To cite this chapter:

Moors, A. (2012). Comparison of affect program theories, appraisal theories, and psychological construction theories. In P. Zachar & R. D. Ellis (Eds.), *Categorical versus dimensional models of affect: A seminar on the theories of Panksepp and Russell* (pp. 257-278). John Benjamins.

CHAPTER 13

Comparison of affect program theories, appraisal theories, and psychological construction theories

Agnes Moors
Ghent University

The water between the emotion theories of Russell (this volume) and Panksepp (this volume) seems deep. The aim of this chapter is to structure the theoretical landscape in the hope of gaining a better understanding of how these theories differ from each other as well as from appraisal theories (e.g., Lazarus, 1991; Scherer, 1984). This should help us decide which empirical research we could design to test these theories. Emotion theories differ in their answer to several questions. Some questions are about the definition of emotion. Others are about the cause of emotion. I discuss both types of questions in turn. After that, I examine which differences await an empirical answer and what kind of research is suitable for providing such answers.

1. Definition of emotion

A concept can be used to refer to a set or to an element or instance of that set. An instance can be defined (a) by pointing at it (i.e., ostensive definition) or (b) by detailing the components (i.e., a componential definition). For example, an instance of a car can be defined by pointing at one, or by detailing components such as wheels, engine, tank, and dashboard. A set can be defined (a) by listing all instances or subsets (i.e., extensional definition) or (b) by stating the necessary and sufficient conditions for a thing to be an instance of the set (i.e., intensional definition).

Intensional definitions rarely contain all the necessary and sufficient conditions. Usually, authors focus on the few conditions or criteria that distinguish the set from certain other sets. This kind of intensional definition can be dubbed a differential definition. It is good to keep in mind that the criteria figuring in differential definitions depend on the to-be-compared sets. For example, to distinguish

© 2012. John Benjamins Publishing Company
All rights reserved

the set of cars from the set of bikes, one may put forward the criteria that cars have four wheels and an engine. To distinguish it from the set of flowers, one may put forward the criteria that cars are non-living objects made out of steel.

In extensional definitions in which subsets are listed, the choice of the subsets can be based on various features. Subsets of cars can be formed on the basis of brand, color, size, and horsepower. The feature chosen as the basis for creating subsets depends on the context. Size matters when transporting a big family. Color matters when organizing a traditional wedding procession.

In the remainder of this section, I examine the questions (a) what are the components of instances of emotions (i.e., componential definition), (b) how to distinguish the set of emotions from other sets (i.e., differential definition), and (c) how to create subsets within the set of emotions (i.e., one type of extensional definition). I discuss the answers given by three (families of) emotion theories: affect program theories (e.g., Ekman, 1972; Izard, 1977; Panksepp, 1998, 2000; Tomkins, 1962), appraisal theories (Frijda, 1986; Lazarus, 1991; Roseman, 1984; Scherer, 1984), and Russell's (2003, this volume) psychological construction theory. Panksepp's theory can be considered a special type of affect program theory. I plead the cause of appraisal theories.

Componential definition

An instance of an emotion can be considered in a broad or a narrow sense. The proponents of all three theories agree that an emotion in the broad sense (also termed emotional episode) consists of various components such as (a) a cognitive component, referring to processing or appraisal of a stimulus, (b) a motivational component, referring to action tendencies, (c) a somatic component, referring to central and peripheral physiological activity, (d) a motor component, referring to behavior (e.g., facial and vocal expressions, gross actions), and (e) a subjective component, referring to experience or feelings. Both Russell (this volume) and Panksepp (this volume) use the term of affect to refer to the feeling component. Most theorists conceive of feeling or affect as the reflection of (aspects of) the other components of the emotion into (first or second order) consciousness. Some theorists, however, leave the possibility open that some feelings are unconscious (e.g., Scherer, 2004, 2005a; Prinz, 2004) or do not require "neocortical readout" (Panksepp, 2007). I should point out that I use the term "component" as shorthand for "change in a component".

There is some disagreement about the exact number and nature of the components to include in the emotion. For example, Parrott (2007) omits the component of physiological responses (because all other components have a central

© 2012. John Benjamins Publishing Company
All rights reserved

physiological basis) and adds a component of emotion regulation. Some theorists argue that the cognitive component refers not only to (changes in) appraisal but also to (changes in) the operation of cognitive faculties like attention and memory (cf. Prinz, 2004). Russell (2003) supplements the standard components of appraisal, somatic responses, and motor behavior with the components of Core Affect, categorization, and emotional meta-experience. Core Affect is a combination of the variables of valence and arousal. It has a somatic side (it is defined as neurophysiological activity) and a mental side (the conscious experience of valence and arousal). The categorization of core affect produces what Russell calls emotional meta-experience, which is not the same thing as what other theories call emotional experience. Emotional meta-experience is more a cold thougt than a hot feeling.

An emotion in the narrow sense consists of a subset of the components described above. Again, there is disagreement about the exact number and nature of the components that are part of the narrow definition of emotion. For example, James (1890) identified emotion in the narrow sense with the feeling component whereas Frijda (1986) identified it with the motivational component plus the feeling component. Others equate emotion with the sum of the somatic component, the motor component, and the feeling component (e.g., Izard, 1972). Still others seem to include all the components of the broad definition in the narrow definition of emotion (Clore & Ortony, 2000). In the remainder of this chapter, I use the term emotion in the broad sense because there is more agreement on the composition of emotion in the broad than in the narrow sense.

Differential definition

The three theories discussed here distinguish the set of emotions from other sets in different ways. According to affect program theories, all and only emotions have a type of neural activity called affect program. The affect program of an emotion is situated in the central part of the somatic component and is put forward as the cause of several other components (motivational, peripheral somatic, motor) in the emotion. The problem is that all mental phenomena have concomitant neural activity somewhere in the brain. The question thus arises which neural activity counts as an affect program and which does not. According to Panksepp (this volume), a neural circuit counts as an affect program if direct stimulation of the circuit serves as a reinforcement or punishment in operant conditioning procedures. This already excludes a host of phenomena, but it does not exclude episodes of hunger and pain. It seems that affect program theories also rely on the additional criterion that all and only emotions have

© 2012. John Benjamins Publishing Company
All rights reserved

one of a fixed set of survival-related functions (e.g., preserving safety, seeking new opportunities for goal fulfilment, and caring for offspring; Panksepp, this volume) or deal with a fixed set of species-constant problems (e.g., loss, danger, frustration; Ekman, 1999a). Each function corresponds to a distinct affect program. The problem with this criterion is that the exact composition of the set of functions is somewhat arbitrary.

According to appraisal theories, appraisal is a necessary or (at least) typical condition for something to be an emotion. Appraisal is a process in which a stimulus is evaluated according to several appraisal variables such as goal relevance, goal congruence, coping potential, and agency. For example, a noise in the hall at night can be appraised as goal relevant, goal incongruent, difficult to cope with, and caused by someone else. In addition to being a necessary or typical part of emotion (in the broad sense), appraisal is also seen as the cause of the other components in the emotion. Appraisal theorists do not, however, think that appraisal is a sufficient condition for something to be an emotion. Appraisal of a stimulus only develops into an emotion when the output of this process is "goal relevant". When the output is "goal irrelevant" (or relevant for a goal of minor importance), no emotion ensues (Moors, 2007). In other words, the criterion that distinguishes emotions from other phenomena is not the presence of appraisal, but the output or content of appraisal.

Russell's (this volume, 2003; Russell & Barrett, 1999) psychological construction theory refuses to point to any one component that differentiates the set of emotion from other sets. Emotion is a folk concept with no scientific value. The fact that Russell does talk about the set of emotions (or emotional episodes) indicates that he acknowledges that there is a feature that is shared by the instances of this set. As I understand it, the criterion for inclusion that Russell uses is "those phenomena that people categorize as emotions". Emotions are collections of components grouped together in people's minds on the basis of shared criteria. Discovering these criteria is one of the modest tasks of emotion research. Russell's criterion for inclusion is objective in that it can be determined in an observer-independent manner which phenomena people categorize as emotions. On the other hand, the criterion is subjective in that the instances that end up in the set are dependent on people's categorizations (they are observer dependent).

It may be recalled that Core Affect or even a change in Core Affect is not specific to emotions. Core Affect is necessary for emotions, but not sufficient; it is ever present (Russell & Barrett, 1999). The presence of or change in Core Affect can thus not be considered as a criterion for demarcating the set of emotions from other phenomena.

© 2012. John Benjamins Publishing Company
All rights reserved

Extensional definition

One type of extensional definition is a list of subsets of emotions. Emotion theories have proposed various criteria for dividing the set of emotions into subsets. I discuss the criteria for differentiation proposed by affect program theories, appraisal theories, and constructivist theories. Each of these criteria can be linked to an emotional component.

Affect program theories divide the set of emotions according to the content of the central part of the somatic component. Subsets of emotions are created on the basis of the specific affect program involved. Several affect program theorists (e.g., Ekman, 1992a) also think that differences in affect program translate into differences in peripheral physiological activity (i.e., peripheral part of the somatic component) and differences in facial expressions (i.e., part of the motor component). As a consequence, the specific pattern of physiological activity and the specific facial expression also serve as criteria for differentiation.

The subsets proposed by affect program theories are often referred to as basic emotions. Some affect program theorists posit that basic emotions fill the entire set of emotions (e.g., Ekman, 1994). They present each basic emotion as a family in which there is room for various shades. For example, anger also covers frustration, rage, and fury. Other affect program theorists hold that the set of emotions contains non-basic emotions in addition to basic ones. They depict non-basic emotions as blends or elaborations of basic ones (Damasio, 1999; Prinz, 2000). For example, contempt has been presented as a blend of anger and disgust (Plutchik, 2001). Jealousy has been presented as an elaboration of anger in that it is anger about a particular event (i.e., another person has what you want). Many affect program theorists who distinguish between basic and non-basic emotions hold that non-basic emotions have the same affect program(s) as the basic emotions from which they derive. Some theorists leave room for non-basic emotions that do not have an affect program (cf. Griffiths, 2004).

Some affect program theorists (e.g., Ekman, 1992a) think their subsets collapse with vernacular subsets of emotions like anger, fear, sadness, happiness, surprise, and disgust. Others reject a complete overlap with vernacular subsets and choose stipulative labels. For example, Panksepp (this volume) chooses the labels SEEKING, RAGE, FEAR, LUST, CARE, GRIEF/PANIC, and PLAY to point at the central part of the somatic component in each of the subsets of emotions, but he has no labels for the subsets themselves.

Appraisal theories form subsets on the basis of appraisal variables such as goal relevance, goal congruence, coping potential, and agency. These make up the content of the cognitive component. Appraisal variables can be treated as discrete (e.g., Roseman, 1991) or continuous (e.g., Scherer, 1994). For example, goal

© 2012. John Benjamins Publishing Company
All rights reserved

relevance can be treated as a discrete variable with the values "goal irrelevant" and "goal relevant", or it can be treated as a continuous variable with an infinite number of values ranging from goal irrelevant to goal relevant. When appraisal variables are treated as discrete variables, their values combine to produce a large but finite number of subsets. For example, when the variables of goal relevance, goal congruence, coping potential, and agency are treated as discrete variables with two values each, their values combine to produce 16 subsets of emotions. When appraisal variables are treated as continuous variables,with an infinite number of values for each variable, their values combine to produce an infinite number of subsets of emotions. It may be noted that some appraisal theorists do not allow all possible combinations of appraisal values to occur. This is because they attempt to map appraisal patterns onto the subsets of emotions put forward by affect program theories or onto a set of vernacular subsets. For example, Lazarus (1991) does not have an emotion subset that belongs to the pattern with the values goal relevant, goal congruent, low coping potential, and self-agent, because there is no vernacular emotion subset that corresponds to it.

Psychological construction theory differentiates emotions on the basis of combinations of valence and arousal, two continuous variables that characterize the content of the feeling component and the central part of the somatic component. The values of these variables combine to produce an infinite number of subsets. Constructivist theorists admit that the variables of valence and arousal do not capture all the variation that is present in the set of emotions. Additional variation may come from any of the other components.

To summarize, the criteria for differentiation put forward by the three theories can all be linked to components. The criterion of affect program theories is linked to the central part of the somatic component, that of appraisal theories to the cognitive component, and those of constructivist theories to several components. Some criteria are variables (e.g., valence, arousal, appraisal variables) whereas other criteria are not easily captured in variables (e.g., affect programs). When criteria are variables and lead to a large or infinite number of subsets of emotions, the subsets are most efficiently described by pointing at the variables. It is easier to list a limited number of variables than a large or infinite number of subsets of emotions. When criteria are not variables (e.g., affect programs) and lead to a small number of subsets of emotions, it is equally efficient to list the criteria and the subsets.

Disagreements among theories have to do with the boundaries of the category of emotions and the way in which to internally organize the set of emotions into subsets. Regardless of whether theorists think that the entire set of emotions as well as subsets of emotions can be distinguished with objective or (inter)subjective criteria, they agree that the instances in the set thus obtained show variation.

© 2012. John Benjamins Publishing Company
All rights reserved

There are many types of variation. I focus on three types. Emotions feel good or bad, and active or passive (Type 1 variation). People also have (or report) feelings of the so-called basic type (like anger and fear; Type 2 variation) and the so-called non-basic type (like guilt and jealousy; Type 3 variation). The same three types of variation can be found in the behavior-related components (motivational, peripheral somatic, and motor). Emotions can involve activity or passivity and the preparation and/or execution of approach or avoidance behavior (Type 1 variation). They can also involve preparation and/or execution of behaviors mentioned for so-called basic emotions (e.g., fleeing and fighting; Type 2 variation) and the so-called non-basic ones (e.g., repairing, Type 3 variation).

2. Emotion causation

The previous section examined questions about the definition of emotion such as what are the components of an instance of emotion and how to distinguish and subdivide the set of emotions. The present section examines questions about the cause of the components in the emotion. The three emotion theories agree that (in normal cases) the components are caused by a stimulus (external or internal) that is somehow processed by the organism. In other words, a stimulus is the remote cause whereas the process(es) in the organism is(are) the proximate cause. The theories differ with regard to the kind of causal process(es) that they put forward. Affect program theories propose affect programs. Appraisal theories put forward appraisal. Russell (this volume) argues that there are separate processes for each component and that none of these processes is specific to emotions. I discuss seven differences between the processes of the three theories.

(1) Theories differ with regard to the component in which they situate their causal process(es). Appraisal belongs to the cognitive component whereas affect programs belong to the central part of the somatic component. Appraisal (in appraisal theories) and affect programs (in affect program theories) instigate action tendencies and peripheral physiological responses, which prepare and support motor responses. Aspects of all these components are reflected in the feeling component. Given Russell's (this volume) proposal that there are separate processes for each component, each of the processes belongs to a different component. Russell does not entertain a priori assumptions about causal influences among these components, but he does not exclude them either.

(2) Theories differ with regard to the level of analysis at which they cast their causal process(es). I follow Marr's (1982) proposal that any process can be described at (minimally) three levels of analysis: a functional level, an algorithmic

© 2012. John Benjamins Publishing Company
All rights reserved

level, and an implementational level. At the functional level, a process is described as a relation between an input and an output. At the algorithmic level, a process is described in terms of the mechanisms that translate input into output. Examples of mechanisms are rule-based computation, the activation of learned associations, and the activation of harwired associations. At the implementational level, a process is described in terms of the brain structures or circuits involved. The three levels are related in that every functional process is based on some mechanism and some kind of neural activity. Yet, distinctions at one level of analysis do not map neatly onto distinctions at another level. For example, one functional process can be based on different mechanisms and different neural circuits.

Affect programs are processes described at the implementational level of analysis. They are cast in neural and/or chemical terms. It is important to note that Marr's (1982) three levels of analysis (functional, algorithmic, implementational) do not align with Panksepp's (this volume) three levels of processing (primary, secondary, tertiary). Panksepp's levels can all be situated at Marr's algorithmic level. Roughly, Panksepp's primary mechanism corresponds to the activation of hardwired associations, his secondary mechanism to the activation of learned associations, and his tertiary mechanisms to rule-based computations. Each of these mechanisms has a neural basis to be described at the implementational level. Panksepp has the assumption that primary mechanisms map onto subcortical areas and tertiary mechanisms onto neocortical areas.[1]

Appraisal is a process described at the functional level of analysis. It takes a stimulus as its input and produces values for one or more appraisal variables (e.g., goal relevance, goal congruence, coping potential, and agency) as its output (Moors, 2010; Reisenzein, 2001). A handful of appraisal theorists (Leventhal & Scherer, 1987; Smith & Kirby, 2001) have also addressed the algorithmic level. The mechanisms that they propose are rule-based, associative (i.e., activation of learned associations), and sensory-motor (i.e., activation of hardwired associations).

In Russell's psychological construction theory (this volume), some processes are cast at the implementational level (e.g., the processes producing Core Affect are defined as neurophysiological activity) whereas others are presumably cast at the functional level (e.g., the appraisal process takes the stimulus as input and produces appraisal values as output; the categorization process takes Core Affect and the stimulus as input and produces emotional meta-experience as output).

1. Another assumption is that affect is tied to the primary mechanism and cognition to the tertiary mechanisms. If the tertiary mechanism is at all capable of eliciting emotions, it is only because it is accompanied by activation of the primary mechanism. Tertiary processing is cognitive elaboration on top of primary processing.

© 2012. John Benjamins Publishing Company
All rights reserved

(3) The causal chain between the stimulus and the emotional components can be split into two parts. The first part deals with the stimulus whereas the second part deals with translating the output of the first part into the remaining components of the emotion. Appraisal deals with the stimulus; hence it covers the first part. Affect programs cover the second part (Moors, 2009; see also Barrett, Ochsner, & Gross, 2007). Theories differ with regard to the part in the stimulus-to-components chain that they focus on. Appraisal theories talk about appraisal whereas affect program theories talk about affect programs. Thus, the natural focus of appraisal theories is on the first part whereas the natural focus of affect program theories is on the second part. This being said, some appraisal theorists have also addressed the second part and some affect program theories have also addressed the first part.

Indeed, some appraisal theories have formulated hypotheses about how the output of the appraisal process influences the other components in the emotion. A first hypothesis (e.g., Lazarus, 1991; Figure 1, upper panel) is that all appraisal values are integrated in a pattern and that this pattern determines the emotion subset that is at stake. After this has been determined, a cluster of components that is typical for the emotion subset is elicited. For example, a sudden noise in the hall yields an appraisal pattern typical for "fear". In turn, this pattern generates action tendencies, peripheral physiological responses, behavior, and feelings that are also typical for fear. In this hypothesis, the influence of the appraisal pattern on the other components is driven by a common process. This process may even take the form of an affect program. Hence, appraisal theories that adopt this hypothesis are compatible with affect program theories.

A second hypothesis (Scherer, 2005b; Figure 1, lower panel) is that all appraisal values have a separate influence on the other components in the emotion. Each of these influences is supported by a separate mechanism so that many mechanisms are operating at the same time.[2] For example, the appraisal value "goal incongruent" influences one piece of the person's action tendencies, responses, and feelings, whereas the appraisal value "low coping potential" determines another piece of each of these components. Together, these pieces shape the components in the emotion. Appraisal theories that adopt this hypothesis are incompatible with affect program theories.

In line with appraisal theories adopting the first hypothesis, some affect program theories (e.g., Ekman, 1992a; Matsumoto & Ekman, 2009) raise the possibility that affect programs are preceded by appraisal. Affect programs translate the

2. Some appraisal theorists (Frijda & Parrott, in press) argue that the influence of appraisal on the somatic and motor components is not direct, but indirect in that it goes via the motivational component.

© 2012. John Benjamins Publishing Company
All rights reserved

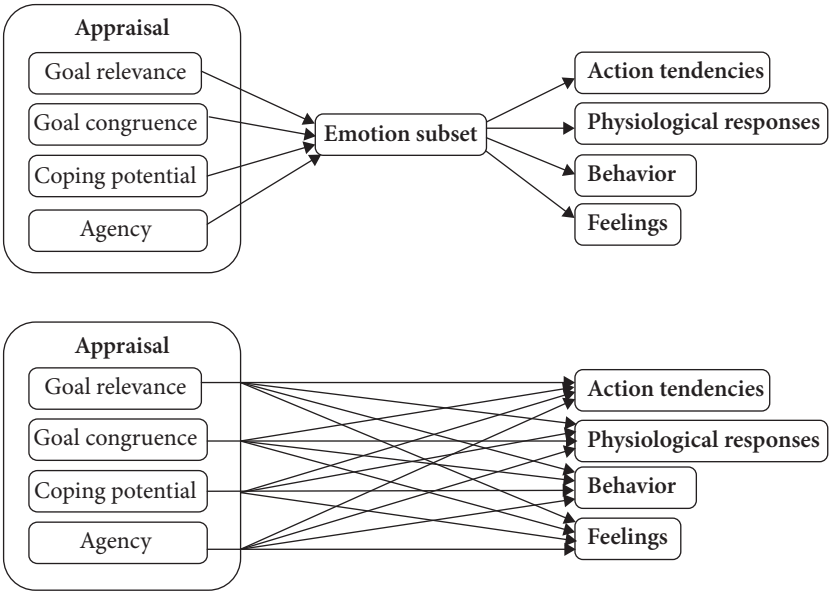

Figure 1. Appraisal theories that are compatible (upper panel) and incompatible (lower panel) with affect program theories

output of appraisal processes into the other components of the emotion. Affect programs are depicted as innate locks that are triggered when the right keys are inserted. The keys are specific appraisal patterns or schemes. In this version of affect program theory, appraisal occupies the first part in the stimulus-to-components chain whereas affect programs occupy the second part.

In Russell's (this volume) psychological construction theory, each component has its own causal chain which should be studied separately. Therefore, this theory does not accommodate easily to the framework in which a first step deals with the stimulus and the second step translates the output of the first step into the other components.

(4) Theories differ with regard to whether they hold that the emotional components are caused by a common process or by separate ones. According to affect program theories, the components in an emotion are caused by a common process, the affect program. There is a limited set of affect programs and each gives rise to its own cluster of values on the other emotional components.

In appraisal theories, appraisal is put forward as the common cause of the remaining components in the emotion. Appraisal may be a single process at the functional level of analysis, but it covers multiple processes (i.e., mechanisms) at the algorithmic level of analysis (rule-based computation and the activation of

© 2012. John Benjamins Publishing Company
All rights reserved

learned and hardwired associations, Leventhal & Scherer, 1987). In addition, appraisal occurs for several variables (e.g., goal relevance, goal congruence, coping potential, agency) and some appraisal theorists have argued that each variable is processed separately (Scherer, 1984; but see Smith & Lazarus, 1993). Furthermore, it is likely that the various mechanisms underlying appraisal rest on different neural processes. In conclusion, appraisal may or may not be a single process at the functional level, but it does not correspond to a single process at the algorithmic level, nor to a single process at the implementational level. This being said, in a given instance of emotion, appraisal may be based on a single process at the algorithmic level (e.g., activation of learned associations) and perhaps a single process at the implementational level.

It may further be recalled that some appraisal theories have not only addressed the first part but also the second part in the stimulus-to-components chain. Appraisal theories that are compatible with affect program theories (e.g., Lazarus, 1991; Figure 1, upper panel) assume that after the appraisal process is completed, the resulting appraisal pattern influences the other emotional components via a common mechanism, like an affect program. Appraisal theories that are incompatible with affect program theories (e.g., Scherer, 2005b; Figure 1, lower panel), on the other hand, assume that each appraisal variable exerts its influence on each of the remaining components via a separate mechanism so that many mechanisms are operating at the same time. In the latter case, the emotional components are not caused by a common process.

Russell (this volume) argues against the existence of a single process (or type of process) that is the common cause of the other emotional components. In his view, each component is caused by its own process, and the study of each component should be handed over to a different branch of psychology. According to Russell, the components are not completely independent but they are more weakly linked than commonly assumed. This view is different from the scenario proposed by appraisal theories that are incompatible with affect program theories (cf. supra; Figure 1, lower panel). In that scenario, each appraisal variable exerts its influence via a separate mechanism, yet each of these mechanisms has a common influence on the other components in the emotion. The appraisal theories that are incompatible with affect program theories parse processes according to appraisal variables whereas Russell parses processes according to components. In other words, Russell proposes component-specific processes.

There is a caveat. It can be questioned whether the links between the components are really that weak. It seems that appraisal does exert a strong influence on action tendencies and responses, and that these do have a strong impact on feelings. Regardless of whether a cluster of these components deserves the label "emotion" or not, the components in the cluster may still entertain strong causal

© 2012. John Benjamins Publishing Company
All rights reserved

relations. I suspect that Russell's suggestion that processes causing components are component specific stems from a failure to distinguish between *causal* processes and *underlying* processes. A causal process precedes a component and it is cast at the same level of analysis as the component, whereas an underlying process occurs simultaneously with the component and is cast at a lower level of analysis. An underlying process simply *is* the component described at a lower level of analysis[3]. To clarify the distinction, consider the example that putting a kettle on the stove causes the water to boil. It is appropriate to say that putting the kettle on the stove causes the water to boil because putting the kettle on the stove occurs temporally prior to the boiling of the water and both events are described at the same level of analysis. The faster movement of the water molecules, on the other hand, is better conceived of as an underlying mechanism than a cause. It is a lower-level description of the boiling of the water. Returning to the case of emotions, I think it makes sense to claim that the processes *underlying* emotional components are component specific (e.g., neural processes underlying behavior may be specific to behavior and neural processes underlying feelings may be specific to feelings) but not that the processes *causing* components are component specific (e.g., if appraisal causes feelings, the process causing feelings is not specific to feelings).

(5) Theories differ with regard to their position regarding the specificity of the proposed causal processes or their outputs for the set of emotions or for some subsets of emotions. Specificity has to do with necessary and sufficient conditions. A is specific for B when A is a necessary and sufficient condition for B.

In affect program theory, the set of affect programs is specific for the set of emotions (Ekman, 1994). Affect programs are necessary and sufficient causes of emotions. In addition, each instance affect program is specific for one subset of emotions. Each instance affect program is necessary and sufficient for one subset of emotions. It is not clear whether affect program theories have the additional assumption that the various affect programs are modular, in the sense that they do not share parts. In any case, the claim that affect programs are specific to subsets of emotions does not imply the claim that they are modular.

In appraisal theory, appraisal is a necessary (or at least typical) but not a sufficient condition for something to count as an emotion (cf. supra, section on definitions). For example, when a noise in the hall at night is appraised as goal irrelevant, no emotion ensues. Thus, appraisal is not specific for emotions; it can also give

3. This is akin to Crisp and Warfield's (2001) distinction between diachronic and synchronic explanations. Both are explanations, but only the former are *causal* explanations. Both Russell and Panksepp tend to treat underlying processes as causal processes, situating all the causal forces at the implementational level of analysis.

© 2012. John Benjamins Publishing Company
All rights reserved

rise to non-emotional phenomena. The mechanisms underlying appraisal at the algorithmic level are not specific for emotions either. Rule-based and associative mechanisms, for example, are standard ingredients of dual mode models in all sorts of domains (cf. Sloman, 1996; Rouder & Ratcliff, 2004). In appraisal theory, it is not appraisal as such that is specific to emotions, but appraisal that produces the output "goal relevant" (cf. supra). In addition, each appraisal pattern is specific for each subset of emotions.

Russell (this volume) not only argues that the processes that cause the emotional components do not have strong causal links (i.e., are component specific, cf. supra), but also that none of these processes is specific for emotions. Feelings that are part of emotions are caused by the same process(es) as feelings that are not part of emotions, and the same goes for categorization, appraisal, physiological responses, and behavior.

(6) Theories differ with regard to the number of causal processes they postulate, and more importantly, the total number of outputs of these processes. The total number of outputs of processes is related to the type of variation (Types 1, 2, and 3) that the processes explain. More outputs yield more variation.

Affect program theories postulate a limited set of affect programs (e.g., Panksepp, this volume, proposes seven) and each has one output, that is, one value for each component[4]. Each cluster of values corresponds to one subset of emotions. In affect program theories, subsets correspond to basic emotions. Thus affect programs account for Type 2 variation. Affect program theories refer to the notions of families, blends, and elaboration based on other processes than affect programs to account for Type 3 variation.

Appraisal theories postulate one appraisal process, but this process has many possible outputs, as many as there are possible appraisal patterns (i.e., combinations of values on appraisal variables). There is variation among individual appraisal theories regarding the number of appraisal variables postulated (cf. Scherer, 2001) and the number of possible values for these variables (a few vs. an infinite number depending on whether the variables are treated as discrete or continuous). As a consequence, appraisal theories vary with regard to the number of appraisal patterns they consider possible. The number of possible appraisal patterns corresponds to the number of possible subsets of emotions. Appraisal theories that postulate a small set of discrete appraisal variables end up with less

4. These outputs can be cast in a broad sense. For example, the output "withdrawal behavior" can refer to all kinds of concrete actions (fleeing, fighting) that have the aim of eventually decreasing contact with the stimulus. So, affect program theories are not committed to one concrete action per emotion subset.

© 2012. John Benjamins Publishing Company
All rights reserved

emotion subsets than appraisal theories that postulate a large set of continuous appraisal variables.

It should be noted that appraisal theories that are compatible with affect program theories tend to limit the set of possible appraisal patterns to those that correspond to basic emotions (cf. supra). In these theories, appraisal accounts for Type 2 variation. Like affect program theories, they have to resort to the notions of families, blends, and elaboration based on other processes than appraisal to account for Type 3 variation. In contrast, appraisal theories that postulate a large/infinite set of appraisal patterns causing a large/infinite set of subsets of emotions naturally account for substantial parts of Type 3 variation in addition to Type 2 variation. But even these theories may have to invoke elaboration for some forms of Type 3 variation (e.g., jealousy). All appraisal theories naturally account for Type 1 variation (valence and arousal). Variation regarding valence comes from the appraisal variables of goal congruence (in all appraisal theories) and intrinsic valence (in some appraisal theories). Variation regarding arousal comes from the appraisal variable of goal relevance (the more relevant a stimulus, the more arousal).

In Russell's theory, the number of possible outputs of the categorization process is limitless. Episodes can be categorized as "emotional," but also as "negative," "anger," or "jealousy." These outputs, however, do not correspond to types of emotional experience (i.e., the feeling component), but to types of emotional meta-experience. The processes producing Core Affect can explain Type 1 variation in the feeling component. The processes producing the behavior-related components are responsible for all three types of variation in these components.

(7) Theories differ with regard to the mechanisms they propose on the algorithmic level of analysis (rule-based computation, activation of learned associations, and activation of hardwired associations) and hence whether the types of variation that they explain are computed on-line, based on learning, or hardwired, Affect programs are said to be hardwired, localized in the subcortical part of the brain, and homologous to the affect programs of other mammals. Thus, Type 2 variation is hardwired. Elaboration responsible for some forms of Type 3 variation is performed by associative and rule-based mechanisms located in the neocortical part of the brain and is based on learning and computation. In addition, Panksepp (2007) suggests that Type 1 variation (at least for valence) is the result of a (neocortical) categorization process based on learning.

As mentioned, appraisal theories that address the algorithmic level propose that all three mechanisms (rule-based, associative, sensory-motor) can underlie appraisal. Leventhal and Scherer (1987) suggested that all appraisal variables can be processed by all three mechanisms. Other appraisal theorists tend to map particular appraisal variables onto particular mechanisms. For example, Reisenzein

© 2012. John Benjamins Publishing Company
All rights reserved

(2009) suggested that appraisal variables like goal relevance and goal congruence rely on hardwired processes. This view resonates with attempts to link appraisal variables to anatomical brain structures (e.g., amygdala as a relevance detector, Sander, Grafman, & Zalla, 2003; e.g., amygdala as a novelty or uncertainty detector, cf. Barrett, 2006a).

Russell (this volume) has not given many details about the mechanisms underlying the components. He does suggest, however, that most components do not stem from the activation of hardwired or even learned associations, but that they are computed or constructed on the spot (and this is what the term "construction" in "psychological construction theory" refers to, Russell, personal communication). Core affect is the only component that can be based on the activation of hardwired associations; hence, Type 1 variation is the only type of variation that can stem from hardwired associations.

3. Empirical research

Now that similarities and differences between constructivist theories, affect program theories, and appraisal theories have been discussed, it is time to consider whether and how the differences can be tested empirically. Similarities and differences pertain to definition and causation.

With regard to definition, the three theories agree that an instance of an emotion consists of various components (componential definition). The theories differ regarding (a) the criteria they propose for demarcating the set of emotions (differential definition), and (b) the criteria they propose for differentiating subsets within the set of emotions (extensional definition). These differences involve an element of choice and therefore cannot be put to an empirical test. This being said, it is recommended that theorists choose criteria that are neither over- nor under-inclusive for the phenomena they have in mind (whether these correspond to the vernacular set of emotions or not) and that they make their criteria explicit. Another remark is that affect program theories and appraisal theories happen to distinguish and internally differentiate the set of emotions on the basis of their purported causal process. Thus, the viability of their definitions depends on empirical evidence for the existence of these processes.

With regard to causation, the three theories agree that (in normal cases) stimuli are somehow processed and that the output of this process is somehow translated into the components of an emotion. The theories differ with regard to (a) the component(s) in which they locate the causal forces, (b) the part in the stimulus-to-components chain that they capitalize on (first vs. second), and (c) the level of analysis at which they describe their causal process(es) (functional vs. algorithmic

© 2012. John Benjamins Publishing Company
All rights reserved

vs. implementational). These are merely differences in focus. Empirical research should concentrate on disagreements among theories. Theories disagree about (a) whether components in an emotion are caused by a common process or by separate ones, (b) whether the causal process(es) or its (their) outputs is (are) specific for the set of emotions or subsets of emotions, (c) the total number of causal processes and/or the total number of outputs of these process(es) – this is related to the type of variation that the process(es) account for –, and (d) the mechanisms they propose at the algorithmic level and and hence the origin of different types of variation. I wish to emphasize that these are separate questions that require separate research programs. Empirical evidence for theories' answers to these questions can be direct or indirect (Ortony & Turner, 1990). Each type of evidence presents its own difficulties.

Direct evidence

In this paragraph, I discuss the kind of direct evidence needed for settling each of the issues of disagreement discussed above. (1) Direct evidence for the claim that the components in an emotion are caused by a common process consists in demonstrating (a) the occurrence of a single process, (b) the occurrence of the remaining components in the emotion, and (c) a causal relation among the process and the components. (2) Direct evidence for the specificity of a causal process for its effect (e.g., an emotion or anger) consists in demonstrating (a) that the process causes the effect, and (b) that the process is necessary and sufficient for the effect. (3) Direct evidence for the existence of a limited number of processes producing one output each (cf. affect program theories) vs. the existence of one process producing a large number of outputs (cf. appraisal theories) consists in (a) identifying the postulated process(es), (b) registering the outputs, and (c) demonstrating a causal relation among process(es) and outputs. (4) Finding direct evidence for the involvement of a mechanism at the algorithmic level is difficult if not impossible because mechanisms cannot be observed directly. Hence, it is difficult to determine what kind of evidence would constitute direct evidence for the hardwired, learned, or computed origin of a process. Some theorists seem to think that a process is hardwired when it can be associated with a particular brain area.

Attempts to obtain direct evidence face several problems. A first problem is that researchers have to agree on what to consider an instance of an emotion. As mentioned, the criterion that affect program theories propose for calling something an emotion is that it should be caused by an affect program. The criterion that appraisal theories propose for calling something an emotion is that it should be caused by an appraisal process that produces the output "goal relevant". By

© 2012. John Benjamins Publishing Company
All rights reserved

delineating the set of emotions on the basis of their cause, claims about the cause of emotions may become circular. For instance, if emotions are defined as those that are caused by an affect program, investigating whether the components in an emotion are caused by a single process comes down to investigating whether components that are caused by an affect program – which is a single process – are caused by a single process. Affect program theories and appraisal theories need to give us additional criteria to determine whether an emotion has occurred.

A second problem is that it is difficult to determine what constitutes a single vs. several processes. It can be argued that, at the functional level, a single process corresponds to a single function. But one function can often be split into several sub-functions. For example, the process of appraisal can be split into sub-processes, each of which has the function of determining the value for a separate appraisal variable. At the algorithmic level of analysis, it is even more difficult to determine what constitutes a single mechanism. Not only can mechanisms be split into sub-mechanisms; they are also unobservable. Hence, each postulated mechanism can be replaced by a combination of alternative mechanisms. Processes described at the implementational level of analysis are difficult to parse as well. When a pattern or network of neural activity is identified, it is difficult to determine whether it constitutes a single process (e.g., an affect program) or several processes. Patterns are composed of parts and it is possible that each of the parts constitutes a separate process. If affect program theorists would be more precise about what kind of neural activity counts as a single neural process, it would be easier to investigate their claims.

A third problem when seeking direct evidence is that causal relations cannot be observed directly. This gap can be bridged to some extent by conducting experimental studies set up to demonstrate dissociations. A process is considered as a cause of an effect when (a) the presence of the process results in the presence of the effect and (b) the absence of the process results in the absence of the effect (i.e., a dissociation). It is not considered as a cause when both the presence and absence of the process result in the presence of the effect (i.e., no dissociation).[5]

A process is considered as a *specific* cause for an effect, when double dissociations are demonstrated. A process is considered as a specific cause for an effect when in addition to the single dissociation described above, presence of the process results in the absence of other effects. To illustrate, the set of affect programs is considered as a specific cause for the set of emotions when (a) activation of

5. It may be noted that in experimental practice, causes are treated as necessary conditions in sufficient sets (this is close to Mackie's, 1975, characterization of a cause as an insufficient but necessary part of a set that is itself unnecessary but sufficient).

© 2012. John Benjamins Publishing Company
All rights reserved

an affect program results in an emotion, but not in non-emotional states, and (b) not activating an affect program results in the non-occurrence of an emotion. Instances of affect programs are considered as specific causes for subsets of emotions when, for instance, (a) activation of the affect program of fear causes the remaining components of fear but not those of other subsets like anger, and (b) activation of the affect program of other subsets like anger causes the remaining components of anger but not those of fear.

The burden of proof is not only on affect program theories. Direct evidence for the claim that appraisal values are specific for parts of the other components in the emotion also requires demonstrating double dissociations. For example, if the appraisal value "goal incongruent" is a specific cause for a furrowed brow, this appraisal value should always result in a furrowed brow but not in other facial action units. In addition, no other appraisal value should result in a furrowed brow[6].

Likewise, direct evidence for Russell's claim that processes are component specific also consists in finding double dissociations. Processes that cause action-related components should not cause the feeling component and processes that cause the feeling component should not cause action-related components.

Indirect evidence

In addition to direct evidence, researchers have adduced indirect evidence for their claims. Indirect evidence is based on inferences that follow from the claims. I give three examples. (1) An inference from the claim that components in the emotion are caused by a common process is that, in the default case, these components should cohere in time and intensity. (2) An inference from the claim that there is a limited set of affect programs leading to a limited set of outputs is that, in the default case, there should be consistency in each of the components across a broad range of stimuli. All stimuli must pass through a limited set of channels. An inference from the claim that the appraisal process produces a large/infinite amount of outputs is that there should be substantial variation in each of the components

6. Only a few appraisal theories that are incompatible with affect program theories (e.g., Roseman, 2001; Scherer, 2009) have fully worked-out hypotheses about relations between appraisal values and parts of the other components and they do not all claim that appraisal outputs are specific for parts of the other components. Another remark is that specificity on the level of parts is not in itself incompatible with specificity on the level of wholes. For theories that assume that the emotions in each subset have a unique part, evidence for the specificity of a process for that part also counts as evidence for the specificity of the process for the subset. For example, if Panksepp's affect programs are specific for certain actions (i.e., parts), this is not a problem as long as it is assumed that each subset of emotions is characterized by a specific action.

© 2012. John Benjamins Publishing Company
All rights reserved

(3) An inference from the claim that affect programs are hardwired is that these should lead to universal signatures in expressive behavior (cf. Ekman, 1992b).

Attempts to obtain indirect evidence faces several problems (for comprehensive reviews, see Barrett, 2006a, 2006b; Ekman, 1999a, 1999b; Ortony & Turner, 1990; Russell, 1994). A first problem is to know how much coherence, consistency/variation, and universality counts as evidence for coherence, consistency/variation, and universality. A second problem is that the default case rarely obtains in real life. The process causing emotional components always interacts with other processes such as regulation processes. This may dampen coherence, consistency, and universality. A possible solution is to keep the need for regulation low by removing external observers from the scene or to manipulate the need for regulation. A third, more serious problem is that indirect evidence falls prey to the fallacy of affirming the consequent. Coherence need not result from a common process; it may also result from the stimulus (independent of appraisal or an affect program). This problem can be circumvented by direct activation of the process in the brain (cf. Panksepp, this volume). Consistency need not result from a limited set of processes or outputs; it may also result from a limited range of stimuli that is generally encountered. Universality need not result from the hardwired nature of the causal process; it may also result from species-constant learning (Ekman, 1999a).

4. Conclusion

The aim of the present chapter was to analyse the complex debate between proponents of affect program theories, appraisal theories, and constructivist theories of emotion. Starting from common ground among the theories (i.e., the premises that emotions in the broad sense are collections of components, that components are caused by process/es, and that different types of variation exist), this analysis revealed several axes on which the theories differ. By disentangling the points of agreement and disagreement, it becomes possible to identify the empirical questions. Each of these questions requires a separate research program and presents its own set of difficulties.

Acknowledgement

Preparation of this chapter was supported by Methusalem Grant BOF09/01M00209 of Ghent University.

© 2012. John Benjamins Publishing Company
All rights reserved

References

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., Ochsner, K. N., & Gross, J. J. (2007). On the automaticity of emotion. Chapter in J. Bargh (Ed.), *Social psychology and the unconscious: The automaticity of higher mental processes.* New York: Psychology Press.

Clore, G. L., & Ortony, A. (2000). Cognition in emotion: Always, sometimes, or never? In R. D. Lane & L. Nadel (Eds.), *Cognitive neuroscience of emotion* (pp. 24–61). New York: Oxford University Press.

Crisp, T. M., & Warfield, T. A. (2001). Kim's master argument. *Noûs, 35* (2), 304–316.

Damasio, A. R. (1999). *The feeling of what happens: Body and emotion in the making of consciousness*. New York: Harcourt Brace.

Ekman, P. (1972). Universals and cultural differences in facial expressions of emotion. In J. Cole (Ed.), *Nebraska Symposium on Motivation 1971*, (Vol. 19, pp. 207–283). Lincoln, NE: University of Nebraska Press.

Ekman, P. (1992a). An argument for basic emotions. *Cognition and Emotion, 6*, 169–200.

Ekman, P. (1992b). Are there basic emotions? *Psychological Review*, *99*, 550–553.

Ekman, P. (1994). All emotions are basic. In P. Ekman, & R. J. Davidson (Eds.), *The nature of emotion: Fundamental questions* (pp. 15–19). Oxford, UK: Oxford University Press.

Ekman, P. (1999a). Basic emotions. In T. Dalgleish and T. Power (Eds.), *The handbook of cognition and emotion* (pp. 45–60). Sussex, U.K.: Wiley.

Ekman, P. (1999b). Facial expressions. In T. Dalgleish and T. Power (Eds.), *The handbook of cognition and emotion* (pp. 301–320). Sussex, U.K.: Wiley.

Frijda, N. H. (1986). *The emotions*. New York: Cambridge University Press.

Frijda, N. H., & Parrott, W. G. (2011). Basic emotions or ur-emotions? *Emotion Review, 3*, x–x.

Griffiths, P. E. (2004). Is emotion a natural kind? In R. C. Solomon (Ed.), *Philosophers on emotion* (pp. 233–249). Oxford and New York: Oxford University Press.

Izard, C. E. (1972). *Patterns of emotions*. New York: Academic Press.

Izard, C. E. (1977). *Human emotions*. New York: Plenum Press.

Lazarus, R. S. (1991). *Emotion and adaptation*. New York: Oxford University Press.

Leventhal, H., & Scherer, K. R. (1987). The relationship of emotion to cognition: A functional approach to a semantic controversy. *Cognition and Emotion, 1*, 3–28.

Mackie, J. L. (1975). Causes and conditionals. In E. Sosa (Ed.), *Causation and conditionals.* Oxford, UK: Oxford University Press.

Marr, D. (Ed.). (1982). *Vision: A computational investigation into the human representation and processing of visual information*. New York: Freeman.

Matsumoto, D., & Ekman, P. (2009). Basic emotions. In D. Sander & K. R. Scherer (Eds.), *Oxford companion to affective sciences* (pp. 69–72). Oxford, UK: Oxford University Press.

Moors, A. (2007). Can cognitive methods be used to study the unique aspect of emotion: An appraisal theorist's answer. *Cognition and Emotion, 21*, 1238–1269.

Moors, A. (2009). Theories of emotion causation: A review. *Cognition and Emotion, 23,* 625–662.

Moors, A. (2010). Automatic constructive appraisal as a candidate cause of emotion. *Emotion Review, 2*, 139–156.

© 2012. John Benjamins Publishing Company
All rights reserved

Ortony, A., & Turner, T. J. (1990). What's basic about basic emotions? *Psychological Review, 97*, 315–331.

Panksepp, J. (1998). *Affective neuroscience: The foundations of human and animal emotions.* New York: Oxford University Press.

Panksepp, J. (2000). Emotions as natural kinds within the mammalian brain. In M. Lewis & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (2nd ed., pp. 137–156). New York: Guilford.

Panksepp, J. (2007). Neurologizing the psychology of affects: How appraisal-based constructivism and basic emotion theory can coexist. *Perspectives on Psychological Science, 2*, 281–296.

Parrott, W. G. (2007). Components and the definition of emotion. *Social Science Information, 46*, 419–423.

Plutchik, R. (2001). The nature of emotions. *American Scientist, 89*, 344–350.

Prinz, J. (2000). Getting to the heart of emotions. New York: paper delivered at American Philosophical Association.

Prinz, J. (2004). *Gut reactions: A perceptual theory of emotions*. Oxford, UK: Oxford University Press.

Reisenzein, R. (2001). Appraisal processes conceptualized from a schema-theoretic perspective. Contributions to a process analysis of emotions. In K. R. Scherer, A. Schorr & T. Johnstone (Eds.), *Appraisal processes in emotion* (pp. 187–201). New York: Oxford University Press.

Reisenzein, R. (2009). Emotions as metarepresentational states of mind: Naturalizing the belief-desire theory of emotion. *Cognitive Systems Research, 10,* 6–20.

Rouder, J. N., & Ratcliff, R. (2004). Comparing categorization models. *Journal of Experimental Psychology: General, 133*, 63–82.

Roseman, I. J. (1984). Cognitive determinants of emotion: A structural theory. *Review of Personality & Social Psychology, 5,* 11–36.

Roseman, I. J. (1991). Appraisal determinants of discrete emotions. *Cognition and Emotion, 5*, 161–200.

Roseman, I. J. (2001). A model of appraisal in the emotion system: Integrating theory, research, and applications. In K. R. Scherer, A. Schorr, & T. Johnstone (Eds.), *Appraisal processes in emotion: Theory, methods, research* (pp. 68–91). New York: Oxford University Press.

Russell, J. A. (1994). Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies. *Psychological Bulletin, 115*, 102–141.

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review, 110,* 145–172.

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.

Sander, Grafman, & Zalla (2003). The human amygdala: An evolved system for relevance detection. *Reviews in the Neurosciences, 14*, 303–316.

Scherer, K. R. (1984). On the nature and function of emotions: A component process approach. In K. R. Scherer & P. Ekman (Eds.), *Approaches to emotion* (pp. 293–317). Hillsdale, NJ: Erlbaum.

Scherer, K. R. (1994). Toward a concept of "modal emotions". In P. Ekman & R. J. Davidson (Eds.), *The nature of emotion: Fundamental questions* (pp. 25–31). Oxford, UK: Oxford University Press.

© 2012. John Benjamins Publishing Company
All rights reserved

Scherer, K. R. (2001). Appraisal considered as a process of multilevel sequential checking. In K. R. Scherer, A. Schorr, & T. Johnstone (Eds.), *Appraisal processes in emotion* (pp. 92–120). New York: Oxford University Press.

Scherer, K. R. (2004). Feelings integrate the central representation of appraisal-driven response organization in emotion. In A. S. R. Manstead, N. H. Frijda, & A. H. Fischer (Eds.). *Feelings and Emotions: The Amsterdam Symposium* (pp. 136–157). Cambridge, UK: Cambridge University Press.

Scherer, K. R. (2005a). Unconscious processes in emotion: the bulk of the iceberg. In P. Niedenthal, L. Feldman-Barrett & P. Winkielman (Eds.), *The unconscious in emotion* (pp. 312–334). New York: Guilford.

Scherer, K. R. (2005b). What are emotions? And how can they be measured? *Social Science Information, 44*, 695–729.

Scherer, K. R. (2009). Emotions are emergent processes: They require a dynamic computational architecture. *Philosophical Transactions of the Royal Society: B, 364*, 3459–3474.

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin, 119*, 3–22.

Smith, C. A., & Lazarus, R. S. (1993). Appraisal components, core relational themes, and the emotions. *Cognition and Emotion, 7*, 233–269.

Smith, C. A., & Kirby, L. D. (2001). Toward delivering on the promise of appraisal theory. In K. R. Scherer, A. Schorr & T. Johnstone (Eds.), *Appraisal processes in emotion* (pp. 121–138). New York: Oxford University Press.

Tomkins, S. S. (1962). *Affect, imagery, consciousness: Vol. 1: The positive affects*. New York: Springer.

© 2012. John Benjamins Publishing Company
All rights reserved