occurrence of one specific emotion. Facial configurations occur in contexts, and they seem to be highly sensitive to changes in such contexts. Five- to seven-month olds, for example, appear to raise their brows more often when presented with attractive toys at above eye level (so that they have to raise their head and gaze to look at them) rather than below. Supposedly the infants are equally interested in the two conditions, and yet the brow raise typical of interest appears more markedly in one condition than in the other. Likewise, it seems that the “BROM” (brow raise open mouth) configuration that is part of the standard surprise expression is more likely to occur together with upward head/gaze shifts of at least 45 degrees (see Camras, 2000, for details).

Camras’s suggestion, in sum, is that facial configurations are best conceived of as context-sensitive coordinative structures that recruit individual facial muscles; they are not (or, she claims, at least not always) the output of one internal cause i.e. of a specific emotion: “coordinative motor structures are available for recruitment for a variety of purposes and [...] unique and exclusive ties may not be formed between some emotions and their corresponding facial expressions” (*op. cit.*, p. 110). Her studies suggest that the face’s shape can be influenced in various ways—from an individual facial action with a specific purpose (e.g. brow raised) to a whole facial Gestalt, as well as from a contextual constraint (i.e. an attractive object presented at above eye level) to individual muscle movements.

Importantly, it is apparent that Camras does not think that a dynamical systems approach is *incompatible* with the notion of discrete emotions; what she rejects are the preformationist and mandatory features of the affect program view, specifically the idea that facial configurations are the effects of internal causes, i.e. the outputs of pre-given instructions. She does not appear to reject the idea that emotions are “discrete” in the more liberal sense of distinct self-organized structures that recruit different processes in different contexts. She concludes, in fact, that “adopting a dynamical systems perspective [...] allows us to retain the idea that facial behavior may express discrete emotions but to view these emotions as a ‘softly assembled’ dynamic system in which no one component is required” (*op. cit.*, p.110).

The idea of discrete emotions as *softly assembled* appears in other developmental-dynamical accounts of emotion. Fogel et al. (1992) suggested that emotions are developmentally open and sensitive to perturbations that may come from the environment from very early on, and should therefore be conceived as “a soft or flexible initial assembly of elements, each of which may developmentally affect the others” (*op. cit.*, p. 127).

In support of this view, Fogel and colleagues have emphasized the variability and context-specificity of the development of facial configurations, in particular of infant smiles. Several different smiles can be identified in infants: *simple smiles* (raised lip corners, with no other facial action), Ekman's prototypical *Duchenne smiles* (raised lip corners and raised cheeks), but also *play smiles* (dropped jaw; also called *open mouth smiles* or *gaping smiles*), and *duplay smiles* (raised lip corners, raised cheeks and dropped jaw) (Fogel et al., 2006). Most smiles in the first month of age are endogenous, namely they occur primarily during REM sleep and during transitions from waking to sleep; there are also early forms of exogenous smiles, which occur irregularly in relation to various forms of sensory stimulation. "Social" smiling elicited by interactions with the human face develops later, at about 2 months of age, together with the ability to maintain eye contact and visual attention for a longer time. The development of these capacities allows for longer face-to-face communication, during which infants appear to explore facial features systematically (see overview in Lavelli & Fogel, 2005).

Importantly, different smiles occur in different contexts (greeting a parent, or greeting a stranger) already from the second month. In addition, smile duration changes according to age and context; and different smiling patterns take place as infants grow older—for example, Duchenne smiles tend to follow simple smiles before the tenth month, whereas later on Duchenne smiles are more frequent when infants greet their mothers after a short separation, than they are when the infants greet strangers. Also, in six- to twelve-month olds, Duchenne smiles are more likely to occur while infants are being tickled than while they play peekaboo (Fogel et al., 2006).

Further appeals to dynamical systems concepts to describe affective phenomena have been made by Klaus Scherer and Marc Lewis. Scherer (2000) invokes dynamical systems concepts in his "componential model," according to which emotions are constituted by several subsystems (appraisal, arousal, motivation, expression, and feeling) that interact with one another; in particular, he distinguishes between five different "appraisal checks" which influence, and are influenced by, the other subsystems.¹² Emotions are here conceptualized as *patterns of synchronization* of the various subsystems; importantly, these patterns are highly variable and depend on context, i.e. on the outcomes of the various appraisal checks. Characteristic of this approach is the attempt to describe the correlations between *individual* appraisal checks, and various emotional manifestations (unlike Discrete Emotion Theory, which merely states that appraisals, broadly characterized, activate specific affect programs; see the discussion in Scherer & Ellgring, 2007).

Lewis, on his part, has used the concept of self-organization to describe the unfolding of emotional episodes at various time-scales, from the origin and micro-development of individual emotional episodes, to the development of moods and personality traits (see e.g. Lewis, 2000). The micro-development of emotional episodes is described in detail by Lewis (2005), who suggests that events in the world initiate a process of reciprocal influences between various constituents—such as attention, evaluation, bodily arousal and action tendencies. Within seconds or a few minutes, these reciprocal influences lead to an *emotional interpretation*, namely the convergence of a cognitive interpretation and an emotional state. Specifically, the emergence of an emotional interpretation begins as a fluctuation in the ongoing stream of intentional action; this fluctuation is triggered by a perturbation (external or internal), which eventually disrupts the orderliness of the current emotional interpretation. Rapid processes of self-amplification through positive feedback ensue, followed by self-stabilization through negative feedback and entrainment, leading to the establishment of a new orderliness in the form of a new momentary emotional interpretation and global intention for action. Lewis (2005) likens the whole process to a movement from one attractor to another in an emotion-cognition state space, and has presented a neuropsychological model of some of the brain areas and large-scale neural-integration processes involved.

6. Dynamical Discrete Emotions: Implications

This overview shows that there are now several works in “dynamical affective science” that, despite differences in emphasis, are all consistent with the suggestion to conceive of an emotion as a “form” or second-order constraint that emerges from the self-organization of various processes, and that entrains these processes into a meta-stable configuration or pattern. These works all underscore the variability and context-sensitivity of emotions, as well as their patterned character; they reject the suggestion that discrete emotions are outcomes of information stored in some kind of central control mechanism (a distinct affect program), and yet they do not deny that there are mechanisms whereby emotions turn out to be coherent patterns with distinguishing features.

From a dynamical systems perspective, it is not only conceivable, but also actually *expected* that emotional episodes will come in several different forms. In the view sketched here, an emotional episode emerges from the reciprocal interactions of several processes, and the way in which such interactions take place is sensitive to contextual conditions. Hence it is not surprising that one does not find consistent patterns of neural and physiological activation in meta-analyses that compared studies in which emotions had been elicited in very different

ways (see above, section 3). From the present perspective, the experimental setting provides control parameters that will influence the mode of organization of various processes, and ultimately bring about an emotion-form. Variations in experimental settings should thus be expected to induce different modes of organization, and thus different emotion-forms.

The picture so far would seem consistent with Russell's and Barrett's considerations and alternative models. But there is a major difference. The dynamical perspective advanced here rejects the suggestion that there isn't anything within the organism that somehow "glues together" various processes into an emotional *episode* (recall Russell's claim that "[t]here is no need to postulate a mechanism that ties all the various components into one package" (Russell, 2003, p. 152)). The idea is rather that neural, bodily, and behavioural processes self-organize into patterns (emotion-forms) that sustain themselves through circular causation; in other words, self-organization is the mechanism that ties components into packages. Self-organization allows emotion-forms to be developmentally open and sensitive to contextual variations, and yet stable and patterned. Crucially, the existence of such patterns explains why we are able to identify emotional episodes and match them with their scripts.

To view emotions as softly assembled self-organizing patterns is consistent and continuous with the approach to emotion traditionally traced back to Darwin. Dynamical affective science situates itself in the Darwin-James-Dewey tradition, according to which emotions can be distinguished on the basis of their physiology, behavioural tendencies, functional relations with the environment, etc. Yet it is not committed to the notion of affect programs; from a dynamical systems perspective, there are no "internal" mechanisms storing instructions that completely specify patterns of emotional expression, action, physiology, etc., and the very notion of an "internal program" is dispensed with.

In addition, from a dynamical systems perspective no emotional episode need be more "basic" than others. If each emotional episode can be seen as a pattern of softly assembled elements, with its own history (trajectory), then emotional episodes that recur reliably across cultures in similar ways appear to be no less complex (or more basic) than idiosyncratic personal emotional reactions. The view is rather that some emotional episodes are phylogenetically old patterns that we share with other primates, whereas idiosyncratic personal emotional episodes are patterns that develop in the course of one's lifetime. In other words, their difference is merely one of history.

Finally, although I have been talking of "emotional episodes" and "emotion-forms" throughout, it is now time to clarify that, as I see it, the suggested dynamical perspective need not posit a clear-cut distinction between emotional and non-emotional episodes, nor need it

posit a clear-cut distinction between emotional and cognitive phenomena. In traditional Discrete Emotion Theory, basic emotions and blends thereof punctuate an otherwise affectively neutral landscape. From a dynamical systems perspective, the picture is rather one in which neural and bodily processes fluctuate and unfold over time, and every now and then converge into patterns that we have come to label as specific emotions. But there is no void between patterns, so to say, and the organism is always undergoing some process and making sense of its environment. Fogel et al. (1992, p. 133) nicely put this thought as follows:

In our view, emotions are continuous and always present. Periods of time that are often called ‘emotional’ reflect particular dynamic processes that motivate a change in the individual’s relationships to the context. These intense periods cannot be isolated from prior and subsequent periods of relative relaxation, characterized by low levels of engagement. These are merely different patterns of organization of the components of the same emotion system. [...] If there are periods that are non-emotional, that would suggest the individual is no longer linked to the environment in any meaningful relationship [...]

Emotions, from this perspective, are not bounded experiences and/or bodily stirrings, strictly distinct from non-emotional phenomena. Instead they are patterns of organization of natural processes that are nevertheless always present, and it is not clear when and where what we identify as “emotions” begin or end. To clarify, this claim is not inconsistent with the view that *there are* coherent patterns or emotion-forms within the organism. The point is rather that the various components that organize themselves into patterns which we call “emotional episodes” are constantly fluctuating and mutually influencing one another, so that various patterns are constantly done and undone. What we call “emotional episodes” are only a subset of the several forms that our complex dynamic organism can realize.

7. Summary and Conclusion

Discrete Emotion Theory is associated with a school of thought that is traced back to Darwin (1872), and that has been largely influenced by the work and ideas of Tomkins, Ekman, Izard and their collaborators. This approach emphasizes that emotions can be identified and differentiated in terms of their distinctive features, but also relies heavily on the relatively recent theoretical construct of “affect programs.” In this paper I have argued that it is possible to do justice to Darwin’s work (1872), and yet not endorse such a construct. This

possibility is illustrated by recent work in dynamical affective science, i.e. the attempt to adopt the conceptual and modelling tools of dynamical systems theory to describe and predict emotional phenomena. This approach, I have argued, offers a perspective on the notion of discrete emotions that does not deny that emotions are coherent patterns, but one that also naturally accommodates variability and context-sensitivity. Dynamical discrete emotions, as I have called them, are coherent “forms” that emerge from the self-organization of various natural processes (neural, autonomic, etc.), and that entrain such processes into meta-stable configurations. Dynamical discrete emotions are thus “softly assembled,” namely constituted by patterns of various processes that organize themselves differently in different contexts (as opposed to being manifested through mandatory configurations of various signature processes).

Thus conceived, discrete emotions are flexible enough to withstand and absorb various criticisms raised to traditional Discrete Emotion Theory. The notion of dynamical discrete emotions does not in fact identify emotions with “internal causes,” and accordingly it need not posit complicated programs to account for variability and context-sensitivity. As softly assembled, dynamical discrete emotions naturally come in various forms; some of these forms will depend more than others on the species’ history, whereas others will be moulded in the course of one’s lifetime.

Significantly, the notion of dynamical discrete emotions is consistent with recent accounts of the structure and organization of biological systems, including the brain. Self-organization, meta-stability and patterned variability are widespread features in the natural world, and accounts of emotion thus ought to take these features into account. Emotions are complex phenomena that engage the whole organism, and it would therefore seem important to try to approach them with conceptual tools that reflect existing views of the structure and organization of living systems.

Acknowledgements

I am grateful (in chronological order) to Jim Russell, Ezequiel Di Paolo and the Life & Mind discussion group at Sussex University, Paul Griffiths, Wayne Myrvold, Klaus Scherer and two anonymous referees for their valuable comments at various stages of preparation of this work.

Notes

1. Russell and Barrett are known for defending a *dimensional* approach to emotion. I should anticipate that I am *not* going to address the dimensional aspect of their views here. This paper will rather discuss (1) their arguments against Discrete Emotion Theory, and (2) their suggestion that there are no “internal mechanisms” within the organism that trigger emotional episodes, and no reliable “patterns” of biological processes associated with specific emotions (see section 4).
2. I am thinking in particular of empirically minded philosophical accounts of emotion such as those of DeLancey (2002), Griffiths (1997) and Prinz (2004), who discuss Discrete Emotion Theory at length, in particular Ekman’s work.
3. Darwin (1872), it may be useful to point out, never used the term “basic” emotions (let alone, of course, “affect programs”).
4. Ekman (1980) already proposed the notion of a *display rule* to account for cultural and individual variability in emotional responsiveness (the study typically mentioned to illustrate the notion of a display rule is Ekman, 1971). He defined display rules as “conventions, norms and habits that develop regarding the management of emotional responses. A display rule specifies who can show what emotion to whom, and when” (Ekman, 1980, p. 87); individuals learn how to regulate their demeanours through observation and/or education, and their expressions “can become so well learned that they operate automatically and are called forth when the affect program is set off” (*op. cit.*, p. 89).
5. It is worth pointing out that Ekman (2003, p. 66) at some points remarks that “[a]ffect programs are, like the emotion databases, a metaphor, for I do not think there is anything like a computer program sitting in the brain, nor do I mean to imply that only one area of the brain directs emotion.” In a footnote, he adds that “[i]t is more popular today to use connectionist models. I don’t disagree with those formulations, but they are more difficult to understand, and for my purposes here I believe the computer metaphor of a program and instructions is more useful” (*op. cit.*, p. 247). This remark suggests that Ekman may not be committed (anymore?) to the notion of a central information storage. Were he to abandon this latter notion, his view would then fit well with the “dynamical” one proposed below.
6. In her discussion of the analogy between colour and emotion, Barrett (2006b, p. 28) actually claims that linguistic categories “create” experiences.
7. For details, see in particular Russell (2003).
8. Compare De Sousa’s (1987) discussion of *paradigm scenarios*, for example. He argues that, in development, we become familiar with the vocabulary of emotion by association with

situations that are described to us as characteristic of specific emotions. On this view, recognizing emotions is a matter of matching an acquired script with the features of a perceived event. The script is here a criterion to which we appeal in order to make sense of our and others' responses, but we are able to appeal to it in the first place because *there are* features in the episode that we identify as matching the script.

9. Some theorists take the absence of quantitative dynamical systems models of emotion as a reason to reject the whole enterprise altogether—see for example Kaup & Clarke's (2005) commentary to Lewis (2005). Yet this rejection is unwarranted, especially given the complexity of the phenomenon at stake. In addition, the notions of “affect program” and “basic emotions” are themselves speculative—they are theoretical constructs that respond to specific descriptive and explanatory needs. So if the absence of quantitative-mathematical models were a reason to dismiss a theoretical framework, then most models proposed in psychology and neuroscience to describe and explain emotional (and other) phenomena would have to be dismissed as well.

10. This terminology is due to Juarrero (1999, pp. 140-144). See also the discussion in Thompson (2007, pp. 424-425).

11. In an interesting recent discussion, Witherington & Crichton (2007) argue that the dynamical systems approach to emotion describes the “efficient and material causes” of emotional episodes; they also compare the dynamical systems with the “functionalist” approach (which defines emotions in terms of the functions they play in the economy of the organism; see e.g. Frijda, 1986), and argue that the latter specifies the “formal and functional causes” of emotional patterning. My characterization of an emotion as a form or second-order constraint is compatible with the functionalist approach.

12. These appraisal checks evaluate (1) the novelty of a stimulus, (2) its pleasantness, (3) its significance for the individual's goals, (4) the individual's ability to cope with it, and (5) its “norm compatibility.”

References

- Aristotle (1986). *De anima (On the soul)*. London: Penguin Classics.
- Barrett, L.F. (2006a). Are emotions natural kinds? *Perspectives on Psychological Science*, *1*, 28-58.
- Barrett, L.F. (2006b). Solving the emotion paradox: Categorization and the experience of emotion. *Personality and Social Psychology Review*, *10*, 20-46.
- Barrett, L.F. (2006c). Valence is a basic building block of emotional life. *Journal of Research in Personality*, *40*, 35-55.
- Barrett, L.F., & Russell, J.A. (1999). The structure of current affect: Controversies and emerging consensus. *American Psychological Society*, *8*, 10-14.
- Cacioppo, J.T., Berntson, G.G., Larsen, J.T., Poehlmann, K.M., & Ito, T.A. (2000). The psychophysiology of emotion. In M. Lewis & J.M. Haviland-Jones (Eds.), *Handbook of emotions. Second edition* (pp. 173-191). New York: Guildford Press.
- Camras, L.A. (2000). Surprise! Facial expressions can be coordinative motor structures. In M.D. Lewis & I. Granic (Eds.), *Emotion, development and self-organization: Dynamic systems approaches to emotional development* (pp. 100-124). Cambridge, UK: Cambridge University Press.
- Darwin, C. (1872). *The expression of the emotions in man and animals*. London: Harper Collins, 1998.
- De Sousa, R. (1987). *The rationality of emotion*. Cambridge, MA: MIT Press.
- DeLancey, C. (2002), *Passionate engines: What emotions reveal about mind and artificial intelligence*. Oxford: Oxford University Press.
- Descartes, R. *The Passions of the Soul*. In J. Cottingham, R. Stoothoff, & D. Murdoch (Eds.)(1985), *The philosophical writings of Descartes*, vol. 1. Cambridge UK: Cambridge University Press.
- Dewey, J. (1894). *The theory of emotion*. In J. Dewey (1882-1898). *The early works*. Southern Illinois University Press.
- Ekman, P. (1971). Universals and cultural differences in facial expressions of emotion. In *Nebraska Symposium on Motivation*, *4*. Lincoln: University of Nebraska Press.
- Ekman, P. (1980). Biological and cultural contributions to body and facial movement in the expression of emotions. In A.O. Rorty (Ed.), *Explaining emotions* (pp. 73-102). Berkeley: University of California Press.
- Ekman, P. (2003). *Emotions revealed: Understanding faces and feelings*. London: Weidenfeld & Nicholson.

- Ekman, P., & Friesen, W.V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, *17*, 124-129.
- Ekman, P., Levenson, R.W., & Friesen, W.V. (1983). Autonomic nervous system activity distinguishes among emotions. *Science*, *221*, 1208-1210.
- Fingelkurts, A.A., & Fingelkurts, A.A. (2004). Making complexity simpler: Multivariability and metastability in the brain. *International Journal of Neuroscience*, *114*, 843-862.
- Fogel, A., Nwokah, E., Dedo, J.Y., Messinger, D., Dickson, K.L., Matusov, E., & Holt, S. (1992). Social process theory of emotion: A dynamic systems approach. *Social Development*, *1*, 122-142.
- Fogel, A., Hsu, H-C., Shapiro, A.F., Nelson-Goens, G.C., & Secrist, C. (2006). Effects of normal and perturbed social play on the duration and amplitude of different types of infant smiles. *Developmental Psychology*, *42*, 459-473.
- Frijda, N.H. (1986). *The emotions*. Cambridge, UK: Cambridge University Press.
- Goldie, P. (2000). *The emotions: A philosophical exploration*. Oxford: Oxford University Press.
- Griffiths, P.E. (1997). *What emotions really are*. Chicago: University of Chicago Press.
- Izard, C.E., & Malatesta, C.Z. (1987). Perspectives on emotional development I: Differential emotions theory of early emotional development. In J.D. Osofsky (Ed.), *Handbook of infant development. Second edition* (pp. 494–554). New York: Wiley.
- Izard, C.E., Ackermann, B.P., Schoff, K.M, & Fine, S.E. (2000). Self-organization of discrete emotions, emotion patterns, and emotion-cognition relations. In M.D. Lewis & I. Granic (Eds.), *Emotion, development and self-organization: Dynamic systems approaches to emotional development* (pp. 15-36). Cambridge, UK: Cambridge University Press
- James, W. (1884). What is an emotion? *Mind*, *9*, 188-205.
- Juarrero, A. (1999). *Dynamics in action: Intentional behavior as a complex system*. Cambridge, MA: MIT Press.
- Kaup, D.J., & Clarke, T.L. (2005). Where's the example? *Behavioral and Brain Sciences*, *28*, 210.
- Kelso, J.A.S. (1995). *Dynamic patterns: The self-organisation of brain and behavior*. Cambridge, MA: MIT Press.
- Lavelli, M., & Fogel, A. (2005). Developmental changes in the relationship between the infant's attention and emotion during early face-to-face communication: The 2-month transition. *Developmental Psychology*, *41*, 265–280.

- Le Van Quyen, M. (2003). Disentangling the dynamic core: A research program for neurodynamics at the large-scale. *Biological Research*, 36, 67-88.
- Levenson, R.W. (2002). Autonomic specificity and emotion. In R.J. Davidson, K.R. Scherer & H.H. Goldsmith (Eds.), *Handbook of affective sciences* (pp. 212-224). New York: Oxford University Press.
- Lewis, M.D. (2000). Emotional organization at three time scales. In M.D. Lewis & I. Granic (Eds.), *Emotion, development and self-organization: Dynamic systems approaches to emotional development* (pp. 37-69). Cambridge, UK: Cambridge University Press.
- Lewis, M.D. (2005). Bridging emotion theory and neurobiology through dynamic systems modelling. *Behavioral and Brain Sciences*, 28, 169-194.
- Lewis, M.D., & Granic, I. (2000)(Eds.), *Emotion, development and self-organization: Dynamic systems approaches to emotional development*. Cambridge, UK: Cambridge University Press.
- Murphy, F.C, Nimmo-Smith, I., & Lawrence, A.D. (2003). Functional neuroanatomy of emotion: A meta-analysis. *Cognitive, Affective, and Behavioral Neuroscience*, 3, 207-233.
- Panksepp, J. (1998), *Affective neuroscience: The foundations of human and animal emotions*. New York: Oxford University Press.
- Phan, K.L., Wager, T.D., Taylor, S.F., & Liberzon, I. (2002). Functional neuroanatomy of emotion: A meta-analysis of emotion activation studies in PET and fMRI. *Neuroimage*, 16, 331-348.
- Plutchik, R. (2001). The nature of emotions. *American Scientist*, 89, 334-350.
- Prinz, J.J. (2004). *Gut reactions: A perceptual theory of emotion*. Oxford: Oxford University Press.
- Russell, J.A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39, 1161-1178.
- Russell, J.A. (1991). Culture and the categorization of emotion. *Psychological Bulletin*, 110, 426-450.
- Russell, J.A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, 110, 145-172.
- Russell, J.A. (2005). Emotion in human consciousness is built on core affect. *Journal of Consciousness Studies*, 12, 26-42.
- Russell, J.A. (2006). Emotions are not modules. *Canadian Journal of Philosophy, Supplement Volume*, 32, 53-71.

- Russell, J.A., & Barrett, L.F. (1999). Core affect, prototypical emotional episodes, and other things called *emotion*: Dissecting the elephant. *Journal of Personality and Social Psychology*, 76, 805-819.
- Scherer, K.R. (1984). On the nature and function of emotion: A component process approach. In K.R. Scherer & P. Ekman (Eds.), *Approaches to emotion* (pp. 293-317). Hillsdale, NJ: Erlbaum.
- Scherer, K.R. (2000). Emotions as episodes of subsystem synchronization driven by nonlinear appraisal processes. In M.D. Lewis & I. Granic (Eds.), *Emotion, development and self-organization: Dynamic systems approaches to emotional development* (pp. 70-99). Cambridge, UK: Cambridge University Press.
- Scherer, K.R. & Ellgring, H. (2007). Are facial expressions of emotion produced by categorical affect programs or dynamically driven by appraisal? *Emotion*, 7, 113-130.
- Scherer, K.R., Johnstone, T., & Klasmeyer, G. (2003). Vocal expression of emotion. In R.J. Davidson, K.R. Scherer & H.H. Goldsmith (Eds.), *Handbook of affective sciences* (pp. 433-456). New York: Oxford University Press.
- Thompson, E. (2007). *Mind in life: Biology, phenomenology, and the sciences of the mind*. Cambridge, MA: Harvard University Press.
- Thompson, E., & Varela, F.J. (2001). Radical embodiment: Neural dynamics and consciousness. *Trends in Cognitive Science*, 5, 418-425.
- Tomkins, S.S. (1962). *Affect, imagery and consciousness: Vol. 1*. New York: Springer.
- Tomkins, S.S. (1963). *Affect, imagery and consciousness: Vol. 2*. New York: Springer.
- Wheeler, M.W. (2005). *Reconstructing the cognitive world: The next step*. Cambridge, MA: MIT Press.
- Witherington, D.C., & Crichton, J.A. (2007). Frameworks for understanding emotions and their development: Functionalist and dynamic systems approaches. *Emotion*, 7, 628-637.