To cite this chapter:

Moors, A. (2024). An overview of contemporary theories of emotions in psychology. In A. Scarantino (Ed.), *Emotion theory: The Routledge comprehensive guide* (pp. 232-259). Routledge Academic. https://doi.org/10.4324/9781315559940-15

An Overview of Theories of Emotions in Psychology

Agnes Moors[1]

Abstract

The chapter reviews five major families of emotion theories in psychology: (a) evolutionary theories, (b) network theories, (c) evaluation theories, which comprise appraisal theories (which qualify as stimulus evaluation theories) and the goal-directed theory (which qualifies as a response evaluation theory), (d) psychological constructionist theories, and (e) social theories. Theories are compared in terms of their working definitions, their constitutive explanations, their causal-mechanistic explanations, whether they think emotions have scientific status, and the way in which they partition the set of emotions.

The aim of this chapter is to provide an overview of contemporary emotion theories in psychology and to shed light on some of the complex debates between these theories by situating them within a comparative analysis. I discuss five families of emotion theories that are influential in contemporary psychology: (a) evolutionary theories, also called (biologically-based) basic emotion theories (Ekman 1999; Ekman and Cordaro 2011; Shiota this volume; Tracy and Randles 2011; Keltner *et al.* 2016), (b) network theories (Bower 1981; Lang 1994; Leventhal 1984; Lewis 2005), (c) evaluation theories, which can be split into stimulus evaluation theories, known as appraisal theories (Ellsworth 2013, this volume; Frijda 1986; Lazarus 1991; Roseman 2013; Scherer 2009), and the less well-known response evaluation theories, in particular the goal-directed theory (Moors 2017a), (d) psychological constructionist theories (Schachter 1964; Barrett 2006, 2012, 2017; Russell 2003, 2009), and (e) social theories (Averill 2012; Mesquita and Parkinson this volume). For an in-depth comparison, the reader is referred to Moors (2022).

The approach to organizing these theories is the following: I start by outlining an idealized path towards theory development that emotion theories follow. During this path, questions arise that different theories have provided different answers to. The questions can be identified as axes and the answers of the theories as the positions that theories occupy on these axes. Combining these axes creates a multidimensional space in which theories can be placed and compared. This approach has the potential to reveal substantive disagreements, which are separate from mere terminological differences. The particular framework offered also reveals disagreements that would otherwise remain unnoticed, and potentially opening the door to theoretical integration once disagreements along a specific axis are resolved.

Theory Development

The development of emotion theories can be described as following a path comprised of the following four stages: In the first step, emotion, which is the explanandum or phenomenon to

be explained, is provisionally demarcated from other phenomena by providing a working definition. This working definition can take the format of a list of superficial properties that instances of the set of emotions are supposed to have or the format of a number of prototypical instances belonging to the set. Working definitions are descriptive definitions in that they describe the way in which laypeople (or scientists as laypeople) understand emotions.

In the second step, an explanation is proposed, in which emotion—the explanandum—is linked to an explaining fact—an explanans. The types of explanations provided by emotion theories are (a) constitutive explanations, which specify the components that are part of emotions, and (b) causal-mechanistic explanations, which specify the entities that cause emotions and/or the mechanisms that mediate between these causes and the emotions. Mechanisms can be cast at the mental level of analysis (i.e., mental mechanisms) or at the neural level of analysis (i.e., neural mechanisms). All theories discussed here provide a mental mechanistic explanation; some theories also work towards providing neural mechanistic explanations, that is, they try to find out how their mental mechanism can be implemented in the brain.

The third step consists of the testing of the explanations in empirical research. If the explanations are sufficiently confirmed, a fourth step can take place in which a scientific definition of emotion is articulated. A scientific definition is a prescriptive definition in that it prescribes how a phenomenon should be understood in scientific discourse. Importantly, the scientific definition derives from the proposed explanans. The ideal is to come up with an explanans that is shared by all instances of the set and not by instances of other sets. In other words, the ideal is that the explanation suggests necessary and sufficient conditions to determine membership to the set from now on. Once a scientific definition of emotions is in place, it allows for the development and testing of new predictions about emotions, such as

how emotions influence other phenomena and how they go astray in different psychopathogies and so on.

Note that in the four-stage path described here, the formulation of a scientific definition only arises at the end of the path, not at the start. To make this idea insightful, consider the same four-stage path if the explanandum is water. In a first stage, water can provisionally be demarcated as transparent, odorless, and tasteless fluid, covering 70 percent of the planet's surface. In a second stage, a constitutive explanation of water, which tries to discover the molecular structure of water, links it to H_2O. After it is verified in the third stage that water indeed has the molecular structure of H_2O, in the fourth stage, water is no longer descriptively defined as a transparent, odorless, and tasteless fluid, but it is prescriptively defined in terms of the proposed explanans: H_2O.

Once a candidate scientific definition is proposed, the adequacy of this definition can be evaluated with criteria such as similarity and fruitfulness (Carnap 1950). The similarity criterion states that the set carved out by the working definition—henceforth, the folk set—and the set carved out by the scientific definition—henceforth, the scientific set—must be similar enough (Green 1992; Scarantino 2012). The fruitfulness criterion states that the set individuated by the prescriptive definition should allow for scientific extrapolation, that is, the generalization of discoveries of one instance to the other instances in the set (Griffiths 2004; Scarantino 2012). This, in turn, requires the set to be homogeneous in a non-superficial way. Instances must share deep features such as a common constitution or a common causal mechanism.

If the folk set is heterogeneous, a trade-off between similarity and fruitfulness is inevitable. This means that maximal similarity comes at a cost to fruitfulness and maximal fruitfulness comes at a cost to similarity. To illustrate, the folk set "jade" is heterogeneous because it consists of two minerals: "jadeite" and "nefrite". Maximizing similarity in this case

would come down to keeping both minerals in the set, but since discoveries of members constituted by jadeite may not generalize to members constituted by nefrite, fruitfulness would be sacrificed. Maximizing fruitfulness in this case, on the other hand, would amount to replacing the folk set "jade" with only a subset of this set, for instance, a subsets that only contains the members constituted by jadeite. This subset may be fruitful but loses similarity with the original folk set. Unfortunately, there are no guidelines for how to weigh these two criteria, so theorists are left with a choice (Swartz 1997).

Instead of maximizing one criterion at the complete cost of the other, more moderate forms of prioritizing one criterion over the other exist. One moderate form of prioritizing *similarity over fruitfulnes*s is to give up the ambition to find a classic definition, which specifies a single condition that is both necessary and sufficient or even a set of necessary conditions that are together sufficient. The alternative is to propose a cluster-type definition in which the conditions are relaxed from being necessary to being typical. This is often illustrated with the folk set of lemons, for which the properties yellow, oval, and acid are typical rather than necessary (some lemons are green, some are round, and some are bitter). Another option is to consider moderate forms of prioritizing *fruitfulness over similarity*. This can be done by trimming the folk set and weed out a small number of exemplars to make it more homogeneous. This is often illustrated with the folk set of fish, from which whales and dolphins were trimmed to arrive at the scientific definition of fish.

The most radical way to deal with the heterogeneity of a set, however, is to eliminate the set altogether (so that nothing from the original set is kept) and to replace it instead with other, more fruitful sets. This can be illustrated with the folk set of air, which was taken by Aristotle to be one of four fundamental elements jointly with water, fire, and earth. Air could be descriptively defined as a transparent, odorless gas that fills the sky and our lungs. However, when scientists discovered that air could not be captured by a unified constitutive

explanation because air is composed of diverse molecules such as oxygen, nitrogen, and carbonite, they concluded that air was not an adequate scientific set and abandoned it. In turn, all the component molecules in air were taken as individual explananda, initiating new scientific paths focused on oxygen, nitrogen, carbonite, and other component molecules.

Emotion theories not only seek to *demarcate* the set of emotions as a whole from other sets, but also to *partition* the set internally. In particular, they want to know whether the prototypical emotions sampled in the working definitions that list such emotions reflect a suitable way to partition this set or whether there are better ways.

Now that the skeleton for the development of emotion theories is in place, it is time to put some flesh on the bone. I will zero in on each of the four stages of theory development. I will identify in these stages the major axes on which theories can differ or align and I will specify where each theory can be situated on these axes. To anticipate on the results of this exercise, five major axes can be identified across the four stages. In Stage 1 (working definition), theories can differ in terms of the working definitions they propose (**Axis 1**). Despite minor differences, however, theories show fair consensus at this stage. This means that there is fair consensus about what the phenomena in need of explanation are. In Stage 2 (explanations), theories differ in terms of the constitutive (**Axis 2**) as well as the causal-mechanistic explanations (**Axis 3**) they propose. Differences in terms of constitutive explanations cut across the families of emotion theories listed above. Differences in terms of causal-mechanistic explanations, on the other hand, form the heart of and determine the very identity of psychological emotion theories. The latter differences transpire in the empirical research programs undertaken by proponents of these theories in Stage 3, as well as in the scientific definitions that they propose in Stage 4. In the latter stage, theories differ in whether the explanations that they proposed allow turning the folk set of emotions into a scientific set or not (**Axis 4**). *Vindicator* theories conclude that this is the case; *skeptical* or *eliminativist*

theories conclude that this is not the case. Thus, vindicator theories believe that emotions are like water, lemons, fish, or jade, whereas skeptical theories believe they are like air. Finally, the explanations of theories may not only serve to demarcate the set of emotions from other sets. They may also provide principles for the partitioning of the set of emotions. Theories with a *discrete* view hold that the set of emotions is best organized according to prototypical emotion subsets (e.g., fear, anger, sadness, joy); theories with a *dimensional* view hold that the variety within the set of emotions is best captured by identifying dimensions and placing emotions in a multidimensional space (**Axis 5**). I will now turn to each of the four stages and discuss differences and similarities between the theories listed at the start of this chapter.

Working Definition

To provisionally demarcate the set of emotions, theories have presented descriptive definitions in the form of a list of typical and apparent properties commonly found among emotions and/or they have relied on samples of prototypical emotion types. Prototypical emotion types are sampled from natural language. Examples included in many lists are fear, anger, sadness, joy, regret, guilt, shame, disgust, surprise, and so forth. Although theories differ in the properties that they eventually include as necessary and real in their scientific definitions of emotions, they show fair (although not perfect) consensus about a number of properties that emotions typically have or at least appear to have (**Axis 1**). These properties are that emotions (a) reside in individuals, (b) show continuity across human adults, on the one hand, and infants and non-human animals, on the other hand (even if they are richer and more complex in the former group), (c) are elicited by important events (physical stimuli external or internal to the body, and mental representations, such as memories and imaginings), (d) have a mental aspect, in particular, that they are about something outside of themselves (i.e., object-directedness, e.g., sadness about losing one's cat, angry about an injust remark, joy about winning a contest), (e) have a bodily aspect in that they are

accompanied by physiological responses, subtle behavior (e.g., facial expressions), coarse behavior (i.e., musculoskeletal behavior), or at least the tendency to engage in behavior (i.e., an action tendency), (f) are characterized by heat, which has a quantitivative part (i.e., intense) and a qualitative part (i.e., positive or negative valence), (g) have control precedence (i.e., they claim priority over mundane goal pursuit), and (h) tend to have an irrational flavor (i.e., they have the capacity to cloud our thinking and make us behave counter to our best interests).

Constitutive Explanation

A constitutive explanation of emotion specifies the parts or components of an emotion. Often-cited components are (a) a cognitive component, with changes in information processing (e.g., appraisal of a stimulus as a threat; categorization of bodily feelings as fear), (b) a motivational component, with changes in action tendencies (e.g., the tendency to fight or flee), (c) a somatic component, with changes in physiological responses (e.g., an adrenaline rush), (d) a motor component, with changes in overt behavior (e.g., a startled facial expression and actual flight behavior), and (e) a subjective component, with changes in experience or feelings (e.g., feelings of fear). The somatic and motor components are physical entities and are therefore directly observable. The cognitive, motivational, and subjective components are mental entities, which are not observable and need to be studied indirectly via observable entities. The somatic component refers to physiological responses. Although these can in principle include both central (i.e., brain) and peripheral activity, some theorists exclude the central part because all mental components are also sustained by neural activity on a lower level of analyis (see Parrott 2007). Information processes in the cognitive component can be spelled out on a high level of analysis as the relation between a certain input and a certain output. For instance, stimulus evaluation or appraisal takes a stimulus and possibly others sources of information as its input and produces a representation with a stimulus evaluation (e.g., dangerous) as its output. An action tendency is a representation of an action that a person

wants to engage in. Feelings are the conscious experience of the contents of certain representations. Feelings can be divided into first-order experience or raw feelings and second-order experience or labeled feelings. Raw feelings occur when representations of one or more of the other components become conscious. Labeled feelings occur when a person becomes conscious of the fact that they have an emotion or a specific emotion (e.g., fear, anger, sadness).

Theories differ with regard to the number and nature of the components that they include in their constitutive explanations of emotions (**Axis 2**). These explanations range from narrow to broad, but as mentioned, this axis cuts across the five families of emotions listed at the start of the chapter. On the extreme narrow end, emotion is equated with a single component. In older theories, for instance, emotions have been equated with feelings (e.g., Schachter 1964). On the extreme broad end, emotions are conceived of as episodes that contain the entire set of components listed above (e.g., Clore and Ortony 2000).

Causal-mechanistic Explanations

I mentioned earlier how explanations in psychology can be constitutive—they specify the components of a phenomenon—or causal-mechanistic—they identify the factors and/or the mechanisms that cause the phenomenon. Theories with a narrow constitutive explanation often treat other entities as causes or consequences of emotions. Theories with a broad constitutive explanation, which include all or most of the listed components inside the emotion, do not always propose causal mechanisms that are separate from and precede the emotion, but often propose causal mechanisms that sit inside the emotion and cause the remaining components inside the emotion. For instance, in several appraisal theories (to be discussed shortly), the cognitive component is part of the emotion but is itself considered as the cause of the motivational, somatic, motor, and feeling components (Moors 2013). As mentioned, the causal-mechanistic explanations represent the key axis on which families of

psychological emotion theories can be differentiated (**Axis 3**). Within this axis, at least three sub-axes can be specified. A first sub-axis (**Axis 3a**) splits theories into personal theories, which focus on mechanisms residing in the individual, and social theories, which tailor these personal mechanisms to the social context and/or consider mechanisms residing at the social level. Personal theories include evolutionary theories, network theories, evaluation theories, and psychological constructionist theories. Social theories can be considered as the social versions of these personal theories. A second sub-axis (**Axis 3b**) further organizes the group of personal theories on the basis of the component that they primarily want to explain. Evolutionary theories, network theories, and evaluation theories are more action-oriented in that they work mainly towards explaining the motor component. Psychological constructionist theories are more feeling-oriented in that they seem primarily interested in explaining the feeling component. This distinction is not a hard one, however. Action-oriented theories have also been concerned with feelings and feeling-oriented theories also have devoted attention to behavior. A third sub-axis (**Axis 3c**) distinguishes theories based on the specific causal mechanisms that they propose. I now turn to a detailed discussion of the causal mechanisms proposed by each of the five emotion-theoretical families, grouped into action-oriented personal theories (evolutionary, network, evaluation theories), feeling-oriented personal theories (psychological constructionist theories), and social theories.

Action-Oriented Theories

To be able to compare the causal mechanisms proposed by action-oriented theories, they must be fitted into a common mold. At a very abstract level, action-oriented theories take the connection between the cognitive and motivational components as the backbone of emotion and they consider the other components as flowing from these. The cognitive component takes care of the extraction of information (i.e., *extraction phase*). After this, the information is somehow translated into the motivational component (i.e., *translation phase*). The

motivational component, in turn, dictates the behavior that the system must undertake. The somatic component recruits the physiological responses to support this behavior and the motor component simply *is* the overt behavior. The feeling component, finally, feeds on one or more of the other components. In the next sections, I will zero in on the way in which each of the action-oriented theories fills in the details for the extraction phase and the translation phase.

Evolutionary theories

Evolutionary theories propose that a limited set of emotions, called basic emotions, are caused (and/or partly constituted) by dedicated mechanisms called affect programs. Affect programs have been installed by evolution. The idea is that our ancestors were exposed to recurrent life tasks (i.e., challenges and opportunities present in the environment) and that the best behavioral solutions for dealing with these life tasks became inherited and anchored in the brain (e.g., Ekman 1999). Each basic emotion corresponds to a specific affect program that suggests a specific response when a particular life task presents itself. For instance, the affect program for fear activates the tendency to flee or avoid in response to threat, that of anger activates the tendency to fight in response to obstacles or competition, that of disgust activates the tendency to eject stimuli that pose a risk for contamination, that of sadness activates the tendency to let go of unfruitful goal striving, that of surprise activates the tendency to prepare for stimuli with potential implications for goals, and that of joy activates the tendency to indulge in successful goal pursuit or to move on to the pursuit of other goals. On the mental level, affect programs can be understood as innate links between a stimulus representation (e.g., snake, danger) and a response representation or action tendency (e.g., to flee, to defend onself), short stimulus-response (S–R) links. Activation of such an S–R link is called a stimulus-driven process (Heyes and Dickinson 1990). In some versions of evolutionary theory, these S–R links are cast on a concrete level, as in links of the type “snake-flee”. In

other versions, they are cast on an abstract level, as in links of the type “danger-defense”. Abstract S–R links allow for more flexibility because different types of stimuli come under the heading of danger and the tendency to defend oneself can be manifested in different concrete behaviors (e.g., not only fleeing, but also fighting and freezing).

On the neural level, affect programs are implemented in dedicated neural mechanisms. These can be specific neural circuits, made up of a set of interconnected brain regions. These can also be specific patterns of activation, which can in principle be distributed across the entire brain (see Shiota this volume). Traditional affect program theory lists six basic emotions: fear, anger, joy, sadness, disgust, and surprise (Ekman 1999); more modern variants progressively expand this list, including for instance pride, shame, contempt, awe, and so on (Keltner *et al.* 2016).

Affect programs in evolutionary theories take care of translation but not of extraction. An account of extraction should specify which stimulus features activate the affect program in the first place and how these features are determined. In evolutionary theories with concrete S–R links, a simple perception process yielding concrete stimulus representations (e.g., wild animals, noise blasts, sudden loss of control) is sufficient to trigger the affect programs. In evolutionary theories with abstract S–R links, an appraisal process is required to provide abstract stimulus representations (e.g., danger, offense, loss) that can activate the affect programs. The latter evolutionary theories can be considered to be hybrids with appraisal theories (to be discussed shortly).

According to the default scenario of evolutionary theories, once the affect program (i.e., the connection between the cognitive and motivational component) of some emotion is activated, it automatically triggers the output components—the somatic, motor, and feeling components—of that emotion. The affect program is considered to be the common cause of the output components, an important implication of which is that the output components are

supposed to form concordant packages. Zooming in on the feeling component, evolutionary theories come in a version in which feelings are produced centrally, by the affect program itself (e.g., Izard 2011), and a version in which feelings are generated peripherally (e.g., Levenson 2011). In the latter version, the somatic and motor responses produce resident effects in the person's body, which can be consciously perceived via interoception or proprioception, and hence are felt.

Evolutionary theories assume that innate affect programs are supplemented with learning processes and computation. On the input side, these supplementary processes can elaborate the number of stimuli that can trigger an affect program. On the output side, they are involved in regulation and planning. Both emotion regulation and planning are supposed to involve goal-directed processes. If an initially generated, emotional action tendency conflicts with a person's long-term goals, the person may try to prevent its manifestation. For instance, if an aggressive tendency is expected to lead to retaliation, the person may try to suppress or conceal it. Even if an emotion action tendency is not regulated, goal-directed processes may still be required to translate a more or less abstract action tendency into a more concrete action tendency or a motor program. If nothing is standing in the way of manifesting one's aggressive tendency into behavior, for instance, the person still needs to decide whether to engage in physical or verbal aggression.

Network theories

Network theories of emotion (e.g., Bower 1981; Lang 1994; Leventhal 1984) originally grew out of associative models of classical conditioning in the learning literature and semantic network models in the memory literature. More recent sources of inspiration are dynamic systems models (e.g., Lewis 2005) and psychometric network models (see Lange et al. 2020). Network theories propose that when an emotion occurs, information about the eliciting stimulus, appraisals, action tendencies, somatic and motor responses, feelings, and other

information (such as emotion labels and societal norms) are encoded in the form of a network. Activation of an emotion network is the principal cause of emotions. Emotion networks may be activated via different entry points: via stimuli that resemble the stimuli represented in the network, but also via representations of the somatic and/or motor responses. For instance, the fear network can be activated by stimuli that elicited fear in the past, but also by putting on a fearful facial expression or by engaging in fleeing behavior. Since most network theories also include information about the meaning or appraisal of stimuli, they naturally count as hybrids with appraisal theories.

Network theories come in a biological version, which remain close to evolutionary theories, and a non-biological version, which distance themselves from these theories. In the biological version, organisms come in the world with primitive innate networks, akin to the S–R links in evolutionary theories. These networks are gradually elaborated and refined via learning principles, but they remain centered around their innate core so that a separate network exists for each specific emotion. Despite their close affinitiy with evolutionary theories, biological network theories still differ from evolutionary theories. The difference is not merely a gradual one, in the sense that evolutionary theories focus more on innate processes and biological network theories more on learning processes. Evolutionary theories tend to see the innate processes as purely emotional whereas biological network theories hold that the information that is added via learning is equally (if not more) emotional.

Non-biological network theories, from their side, are not committed to the idea that each emotion network is tied to an innate core. These theories lean on psychometric network theories, assuming that if components cluster into certain packages this is entirely and solely due to the interactions among the components themselves (Lange *et al.* 2020) . It is thereby not excluded that the components configure into the packages hypothesized by evolutionary theories, but there is also no guarantee that they will configure into these packages rather than

into others. Non-biological network theories also lean on principles of dynamic systems models of the mind. Rather than portraying processes as linear input-output units, these models describe the unfolding activity of a system over time, and as emerging from complex, nonlinear interactions among components. The components constrain each other via numerous feedback loops until a stable solution emerges.

Early formulations of network theories (which happen to belong to the biological version; Lang 1994; Leventhal 1984) have proposed that networks have multiple layers, of which only some are capable of producing emotions. This feature was added to these theories to be able to account for dissociations between cold retrieval without emotion, and hot retrieval with emotion. Thus, these theories distinguish between an emotional layer that houses automatic processes that preferrably operate on perceptual representations and/or representations with response-related content, on the one hand, and a non-emotional layer that houses nonautomatic processes that operate on conceptual representations, on the other.

Evaluation theories

Evaluation theories are a family of emotion theories that interpose a process of evaluation between the incoming stimulus and the remaining components. The family immediately branches out into the well-known subfamily of appraisal theories, which count as stimulus evaluation theories (e.g., Arnold 1960; Lazarus 1991; Scherer 2009; see Moors et al. 2013), and the recently developed goal-directed theory, which count as a response evaluation theory (Moors 2017a, 2022). As the terms already suggest, stimulus evaluation theories concentrate on the information that can be extracted from the stimuli, whereas response evaluation theories focus on the evaluation of different response options. I discuss these two subfamilies in turn.

Appraisal theories. Appraisal theories propose that stimuli cause emotions via a process called appraisal. This process evaluates incoming stimuli on a number of criteria such

as the extent to which they are goal-relevant, goal-in/congruent (hence negative/positive), un/expected, un/certain, easy/difficult to control, and caused by internal/external causes (i.e., agency) (see also Ellsworth this volume). Together, the values on these appraisal criteria determine the values of the other emotional components. For instance, an appraisal pattern with the values goal-relevant, goal-incongruent, unexpected, uncertain, difficult to control, and externally caused, leads to the tendency to flee, the physiological activity required to sustain this behavior such as an increased heart rate and blood flowing to the legs, overt flight behavior, and certain feelings that can be labeled as fear. A small change in the appraisal pattern (e.g., appraising a stimulus from difficult to easy to control) may lead to a different action tendency (e.g., from the tendency to flee to the tendency to fight), with the attendant changes in somatic, motor, and feeling components.

The appraisal process takes care of the extraction of appraisal values, which are abstract stimulus features. Appraisal is a process that is defined in terms of the types of information that it processes, or differently put, the *content* of its resulting representations. Appraisal theories do not place restrictions on the *formats* of these representations nor on the types of *operations* that can produce them. For instance, appraisals can be verbal-like or image-like, and the operations producing them can be a rule-based computation or the activation of a previously computed and stored association between a specific stimulus and a pattern of appraisal values (Smith and Kirby 2001; see Moors *et al.* 2013).

Once the appraisal values are determined in the extraction phase, the translation of these values into the other components is again a fairly rigid affair. Like network theories, appraisal theories also come in a biological and a non-biological version. These two versions offer differing proposals about the operations involved in the translation step. Biological appraisal theories (a.k.a. Flavor 1 appraisal theories, Moors 2014) assume that appraisal values are combined into a pattern and linked to a summary appraisal value. For instance, a

pattern of the appraisal values goal-relevant, goal-incongruent, difficult to control, and externally caused is summarized as dangerous. This summary appraisal value is sometimes referred to as the *molar* appraisal and the individual appraisal values as *molecular* appraisals (Moors 2013). The summary appraisal pattern subsequently activates an abstract S–R link or affect program, of the type "danger-defense", which fixes the identity of the emotion (i.e., fear). This, in turn, determines the values of the remaining components: somatic responses that prepare for defense, actual defense, and feelings of fear. Biological appraisal theories can be considered as the counterpart of evolutionary theories that are hybrids with appraisal theories. The difference between the two mainly lies in the fact that appraisal theories focus on the extraction part whereas evolutionary theories focus on the translation part.

Non-biological appraisal theories (a.k.a. Flavor 2 appraisal theories, Moors 2014), on the other hand, assume that each appraisal value separately determines aspects of the motivational component, that is, the action tendency. For instance, an appraisal of goal relevance determines the *intensity* of the action tendency (the more relevant the stimulus, the more intense the tendency), goal in/congruence determines its *direction* (increasing contact if congruent, decreasing contact if incongruent), and low/high control determines its *direction of adaptation* (e.g., a weaker opponent will adjust to the stronger one, the stronger opponent makes the weaker one adjust to them). The action tendency mobilizes physiological responses that prepare and support overt behavior, and aspects of all components make up the content of feelings (Moors and Scherer 2013). Here the organism does not need to determine at any point that an emotion is present and which one; the emotion simply emerges from the integrated sum of the values of all components. Note, however, that each link between an appraisal value (e.g., goal-incongruent) and a particular action tendency value (e.g., decrease contact) is still fixed, although now at the molecular rather than the molar level.

Similar to evolutionary theories, both biological and non-biological appraisal theories add others sources of variability to the mix, such as regulation and planning. This leads to an imperfect relation between appraisals and overt behavior. Suppose a spider phobic appraises the spider in the corner of the room as dangerous and therefore activates a defensive tendency. Acting on this tendency may nevertheless interfere with the goal to finish a task in that room, prompting attempts to inhibit this tendency (i.e., regulation). Even if nothing is standing in the way of defense, however, the person still needs to decide about whether to flee or to hide (i.e., planning).

To summarize the theories discussed so far, in non-hybrid evolutionary theories, extraction of stimulus information is implemented by the transduction of purely concrete features. In appraisal theories and hybrid evolutionary theories, extraction is done by an appraisal process that produces abstract stimulus features like goal in/congruence, un/expectedness, un/certainty ease/difficulty of control, and internal/external cause. Most network theories incorporate appraisals. They still differ from appraisal theories in that activation of an emotion network can happen via the response side and does not have to travel via appraisal before the other components can occur. It is true that contemporary appraisal theories do not insist on appraisal being a necessary cause of emotions (or of other emotional components), but the role of appraisal is less optional than in network theories. In all three theoretical families, the translation of stimulus features in the other components is covered by a stimulus-driven process, that is, the activation of a fixed link between the cognitive and the motivational components. In addition to the stimulus-driven process that is assumed to undergird emotions, the theories discussed so far leave room for a goal-directed process that takes care of regulation and planning. Consistent with classic dual-process models in the domain of motivation/behavior, the dichotomy between stimulus-driven and goal-directed processes is aligned with other dichotomies such as that between automatic and nonautomatic.

In addition, the division of labor between both types of processes follows a default-interventionist architecture in which it is assumed that the stimulus-driven processes involved in emotion causation are automatic and therefore the default process whereas the goal-directed processes involved in regulation and planning are nonautomatic and can occasionally intervene, provided that there is sufficient opportunity, capacity, and motivation to engage in these processes. The goal-directed theory, to be discussed in the next section, breaks with this architecture and proposes instead that not only regulation and planning but also the phenomena of emotions themselves rely on goal-directed processes.

Goal-directed theory. The goal-directed theory recently proposed by Moors (2017a, 2022; Moors, Boddez and De Houwer al. 2017) shifts emphasis from stimulus evaluation to response evaluation. Stimulus evaluation still has a place in this theory, but only to detect whether there is a need for behavior and to help select the behavior that is most suitable in the current context. According to this theory, the organism's first task is not to process information but to satisfy goals. To this end, organisms are constantly engaged in goal-directed cycles that are at the service of their goals. A single cycle proceeds roughly as follows: When a discrepancy with some goal is detected, the organism develops a further goal to reduce this discrepancy. This reduction can be achieved via one of three broad strategies: The first strategy is assimilation (of the stimulus to the goal). This involves engaging in overt behavior and requires the further selection of a behavioral goal or action tendency. The second strategy is accommodation (of the goal to the stimulus). This involves replacing the initial goal by a different goal. The third strategy is immunization. This involves reinterpreting the stimulus so that it is less discrepant with the initial goal. To illustrate, suppose a child wants to play with a toy that another child is holding. The child can try to get access to the toy by resorting to different behaviors: yanking it out of the hands of the other child, luring the other child away, or asking for the toy nicely. The child can also give up the

goal to play with that particular toy and seek out a different toy. Finally, the child can remember playing with the toy before and in this way realize that the goal to play with the toy is therefore already satisfied or partially so. Both the selection of the broad strategies of assimilation, accommodation, and immunization, as well as the concrete action tendencies in the case of assimilation are based on a weighing of the expected utilities of these strategies and options. The expected utility of a strategy/behavior depends on the value of the outcome of the strategy/behavior and the expectancy that the strategy/behavior will lead to this outcome. Crucially, the theory proposes that all goal striving, whether seen as mundane or emotional, is based on a goal-directed process. Stimulus-driven processes may still exist, but their role in behavior causation, including emotional behavior causation, is negligeable. The theory further assumes that organisms have multiple goals, organized in a goal hierarchy, and that some of these may conflict and vie for satisfaction.

Feeling-oriented Theories: Psychological constructionist theories

The next theoretical family to be discussed are psychological constructionist theories. These theories count as feeling-oriented theories because they were, initially at least, formulated to explain the feeling component. This being said, more recent versions of psychological constructionist theories also address the behavior-related components of emotional episodes. Constructionist theories deny the existence of affect programs or any other special-purpose mechanism, that is, a mechanism dedicated to emotions. Instead, they propose that the phenomena called emotions are caused by two or more general-purpose mechanisms. Within the family of psychological constructionist theories, it is worth distinguishing between two versions. The first version is that of Barrett (2006, 2012, 2017), which is reminiscent of Schachter's (1964) two-factor theory. A first factor consists of relatively undifferentiated bodily feelings or core affect. A second factor is a construction process that binds these bodily feelings to an external stimulus and thereby produces a representation with the content

"emotion" or a specific emotion such as "fear", "anger", "sadness", "joy", and so forth. For instance, meeting a growling dog and being insulted by a colleague produce undifferentiated bodily feelings, which become meaningful as fear vs. anger after these bodily feelings have been attributed to their presumed causes and have accordingly been labeled as fear vs. anger. Building further on these seminal ideas, Barrett (2012) has presented the construction process as a categorization process in which sensory input from inside (i.e., core affect) and outside the body (i.e., external stimuli) are integrated with learned knowledge about emotion categories, called emotion concepts. An emotion concept is a collection of situation-specific scripts. For instance, the anger concept is a collection of scripts for anger in traffic, anger in competition, anger about betrayal, and so forth. Each script contains information about a situation in which an emotion label was used together with behavior and its sensory consequences. Scripts come in multimodal and embodied format so that activation of a script results in the re-experience and partial re-enactment of its components (Wilson-Mendenhall *et al.* 2011). Moreover, Barrett (2012) has characterized this categorization process on a lower level of analysis as a pattern completion process in which a script comprised of multiple elements gets activated by a small piece of that script (Barrett 2006), much like what has been proposed in network theories.[2] More recently, she has added that the script functions as a prediction that can be confirmed or disconfirmed, leading to a readjustment of the prediction (Barrett 2017), in line with predictive processing theories first developed in the domain of perception (Clark 2013).

The second version of psychological constructionist theories is that of Russell (2003, 2009, 2012), a theory that can more aptly be described as a multi-factor theory. The phenomena called emotions house a range of components: all kinds of information processes (i.e., cognitive component), action tendencies (i.e., motivational component), physiological responses (i.e., somatic component), subtle and coarse behavior (i.e., motor component), core

affect (i.e., undifferentiated feelings), and emotional meta-experience (i.e., labeled feelings). In this theory, construction refers to the fact that components in the episodes that people call emotional are not preset (neither by an innate affect program nor by a learned network) but are instead configured on the fly. Russell (2012) argues that strong causal ties among components cannot be assumed on an a priori basis and therefore that each component must be studied in its own right by handing it over to its respective research area: Action tendencies and overt behavior must be studied by motivation/behavior research; experience is the subject matter of consciousness research. The two-factor idea of Schachter (1964) can still be recognized if we zoom in on the component of emotional meta-experience—the experience of having an emotion—which is the result of the categorization of all other components in terms of a discrete emotion. In Russell's (2005) theory, the categorization process is based on the resemblance of a configuration of components to a stored mental script for a particular emotion, which will be a prototypical instance with prototypical components. While Barrett (2012) believes that the result of the categorization process is an emotion, Russell (2009, 2014) believes that the result of the categorization process is nothing but an emotional self-ascription, which is not an emotion. Thus, Barrett (2012) believes that people who categorize their bodily feelings as fear *have* the emotion of fear whereas Russell (2009) believes that people who categorize their bodily and mental changes as fear merely *perceive themselves as having* the emotion of fear.

Social theories

Each of the previously discussed personal theories acknowledges the fact that emotions often occur in a social context. While the personal versions of these theories do not go far beyond acknowledging this fact and sampling illustrations of emotional events from the social realm, some theorists have developed social versions of these theories in which the implications of the social context of emotions are worked out in more detail. Social emotion theorists have

expanding the focus from a single individual to several individuals: dyads, groups, and cultures. While dyads and groups involve actual interactions among individuals, the cultural level is more concerned with cultural norms about emotions. When actual interactions are considered, another natural expansion is from a single interaction unit (from sender to receiver) to multiple interaction units (from sender to receiver and back). In addition to an increased attention to complexity in terms of numerosity (i.e., number of interaction partners) and time (i.e., number of interaction units across time), this exercise has also led to the proposal of specifically social mechanisms involved in emotion causation. Below, I highlight a few social extensions for each of the personal theories discussed above.

In evolutionary theories, a first social extension is concerned with the evolved social functions of emotions in dyads, groups, and cultural societies (Keltner and Haidt 1999). Emotions in dyads help coordinate the interactions among the interaction partners. Emotions in groups help to establish group identities and group roles. Socio-cultural norms about emotions, finally, regulate the expression of emotions. A second social extension focuses on the communication of emotions in social interactions, in particular on the evolutionary origin of emotional expressions (Shariff and Tracy 2011; Sauter and Russell this volume). Based on their reading of Darwin (1872), evolutionary theories propose that in a first stage of evolution, certain facial displays developed as adaptations, that is, behaviors that prepare for emotional behaviors. An example is baring the teeth as a preparation for biting an opponent. In a later stage of evolution, these facial displays developed into exaptations, that is behavior that adopted a novel function, namely communication. Although initially, baring the teeth did not evolve to communicate anger, it allowed receivers to infer that the sender is angry. In a later stage of evolution, baring the teeth was co-opted to serve the novel function of communicating anger (but see Fridlund 1994, 2017, for a critique).

In network theories, a notable social extension is the proposal of a social mechanism for emotion causation known as emotional contagion (Hatfield et al. 1994). A receiver mimics the emotional responses (e.g., facial expressions) of a sender, after which somatosensory feedback of these responses induces a similar emotion in the receiver. For instance, if a sender looks scared, the receiver mimics this expression, which produces fear in the receiver in turn. This mechanism fits with network theories, and in particular, their assumption that emotions can be activated via the response side, via the artificial induction of emotional responses.

In appraisal theories, a first social extension has explored whether some appraisal criteria should receive different weights in different cultures and even whether the set of appraisal criteria must be expanded with culture-specific ones (e.g., Markus and Kitayama 1991; see Mesquita and Ellsworth 2001). A second social extension is the proposal of a social mechanism of emotion causation known as social appraisal (Manstead and Fischer 2001). A receiver infers how to appraise a stimulus based on the emotional responses (e.g., facial expression) of a sender, and this leads to a similar or complementary emotion. If a sender looks scared, for instance, the receiver may infer that some danger lurks around the corner and start feeling scared as well. A final social extension of appraisal theories is captured in the mechanism of distributed appraisal (Parkinson 2012). Here, the various appraisal checks are not made all at once, but they are negotiated across multiple interaction units. Thus, a person may first appraise a situation as goal-congruent, to only later find out that it is goal-incongruent, and even later still that the situation is difficult to control, but that no one is to blame.

In the goal-directed theory, a first social extension proposes to analyze communicative and other social behavior in terms of goal-directed processes (e.g., Fridlund 2017; Moors 2022). Rather than assuming that people's facial behavior expresses or communicates emotions, it may be seen as instrumental for reaching a certain goal. For instance, crying does

not express sadness but may solicit attention and comfort from others, a frown does not express anger but may be a way to make others adjust their tone, and smiling does not express happiness but may solicit sympathy from others. A second social extension of the goal-directed theory proposes a distributed goal-directed process in which the weighing of the expected utilities of action options does not occur in a single shot but is spread out and adjusted across multiple interaction units (Moors 2022).

Social extensions of psychological constructionist theories are known as social constructionist theories. While the categorization process in psychological constructionism is located in the individual, social constructionism works out the role of culture in shaping the category knowledge used in this process. Some social constructionists (e.g., Averill 2012) show affinity with Russell's (2009) theory whereas others (e.g., Mesquita, in Mesquita and Parkinson this volume) find themselves in closer connection with Barrett's (2017) theory. Despite the compatibility that social constructionist theories have with psychological constructionist theories, the former did not exactly *grow out* of the former, and some social constructionist theorists are not committed to the mechanisms proposed in existing psychological constructionist theories. For instance, Parkinson (2012) subscribes to the general idea that emotions are constructions in the sense that they are made up of mechanisms that are not emotional per se. These mechanisms may not be restricted to the categorization of bodily feelings as proposed by Barrett (2006), but may extend to certain appraisal processes proposed by non-biological appraisal theories.

Empirical Testing

In this section, I discuss the empirical evidence marshalled in favor of the mechanisms proposed in the emotion theories considered. I also introduce some of the critiques received by the cited research programs, jointly with some replies offered by the theories' proponents. For lack of space, this discussion is cursory.

Evolutionary theories

Affect program theorists have sought and claimed to have found direct and indirect evidence for the existence of affect programs dedicated to basic emotions. Direct evidence is evidence for the existence of these affect programs in the brain, either in the form of specific neural circuits (Johansen *et al.* 2011; Öhman and Mineka 2001) or specific neural patterns (Saarimäki *et al.* 2016). Indirect evidence is based on the following reasonings (Ortony and Turner 1990). First, if each basic emotion has an affect program with a unique function, it should have distinct components (short, *distinctness*) such as a specific physiological response pattern (reviews in Ekman 1992; Kreibig 2010; Levenson 2014) and a specific expressive behavior (reviews in Ekman 1999; Keltner *et al.* 2016; Matsumoto *et al.* 2008). Second, if each basic emotion has an affect program, which is innate, then it should be universal and so should be the components of that emotion (short, *universality*). That is, there should be universal physiological responses and universal facial expressions. Evidence for universality has been collected across cultural groups (Ekman et al. 1969; Elfenbein and Ambady 2002; Tracy *et al.* 2013) but also across seeing and congenitally blind people (Matsumoto and Willingham 2009, but see Galati *et al.* 1997). Third, if each basic emotion has an affect program, which causes a number of output components, these output components should show strong correlations (short, *concordance*). For instance, fighting behavior should correlate with an angry facial expression and angry feelings (see reviews in Levenson 2014; Matsumoto *et al.* 2008).

Empirical evidence for the existence of affect programs, whether direct or indirect has been the topic of huge controversy. A first line of criticism concerns the empirical status of the evidence to date. Critics have pointed at meta-analyses that have not been able to present evidence for distinctness, universality, and concordance across studies, but show evidence of variety instead (e.g., Cacioppo *et al.* 2000; Murphy, Nimmo-Smith and Lawrence 2003;

Larsen *et al.* 2008; Phan *et al.* 2002; Quigley and Barrett 2014). Critics have also pointed their arrows at methodological shortcomings (Nelson and Russell 2013; Russell 1994; Russell, Bacharowski and Fernández-Dols 2003; Gendron *et al.* 2014). For instance, it has been argued that many studies designed to test the existence and universality of emotion-specific facial expressions make use of a recognition method, while they should be using a production method instead, preferably one in which events elicit spontaneous facial expressions. If participants are asked to recognize emotions from posed facial expressions or to produce artificial facial expressions, they may tap into their learned stereotypical scripts to solve the task (Barrett 2011; Fridlund 1994; Lindquist and Gendron 2013; Parkinson 2013). Furthermore, recent spontaneous production studies do not seem to replicate previous evidence, neither in the laboratory (Reisenzein, Studtmann and Horstmann 2013) nor in the field (Fernández-Dols and Crivelli 2013).

Evolutionary theorists have insisted on a more favorable reading of the evidence (Ekman 1999; Keltner *et al.* 2016; Shiota this volume). In addition, they have argued that part of the variety observed in certain studies is already predicted or acknowledged by evolutionary theories. Sources of variety include (a) the combination of affect programs in mixed emotions (e.g., if a stimulus elicits both fear and sadness, components of both emotions may occur), (b) learning processes and computation, which create variety within emotion families in terms of the stimuli that elicit the emotion and the somatic and motor responses in which they are manifested (e.g., the disgust family contains primitive disgust and moral disgust, which may be characterized by different physiological responses and behavior), (c) emotion regulation, which can create variety in the somatic, motor, and/or feeling components of emotions (e.g., suppressed and non-suppressed anger will vary in their physiological and behavioral profile), (d) other non-emotional processes (e.g., somatic responses also serve homeostasis, and facial behavior is also used for verbal communication), and (e)

methodological constraints of past research such as weak methods for the induction of emotions (e.g., due to ethical concerns) and insensitive methods for the measurement of emotional components. Finally, evolutionary theorists have argued that distinctness does not need to be manifested in each and every component. For instance, some emotions do not have a distinct facial expression, but rather a distinct vocal or gestural expression (e.g., pride is characterized by a tilted head and expanded chest; Tracy, Shariff and Cheng 2010), or they have a distinct expressive pattern across various expressive channels and even components (e.g., love is characterized by smiling, a tilted head, open hand gestures, and oxytocin release) (Gonzaga *et al.* 2006; Keltner *et al.* 2016)

A second line of criticism raised against evolutionary theories concerns the *interpretation* of the data that do seem to support distinctness, universality, and concordance. Alternative explanations are that the brain mechanisms and responses found are not specific to emotions but to emotional components or parts thereof (e.g., frowning may be specific to goal-incongruent situations rather than that an entire scowling face is specific to anger) or to non-emotional phenomena (e.g., frowning may be specific to effort rather than to anger).

Even if it would be shown that the brain mechanisms and responses found are specific to emotions, this is not yet evidence that these emotions are caused by the affect programs proposed by evolutionary theories rather than by the learning mechanisms proposed in non-biological network theories or by the links between appraisal values and action tendency values in non-biological appraisal theories. For instance, evidence for the universality of a facial expression is not proof of the existence of an affect program because universality may also indicate that different cultures developed the same behavioral solutions to similar problems (convergent cultural evolution; Fridlund 1994). Likewise, critics argue that evidence that congenitally blind people produce the same facial expressions as sighted people is not

proof of the existence of innate affect programs, because most likely, blind people also have a history of being rewarded for producing culturally appropriate facial expressions.

Network theories

Some proponents of network theories have been active in discovering the learning principles involved in the elaboration of emotion networks (see research on fear conditioning; Beckers *et al.* 2013; Vervliet and Boddez 2020). Others have investigated the interplay between emotion and cognitive functions such as attention and memory and their role in the development of psychopathologies such as anxiety disorders and depression (e.g., MacLeod, Grafton and Notebaert 2019; Teachman *et al.* 2019).

Other research carried out under the flag of network theories has addressed the hypothesis that emotions can be elicited via the response side. Some studies focus on emotion induction via the administration of physical substances (i.e., drugs). Other studies focus on emotion induction via the adoption of artificial facial expressions (i.e., facial feedback hypothesis; Allport 1924; Niedenthal, Krauth-Gruber and Ric 2006). The latter research is illustrated by the often-cited study of Strack, Martin and Stepper (1988) who reported that participants judged cartoons as funnier when they held a pen between the teeth—which induced a happy facial expression—than when they held a pen between the lips—which induced a sad facial expression.

Here again, critics have questioned the robustness of the data. For instance, a large-scale meta-analysis of newly conducted replication studies of the Strack *et al.* (1988) study turned out to be unable to replicate the effect (Wagenmakers *et al.* 2016). In addition, critics have questioned the interpretation of the data. Even if future research would show robust evidence that emotions can be induced or influenced via artificial somatic and motor responses, there is no reason to believe that this influence does not travel via appraisals (as

appraisal theories would have it; Arnold 1960) or via the categorization of core affect (as certain psychological constructionist theories would have it; Barrett 2006).

Evaluation theories

This section offers a glimpse of the research programs carried out by stimulus evaluation theories such as appraisal theories and response evaluation theories such as the goal-directed theory.

Appraisal theories. Research testing appraisal theories can be split into several research lines. The first of these has addressed the hypothesis of radical versions of appraisal theory (e.g., Lazarus 1991), which hold that appraisal is a necessary cause of emotion. Critics have pointed at various types of observations that supposedly disprove this idea. Appraisal theorists, in turn, have invalidated several of these arguments (Moors 2013). To illustrate, a number of studies conducted by Zajonc and his colleagues (Kunst-Wilson and Zajonc 1980; Murphy and Zajonc 1993) showed that affect, which is considered to be an inchoate form of emotion, can be elicited without cognitive processing. Appraisal theorists, however, have argued that these studies merely show that affect can be elicited without non-evaluative forms of cognition that do not qualify as appraisal. Another illustration is provided by studies already discussed above, which try to show that emotions can be elicited via artificial induction of emotional responses (e.g., taking drugs or adopting artificial facial expressions). As mentioned, however, these studies do not produce robust, replicable effects, and they are not immune to the alternative explanation that the responses bypass appraisal. Other illustrations are evidence for LeDoux's (1996) fast route by which stimuli activate the amygdala—which is considered to be the fear center—while bypassing cognitive processing in the neocortex, and emotions that are supposed to be triggered via direct brain stimulation. Here again, the generality, validity, and relevance of the findings has been called into question. LeDoux's (1996) fast route was found to be extremely rare, especially in humans

but even in animals (Pessoa and Adolphs 2010; Storbeck, Robinson and McCourt 2006). Evidence for emotions caused by direct brain stimulation has been judged as unconvincing (Barrett *et al.* 2007). But even if it were convincing, appraisal theorists have argued that activating the proximal cause of emotions via direct brain stimulation does not falsify appraisal theory' idea that this proximal cause is in normal circumstances triggered by a remote cause, which is the appraisal process.

The second line of research has studied hypotheses about causal relations between appraisals and emotions or emotional components. In early appraisal research—conducted by biological appraisal theorists—the aim was to study the relation between specific appraisal patterns and specific discrete emotions, such as the hypothesis that goal-incongruent stimuli that are difficult to control lead to fear, whereas goal-incongruent stimuli that are easy to control lead to anger (Ellsworth and Scherer 2003; Roseman, Spindel and Jose 1990). These relations were often studied using self-report methods. Participants are asked to remember or imagine an emotion and to rate the appraisals that caused or would cause it. Alternatively, participants are presented with descriptions of appraisal patterns and they are asked which emotions these would evoke. Using these methods, individual appraisal researchers have been able to confirm part of their hypotheses and it has also led to refinements and adjustments of the original hypotheses.

This early appraisal research program has received extensive criticism, especially on methodological grounds. It has been argued that the use of verbal methods and self-report in these studies invites participants to tap into stereotypic scripts linking specific appraisal patterns with emotions rather than that it reveals anything about causal relations between actual appraisals and actual emotions (Frijda and Zeelenberg 2001; Moors and Scherer 2013; Parkinson 1997). In addition, this research faces a similar difficulty as research testing evolutionary theories, namely that any empirical study that tries to examine the relation

between a component (in this case, appraisal) and an emotion ends up measuring the relation between that component and another component that is supposed to measure the emotion. Thus, much research aimed at testing the relation between appraisals and emotions can be reframed as research testing the relation between appraisals and the feeling component, because emotion words coincide with feeling words (Moors 2017a).

Later appraisal research—conducted by non-biological appraisal theorists—has embraced this latter fact and has formulated the explicit aim to study relations between appraisals and other components, without linking these back to discrete emotions. For instance, researchers have tried to discover the appraisal patterns that cause the tendencies to fight vs. flee without linking these to the discrete emotions of anger vs. fear. These researchers, moreover, strive to manipulate appraisals via real situations (in the lab or in the field) instead of via verbal descriptions of appraisals and to measure components with objective measures instead of subjective (i.e., self-report) ones. Among the components that can be measured in this way are action tendencies (Moors *et al.* 2019), physiological responses (Kreibig, Gendolla and Scherer 2012), subtle behavior in the face (Mehu and Scherer 2015; Scherer *et al.* 2018), voice (Goudbeek and Scherer 2010; Laukka and Elfenbein 2012), and gestures (Dael, Goudbeeck and Scherer 2013). Large-scale meta-analyses of this type of appraisal research are lacking, but the evidence is mixed (see Scherer and Moors 2019, for a review). For instance, research testing the influence of appraisals on action tendencies often failqs to confirm the hypotheses put forward by appraisal theories. For instance, Fischer *et al.* (2022) manipulated the degree of control that participants had over a goal-incongruent stimulus and found faster tendencies to fight when participants had low control than when they had high control. This goes against the hypothesis that high control leads to the tendency to fight and low control to the tendency to flee.

Goal-directed theory. Empirical research testing the goal-directed theory is still in its early days. A few illustrations are nevertheless worth citing. Parkinson (2001) showed that drivers report more aggressive behavior (e.g., honking) when they are in their car than when they are not, and that this cannot be attributed to differences in appraisals of the situation but rather to the expected utility of this behavior. Likewise, Fischer, Kuppens and Moors (2020) showed that socially excluded participants who were given the choice between aggressive and prosocial action options preferred the aggressive option whereas those who could also choose a moralizing action option preferred the moralizing option. This suggests that the tendency to engage in aggression depends on the expected utility of that behavior relative to that of other behaviors. More recent studies using neuroscientific methods have found preliminary support for the assumption that goal-directed processes can determine initial action tendencies rather than only being involved in emotion regulation or planning (e.g., Fischer *et al.* 2020; Moors *et al.* 2019).

Psychological constructionist theories

Two-factor versions of psychological constructionism try to demonstrate that Factor 1, consisting of undifferentiated bodily feelings (arousal or core affect), needs to be combined with Factor 2, which consists of a construction process (attribution and/or categorization), in order to produce specific emotions. This involves showing that each factor alone does not produce a specific emotion while their joint presence does. One approach has been to conduct misattribution studies in which bodily feelings were manipulated independently from the presence of a salient source to which these bodily feelings could be (falsely) attributed, and in which both elements were shown to contribute to labeled feelings and/or emotional behavior (Dutton and Aron 1974; Schachter and Singer 1962). In these studies, it was shown that only participants in whom strong bodily feelings were induced and who were exposed to a salient stimulus in the environment that allowed them to interpret their bodily feelings in emotion-

specific terms reported having specific feelings and displayed emotional behavior. Recent research has revisited this line of work, even if it has shifted emphasis from attribution to categorization (as in Barrett's 2006 theory). Support for the role of categorization comes from studies in which manipulation of the accessibility of emotion categories (e.g., via priming vs. satiation procedures) influenced the emotions that participants experienced themselves (e.g., Lindquist and Barrett 2008) or perceived in others (e.g., Lindquist *et al.* 2006). Other research has concentrated more on the role of bodily feelings (i.e., Factor 1) in emotions. This research has shown, for instance, that the intensity of bodily feelings correlates with the intensity of emotions (MacCormack and Lindquist 2017).

Note that much of the evidence adduced for the two-factor version of psychological constructionism endorsed by Barrett (2012) also counts count as evidence for Russell's (2009) version of psychological constructionism. The only difference is that the former version takes self-reported emotions to be evidence of real emotions whereas the latter version takes them to be mere self-ascriptions. The self-report methods used to measure emotions cannot tell real emotions apart from self-ascribed ones.

Critics of the psychological constructionist research program have pointed out that evidence for the influence of arousal on the intensity of emotions and for the influence of categorization on the quality of emotional experience or perception does not provide unique support for psychological constructionism because it is also compatible with other emotion theories. Appraisal theories, for instance, accept that once the emotion has occurred, the subsequent labeling of the emotion can further shape the emotional experience after the fact. Two-factor versions of constructionism, however, make the more radical claim that the two factors of bodily feelings and the construction process are *necessary* for emotions. If one of these is absent, there can be no emotion-specific quality. As I have argued elsewhere (Moors, 2016), however, researchers can never definitively demonstrate the necessity of something.

Showing the necessity of the construction process for obtaining emotion-specific quality in one setting is not enough and demonstrating it in all possible settings is not feasible. Claims of necessity can only be plausibly maintained until there is evidence to the contrary. Rivalling emotion theories, such as evolutionary theories, network theories, and appraisal theories, are convinced that their research shows evidence of emotion-specific quality prior to the intervention of any construction process.

Social theories

Social theories have set up various lines of research. Some of these apply the mechanisms of personal theories in a social context in an attempt to unravel dynamic patterns across multiple interaction partners and over time. Other lines of research have specifically focused on testing the social mechanisms that have been proposed in the various social theories such as emotional contagion, social appraisal, and distributed forms of social appraisal (see Parkinson and Manstead 2015). Still other lines of research have addressed the question whether appraisals are weighed differently across cultures, whether there are appraisals that are unique to certain cultures, and whether there are emotions that are unique to certain cultures (see Mesquita and Ellsworth 2001; Mesquita and Parkinson this volume).

Scientific Definitions

While the research programs conducted under the umbrellas of the various emotion theories have not provided definitive evidence so far for any of these theories, this has not stopped several of them from formulating scientific definitions for the purpose of *demarcating* the set of emotions and from proposing principled ways for *partitioning* this set. Let us start with demarcation.

Demarcation

It may be recalled that scientific definitions ideally list the necessary and sufficient conditions or criteria for an instance to belong to the set.[3] Conditions can be sampled from the

components listed in the constitutive explanations and/or the causal-mechanistic explanations proposed by theories.

In none of the theories discussed so far, the mere presence of the components that they assume to be part of the emotion suffices to distinguish emotions from non-emotional phenomena. Indeed, mundane events like dropping a pen on the floor give rise to all the components that have been identified as being part of emotions, such as appraisal of the event as goal-incongruent and easy to control, the tendency to pick up the pen, physiological responses that prepare the body for picking up the pen, the actual picking up of the pen, and traces of all of these components integrated in conscious feelings. Yet few theorists will classify this as an emotion, at least not a strong one. To demarcate the set of emotions from other sets, more is needed. To this end, theories have sampled additional necessary conditions from their causal-mechanistic explanations. I will review the criteria that flow from the causal-mechanistic explanations proposed by the theories, and examine whether or not they buy the set of emotions scientific status (**Axis 4**). I will restrict my discussion to the personal theories because the social versions of these theories can be said to follow in the footsteps of the personal ones.

Evolutionary theories propose that emotions are sets of components that are caused by an affect program. For instance, Ekman and Cordaro (2011) stated that affect programs are the sine qua non of emotions. This criterion is straightforward. If empirical research would come to confirm the existence of dedicated affect programs, they would turn the set of emotions into a scientific set.

Few network theories have explicitly addressed the demarcation problem, but if we follow their logic, emotions are sets of components that are caused by an emotion network. This raises the question what makes an emotion network emotional. In the biological brand of network theories at least, an emotional network is one that is built around an affect program.

Moreover, these theories argue that only a certain layer (which automatically activates salient perceptual nodes or nodes with a response content) is capable of eliciting true emotions (Lang 1994; Leventhal 1984).

Appraisal theories typically hold that emotions are sets of components that are caused by the appraisal of a stimulus as being highly goal-relevant, in the sense that it has an impact on a goal of high importance (Frijda 1986; Lazarus 1991; Moors 2007; Oatley and Johnson-Laird 1987). In the pen example, the goal to hold on to the pen could be of minor importance, which is why it does not generate an emotion. An appraisal pattern including high goal relevance translates into an action tendency with control precedence and perhaps into strong synchronicity or coherence among all components (i.e., quantitative changes in these components occur within the same time window, directed towards the fulfilment of the aim of the emotion).[4] It should be noted, however, that the criterion of goal relevance is gradual in nature (because goal relevance is a dimension and not a matter of all or nothing), thereby allowing only relative statements about whether something is an emotion or not.

The goal-directed theory (Moors 2017a) proposes that episodes that are considered to be emotions are undergirded by the exact same goal-directed mechanism as episodes of mundane goal pursuit. Thus, the causal-mechanistic explanation of emotion proposed by the goal-directed theory does not allow demarcating the set of emotions from other sets. This being said, the theory recognizes that there may be descriptive differences between emotions and mundane goal pursuit. Similar to appraisal theories, the theory suggests that in the phenomena called emotions the goals at stake tend to have higher values, which leads to action tendencies that are more intense and have more control precedence. In contrast with appraisal theories, however, the goal-directed theory does not take the criterion of goal relevance to be deep enough to grant the set of emotions scientific status.

Psychological constructionist theorists have emphasized that the mechanisms proposed in their theories are general-purpose, in the sense that they are not dedicated to emotions. In Barrett's (2006) theory, the bodily feelings (arousal or core affect; Factor 1) and the construction process (categorization; Factor 2) are not unique to emotions. We always have bodily feelings, and we engage in categorization processes for many purposes that have no emotional outcome. The categorization process proposed in these theories indeed qualifies as general-purpose when it is described in terms of the *operations* involved. This is no longer the case, however, if this process is described in terms of the *content* of the representations involved, that is, as the transition from a representation of bodily responses to a representation of a specific emotion. This is because not all categorizations end up in such a representation. Moreover, the fact that Barrett (2006, 2012) takes the representation resulting from this categorization process to be a real emotion, ironically turns this process into a mechanism that is again dedicated to emotions.

Russell (2003, 2009), on the other hand, considers the representation resulting from his categorization process to be a mere self-ascription of an emotion, which is not the same thing as an emotion. It is not because one ascribes an emotion to oneself, that one is actually having an emotion, or that emotions even exist. Compare it to supernatural powers. It is not because one ascribes supernatural powers to oneself, that one actually has these powers, or that supernatural powers even exist (Moors 2017b). In conclusion, while the mechanisms proposed by Barrett (2006) seem to be set up to grant the set of emotions scientific status, Russell (2003) explicitly denies such a scientific status.

In sum, evolutionary theories, and in their wake, biological versions of network theories and appraisal theories are vindicator theories in that they believe that a fruitful scientific definition can be proposed for the set of emotions. While some individual theories within these families take emotions to be like "water", some of them prefer thinking of

emotions as "lemons" and propose a cluster-type definition rather than a classical definition. Rather than treating the conditions or criteria cited above as necessary, they treat them as typical instead. Still others prefer thinking of emotions as "fish" or "jade" in the sense that they trim or regiment the folk set by throwing out a smaller or larger portion of the original set.

In sharp contrast with the above theories, the goal-directed theory and Russell's (2003) psychological constructionist theory count as skeptical or eliminativist theories. They take emotions to be like "air", that is, a folk set that cannot reach the status of a scientific set. Finally, both non-biological appraisal theories and Barrett's (2012) psychological constructionist theory can be said to take an intermediate position on this axis. The mechanisms they propose do not qualify as emotion-specific when only the operations involved in them are considered (appraisal in appraisal theories; bodily feelings and categorization in Barrett's theory). However, the content or output of these mechanisms does qualify as emotion-specific (appraisal of a stimulus as highly goal-relevant in appraisal theories; categorization of bodily feelings as a specific emotion in Barrett's theory).

Partitioning

The causal-mechanistic explanations proposed by emotion theories also serve as the basis for partitioning of the set of emotions into subsets (i.e., **Axis 5**). In evolutionary theories, each basic emotion is caused by its own affect program; in network theories, each emotion is caused by its own network; and in appraisal theories, each emotion is caused by its own appraisal pattern. In evolutionary theories, the number of affect programs is limited, and so are the number of networks in biological versions of network theories as well as the number of appraisal patterns in biological appraisal theories. As a consequence, these theories are happy to subdivide the set of emotions into subsets corresponding to discrete emotions—they endorse a discrete view.

In the non-biological version of network theories, by contrast, networks need not remain or center around discrete emotion categories, and in the non-biological version of appraisal theories, the number of possible appraisal patterns is infinite and hence so is the number of possible emotions flowing from these patterns. As a consequence, these theories believe that the variety within the set of emotions is more parsimoniously organized by pointing at dimensions on which emotions can vary—they endorse a dimensional view. Whereas network theories with a dimensional view make no specific proposals about the nature of the relevant dimensions, appraisal theories single out dimensional appraisal criteria as the organizing principles.

Psychological constructionists such as Barrett (2006) and Russell (2003) reject a discrete view. They have argued that the evidence to date does not support the existence of dedicated affect programs for discrete emotions but rather that it suggests dedicated neural substrates only for valence and arousal. It is precisely because of their focus on valence and arousal, that these authors have been misunderstood as holding a dimensional view in which all the variety within the set of emotions can be captured by the dimensions of valence and arousal. In reality, however, Barret, Gendron and Huang (2009) have proposed combining a dimensional with a discrete view. If Factor 1 (core affect, composed of valence and arousal) is considered, a dimensional view is fitting, whereas when Factor 2 (categorization of core affect in terms of discrete emotions) is considered, a discrete view is more appropriate. Russell (2003), from his side, does not commit to either view. The phenomena that people call emotions can differ in various ways. These can be organized using all kinds of categories or dimensions depending on the question at hand.

Conclusion and Future Directions

To summarize, five families of emotion theories—evolutionary theories, network theories, evaluation theories (including appraisal theories and the goal-directed theory),

psychological constructionist theories, and social theories—were compared with the help of a framework outlining four stages in theory development. This yielded the following five axes for the comparison of these theories. **Axis 1** compares theories with respect to the working definition that they propose, an axis on which theories turned out to present more similarities than differences.

Axis 2 organizes theories according to the number and nature of the components that they include in an emotion. Most theoretical families present within-family variation regarding this axis. This means that in a single family, the constitutive explanations of individual theories can range from narrow (including few components) to broad (including many components). It may be noted, however, that within the realm of psychological theories of emotion, the breadth of the constitutive explanation is not considered to be of major importance. Whether or not a component is part of the emotion proper or is treated as a cause or a consequence instead is looked upon as a fairly arbitrary matter in this discipline. To avoid that differences in constitutive explanations would pose a hurdle in bringing theories closer together, several theorists have switched talk of emotions to talk of emotional episodes (which extend beyond the emotion proper to also include the immediate causes and consequences).

The true differences between psychological emotion theories reside in the causal-mechanistic explanations that they propose. This is **Axis 3**. The central mechanisms for emotion causation are innate affect programs in evolutionary theories, learned emotion networks in network theories, an appraisal process with the output goal-relevant in appraisal theories, a goal-directed process in the goal-directed theory, the categorization of bodily feelings into discrete emotion categories in certain psychological constructionist theories, and extensions of several of these mechanisms for social contexts in social theories (e.g., social contagion, social appraisal, distributed appraisal).

Several theories have used their central mechanism for emotion causation as the basis for the demarcation and partitioning of the set of emotions. **Axis 4** captures the difference between vindicator theorists, who believe to have found a mechanism (or set of mechanisms) that allows grounding the set of emotions in such a way that it acquires scientific status, and skeptical or eliminativist theorists, who explicitly reject this possibility. **Axis 5**, finally, captures the difference between theories with a discrete view, which hold that the set of emotions is best partitioned into discrete emotion subsets, theories with a dimensional view, which hold that this set is more parsimoniously organized by identifying multiple dimensions, and theories that combine a discrete view with a dimensional view.

The sharp distinctions between psychological emotion theories drawn so far can be softened by pointing at the existence of several hybrid theories such as biological versions of network theories and appraisal theories that incorporate affect programs, and evolutionary theories and network theories that incorporate a role for appraisal. Differences between non-hybrid theories, however, remain the topic of intense controversy and ongoing empirical research. Integration between existing psychological theories therefore does not seem within immediate reach. This does not exclude that future players on the theoretical scene might consider adopting an increased openness towards integration rather than towards competition. This would involve searching for common ground and/or whether theorists are dealing with different aspects of the explanandum (i.e., behavior vs. experience) opening the door for co-existence.

References

Allport, F. H. (1924) *Social Psychology*, Boston: Houghton Mifflin.

Arnold, M. B. (1960) *Emotion and Personality*, New York: Columbia University Press.

Averill, J. R. (2012) "The Future of Social Constructionism: Introduction to a Special Section of Emotion Review," *Emotion Review 4*(3), 215-220.

Barrett, L. F. (2006) Solving the Emotion Paradox: Categorization and the Experience of Emotion. *Personality and Social Psychology Review, 10*, 20-46.

---- (2012) "Emotions are Real," *Emotion, 12*(3), 413-429.

---- (2017) "The Theory of Constructed Emotion: An Active Inference Account of Interoception and Categorization," *Social Cognitive and Affective Neuroscience, 12*(1), 1-23.

Barrett, L. F., Gendron, M. and Huang, Y. M. (2009) "Do Discrete Emotions Exist?" *Philosophical Psychology*, *22*(4), 427-437.

Barrett, L. F., Lindquist, K. A., Bliss-Moreau, E., Duncan, S., Gendron, M., Mize, J. and Brennan, L. (2007) "Of mice and Men: Natural Kinds of Emotions in the Mammalian Brain? A Response to Panksepp and Izard," *Perspectives on Psychological Science*, *2*(3), 297-312.

Beckers, T., Krypotos, A. M., Boddez, Y., Effting, M. and Kindt, M. (2013) "What's Wrong with Fear Conditioning?" *Biological Psychology*, *92*(1), 90-96.

Bower, G. H. (1981) "Mood and Memory," *American Psychologist, 36*, 129-148.

Cacioppo, J. T., Berntson, G. G., Larsen, J. T., Poehlmann, K. M. and Ito, T. A. (2000) "The Psychophysiology of Emotion. in M. Lewis & J. M. Haviland-Jones (eds.) *Handbook of Emotions.* 2nd ed., New York: Guilford.

Carnap, R. (1950) *Logical Foundations of Probability*, Chicago: University of Chicago Press.

Clark, A. (2013) “Whatever Next? Predictive Brains, Situated Agents, and the Future of Cognitive Science,” *Behavioral and Brain Sciences, 36*(3), 181–204.

Clore, G. L. and Ortony, A. (2000) “Cognition in Emotion: Always, Sometimes, or Never?” in R. D. Lane & L. Nadel (eds.) *Cognitive Neuroscience of Emotion,* Oxford: Oxford University Press.

Dael, N., Goudbeek, M. and Scherer K. R. (2013) “Perceived Gesture Dynamics in Nonverbal Expression of Emotion,” *Perception, 42(6,)* 642 – 657.

Darwin, C. (1872) “*The Expression of the Emotions in Man and Animals by Charles Darwin*,” John Murray.

Dutton, D. G. and Aron, A. P. (1974) “Some Evidence for Heightened Sexual Attraction under Conditions of High Anxiety,” *Journal of Personality and Social Psychology, 30*(4), 510-517.

Ekman, P. (1992a) “An Argument for Basic Emotions,” *Cognition and Emotion, 6*(3-4), 169–200.

---- (1999) “Basic emotions,” in T. Dalgleish & M. J. Power (eds.), *Handbook of Cognition and Emotion*, Chichester: Wiley.

Ekman, P. and Cordaro, D. (2011) “What is Meant by Calling Emotions Basic,” *Emotion Review*, *3*(4), 364-370.

Ekman, P., Sorenson, E. R. and Friesen, W. V. (1969) “Pan-cultural Elements in Facial Displays of Emotion,” *Science*, *164*(3875), 86-88.

Ellsworth, P. C. (2013) “Appraisal Theory: Old and New Questions,” *Emotion Review*, *5*(2), 125-131.

----- (this volume) “Appraisal Theories,” in A. Scarantino (ed.), *Routlegde Handbook of Emotion Theory*. London: Routledge.

Ellsworth, P. C. and Scherer, K. R. (2003) “Appraisal Processes in Emotion,” in R. J. Davidson, K. R. Scherer and H. H. Goldsmith (eds.), *Handbook of Affective Sciences*, Oxford: Oxford University Press.

Elfenbein, H. A. and Ambady, N. (2002) “On the Universality and Cultural Specificity of Emotion Recognition: A Meta-Analysis,” *Psychological Bulletin, 128*(2), 203–235.

Fernández-Dols, J. M. and Crivelli, C. (2013) “Emotion and Expression: Naturalistic Studies,” *Emotion Review, 5*(1), 24–29.

Fischer, M., Bossuyt, E., Köster, M., Buabang, E. and Moors (2022) “On the Interplay Between Stimulus-Driven and Goal-Directed Processes in the Decision to Fight or Flee,” Manuscript submitted for publication.

Fischer, M., Fini, C., Brass, M. and Moors, A. (2020) “Early Approach and Avoidance Tendencies can be Goal-Directed: Support from a Transcranial Magnetic Stimulation Study,” *Cognitive, Affective, and Behavioral Neuroscience*, *20*, 648-657.

Fischer, M., Kuppens, P. and Moors, A. (2020) “When Socially Excluded People Prefer Moralizing to Anti- and Prosocial Behavior: Support for a Goal-Directed Account,” *Motivation and Emotion*, *44*, 508-524.

Fridlund, A. J. (1994) *Human Facial Expression: An Evolutionary View*. San Diego: Academic Press.

---- (2017) “The Behavioral Ecology View of Facial Displays, 25 Years Later,” in J. M. Fernández-Dols & J. A. Russell (eds.), *The Science of Facial Expression*, Oxford: Oxford University Press.

Frijda, N. H. (1986) *The Emotions*. Cambridge: Cambridge University Press.

---- (2007) *The Laws of Emotion.* Erlbaum.

Frijda, N. H. and Zeelenberg, M. (2001) "What Is the Dependent?" in K. R. Scherer, A. Schorr and T. Johnstone (eds.), *Appraisal Processes in Emotion: Theory, Methods, Research*, Oxford: Oxford University Press.

Galati, D., Scherer, K. R. and Ricci-Bitti, P. E. (1997) "Voluntary Facial Expression of Emotion: Comparing Congenitally Blind with Normally Sighted Encoders," *Journal of Personality and Social Psychology, 73*(6), 1363-1379.

Gendron, M., Roberson, D., van der Vyver, J. M. and Barrett, L. F. (2014b) "Perceptions of Emotion from Facial Expressions Are Not Culturally Universal: Evidence from a Remote Culture," *Emotion, 14*(2), 251–262.

Gonzaga, G. C., Turner, R. A., Keltner, D., Campos, B. and Altemus, M. (2006) "Romantic Love and Sexual Desire in Close Relationships," *Emotion, 6*(2), 163–179.

Goudbeek, M. and Scherer, K. (2010) "Beyond Arousal: Valence and Potency/Control Cues in the Vocal Expression of Emotion," *The Journal of the Acoustical Society of America, 128*(3), 1322-1336.

Green, O. H. (1992) *The Emotions: A Philosophical Theory*. Kluwer.

Griffiths, P. E. (2004) "Is Emotion a Natural Kind?" in R. C. Solomon (ed.), *Thinking About Feeling: Contemporary Philosophers on Emotions*, Oxford: Oxford University Press.

Hatfield, E., Cacioppo, J. T. and Rapson, R. L. (1994) *Emotional Contagion?* Cambridge: Cambridge University Press.

Heyes, C. and Dickinson, A. (1990) "The Intentionality of Animal Action," *Mind and Language, 5*(1), 87-103.

Izard, C. E. (2011) "Forms and Functions of Emotions: Matters of Emotion–Cognition Interactions," *Emotion Review, 3*(4), 371-378.

Johansen, J. P., Cain, C. K., Ostroff, L. E. and LeDoux, J. E. (2011) "Molecular Mechanisms of Fear Learning and Memory," *Cell, 147*(3), 509-524.

Keltner, D. and Haidt, J. (1999) "Social Functions of Emotions at Four Levels of Analysis," *Cognition and Emotion, 13*(5), 505-521.

Keltner, D., Tracy, J., Sauter, D. A., Cordaro, D. C. and McNeil G. (2016) "Expression of Emotion," in L. F. Barrett, M. Lewis, & J. M. Haviland-Jones (eds.), *Handbook of Emotions.* 4th ed, New York: Guilford.

Kreibig, S. D. (2010) "Autonomic Nervous Activity in Emotion: A Review," *Biological Psychology, 84*(3), 394-421.

Kreibig, S. D., Gendolla, G. H. and Scherer, K. R. (2012) "Goal Relevance and Goal Conduciveness Appraisals Lead to Differential Autonomic Reactivity in Emotional Responding to Performance Feedback," *Biological Psychology, 91*(3), 365-375.

Kunst-Wilson, W. R. and Zajonc, R. B. (1980) "Affective Discrimination of Stimuli that Cannot be Recognized," *Science, 207*(4430), 557-558.

Lang, P. J. (1994) "The Motivational Organization of Emotion: Affect-Reflex Connections," in S. H. M. van Goozen, N. E. Van de Poll and J. A. Sergeant (eds.), *Emotions: Essays on Emotion Theory,* New York: Psychology Press.

Lange, J., Dalege, J., Borsboom, D., van Kleef, G. A. and Fischer, A. H. (2020) "Toward an Integrative Psychometric Model of Emotions," *Perspectives on Psychological Science, 15*(2), 444-468.

Larsen, J. T., Berntson, G. G., Poehlmann, K. M., Ito, T. A. and Cacioppo, J. T. (2008) "The Psychophysiology of Emotion," in M. Lewis, J. M. Haviland-Jones and L. F. Barrett (eds.), *Handbook of Emotions,* New York: Guilford.

Laukka, P. and Elfenbein, H. A. (2012) "Emotion Appraisal Dimensions can be Inferred from Vocal Expressions," *Social Psychological and Personality Science, 3*(5), 529-536.

Lazarus, R. S. (1991) *Emotion and Adaptation,* Oxford: Oxford University Press.

LeDoux, J. E. (1996) *The Emotional Brain: The Mysterious Underpinnings of Emotional Life*, New York: Simon & Schuster.

Levenson, R. W. (2011) "Basic Emotion Questions," *Emotion Review*, *3*(4), 379-386.

---- (2014) "The Autonomic Nervous System and Emotion," *Emotion Review*, *6*(2), 100-112.

Leventhal, H. (1984) "A Perceptual-Motor Theory of Emotion," in L. Berkowitz (ed.), *Advances in Experimental Social Psychology*. vol. 17, San Diego: Academic Press.

Lewis, M. D. (2005) "Bridging Emotion Theory and Neurobiology Through Dynamic Systems Modeling," *Behavioral and Brain Sciences*, *28*(2), 169-245.

Lindquist, K. A. and Barrett, L. F. (2008) "Constructing Emotion: The Experience of Fear as a Conceptual Act," *Psychological Science*, *19*(9), 898-903.

Lindquist, K. A., Barrett, L. F., Bliss-Moreau, E. and Russell, J. A. (2006). "Language and the Perception of Emotion," *Emotion*, *6*(1), 125-138.

MacCormack, J. K. and Lindquist, K. A. (2017) "Bodily Contributions to Emotion: Schachter's Legacy for a Psychological Constructionist View on Emotion," *Emotion Review*, *9*(1), 36-45.

MacLeod, C., Grafton, B. and Notebaert, L. (2019) "Anxiety-Linked Attentional Bias: Is It Reliable?" *Annual Review of Clinical Psychology*, *15*, 529-554.

Manstead, A. S. and Fischer, A. H. (2001) "Social Appraisal," *Appraisal Processes in Emotion: Theory, Methods, Research,* Oxford: Oxford University Press.

Markus, H. R. and Kitayama, S. (1991) "Culture and the Self: Implications for Cognition, Emotion, and Motivation," *Psychological Review*, *98*(2), 224–253.

Matsumoto, D., Keltner, D., Shiota, M. N., O'Sullivan, M. and Frank, M. (2008) "Facial Expressions of Emotion," in M. Lewis, J. M. Haviland-Jones and L. F. Barrett (eds.), *Handbook of Emotions*, New York: Guilford.

Matsumoto, D. and Willingham, B. (2009) “Spontaneous Facial Expressions of Emotion of Congenitally and Noncongenitally Blind Individuals,” *Journal of Personality and Social Psychology*, *96*(1), 1-10.

Mehu, M. and Scherer, K. R. (2015) “Emotion Categories and Dimensions in the Facial Communication of Affect: An Integrated Approach,” *Emotion, 15*(6), 798–811.

Mesquita, B. and Ellsworth, P. C. (2001) “The Role of Culture in Appraisal,” in K. R. Scherer, A. Schorr and T. Johnstone (eds.), *Appraisal Processes in Emotion: Theory, Methods, Research*, Oxford: Oxford University Press.

Mesquita, B. and Parkinson, B. (this volume) “Social Constructionism,” in A. Scarantino (ed.), *Routlegde Handbook of Emotion Theory*, London: Routledge.

Moors, A. (2007) “Can Cognitive Methods Be Used to Study the Unique Aspect of Emotion: An Appraisal Theorist's Answer,” *Cognition and Emotion*, *21*(6), 1238-1269.

---- (2013) “On the Causal Role of Appraisal in Emotion,” *Emotion Review, 5*(2), 132-140.

---- (2014) “Flavors of Appraisal Theories of Emotion,” *Emotion Review, 4*, 303-307.

---- (2016) “Automaticity: Componential, Causal, and Mechanistic Explanations,” *Annual Review of Psychology*, *67*, 263-287.

---- (2017a) “Integration of Two Skeptical Emotion Theories: Dimensional Appraisal Theory and Russell's Psychological Construction Theory,” *Psychological Inquiry, 28*(1), 1-19.

---- (2017b) “The Integrated Theory of Emotional Behavior Follows a Radically Goal-Directed Approach,” *Psychological Inquiry*, *28*(1), 68-75.

---- (2022) *Demystifying Emotions: A Typology of Theories in Psychology and Philosophy,* Cambridge: Cambridge University Press.

Moors, A., Boddez, Y. and De Houwer, J. (2017) “The Power of Goal-Directed Processes in the Causation of Emotional and Other Actions,” *Emotion Review*, *9*(4), 310-318.

Moors, A., Ellsworth, P., Scherer, K. R. and Frijda, N. H. (2013) "Appraisal Theories of Emotion: State of the Art and Future Development," *Emotion Review*, *5*(2), 119-124.

Moors, A., Fini, C., Everaert, T., Bardi, L., Bossuyt, E., Kuppens, P. and Brass, M. (2019) "The Role of Stimulus-Driven versus Goal-Directed Processes in Fight and Flight Tendencies Measured with Motor Evoked Potentials Induced by Transcranial Magnetic Stimulation," *PLoS one*, *14*(5), e0217266, 1-22.

Moors, A. and Scherer, K. R. (2013) "The Role of Appraisal in Emotion," in M. Robinson, E. Watkins and E. Harmon-Jones (eds.), *Handbook of Cognition and Emotion*, New York: Guilford.

Murphy, F. C., Nimmo-Smith, I. A. N. and Lawrence, A. D. (2003) "Functional Neuroanatomy of Emotions: A Meta-Analysis," *Cognitive, Affective, & Behavioral Neuroscience*, *3*(3), 207-233.

Murphy, S. T. and Zajonc, R. B. (1993) "Affect, Cognition, and Awareness: Affective Priming with Optimal and Suboptimal Stimulus Exposures," *Journal of Personality and Social Psychology*, *64*(5), 723–739.

Nelson, N. L. and Russell, J. A. (2013) "Universality Revisited," *Emotion Review*, *5*(1), 8-15.

Niedenthal, P. M., Krauth-Gruber, S. and Ric, F. (2006) *Psychology of Emotion: Interpersonal, Experiential, and Cognitive Approaches*, New York: Psychology Press.

Oatley, K. and Johnson-Laird, P. N. (1987) "Towards a Cognitive Theory of Emotions," *Cognition and Emotion*, *1*(1), 29-50.

Öhman, A. and Mineka, S. (2001) "Fears, Phobias, and Preparedness: Toward an Evolved Module of Fear and Fear Learning," *Psychological Review, 108*(3), 483–522.

Ortony, A. and Turner, T. J. (1990) "What's Basic About Basic Emotions?" *Psychological Review*, *97*(3), 315–331.

Parkinson, B. (1997) “Untangling the Appraisal-Emotion Connection,” *Personality and Social Psychology Review, 1*(1), 62-79.

---- (2001) “Anger on and off the Road,” *British Journal of Psychology, 92*(3), 507-526.

---- (2013) “Contextualizing Facial Activity,” *Emotion Review 5*(1), 97-103.

Parkinson, B. and Manstead, A. S. (2015) “Current Emotion Research in Social Psychology: Thinking About Emotions and Other People,” *Emotion Review*, *7*(4), 371-380.

Parrott, G. W. (2007) “Components and the Definition of Emotion,” *Social Science Information, 46*, 419-423.

Phan, K. L., Wager, T., Taylor, S. F. and Liberzon, I. (2002) “Functional Neuroanatomy of Emotion: A Meta-Analysis of Emotion Activation Studies in PET and fMRI,” *Neuroimage, 16*(2), 331-348.

Pessoa, L. and Adolphs, R. (2010) “Emotion Processing and the Amygdala: From a ‘Low Road’ to ‘Many Roads’ of Evaluating Biological Significance,” *Nature Reviews Neuroscience, 11*(11), 773-782.

Quigley, K. S. and Barrett, L. F. (2014) “Is There Consistency and Specificity of Autonomic Changes During Emotional Episodes? Guidance From the Conceptual Act Theory and Psychophysiology,” *Biological Psychology*, *98*, 82-94.

Reisenzein, R., Studtmann, M. and Horstmann, G. (2013) “Coherence between Emotion and Facial Expression: Evidence from Laboratory Experiments,” *Emotion Review*, *5*(1), 16-23.

Roseman, I. J. (2013) “Appraisal in the Emotion System: Coherence in Strategies for Coping,” *Emotion Review, 5*, 141-149.

Roseman, I. J., Spindel, M. S. and Jose, P. E. (1990) “Appraisals of Emotion-Eliciting Events: Testing a Theory of Discrete Emotions,” *Journal of Personality and Social Psychology*, *59*(5), 899-915.

Russell, J. A. (1994) “Is There Universal Recognition of Emotion from Facial Expression? A Review of the Cross-Cultural Studies,” *Psychological Bulletin, 115*(1), 102–141.

---- (2003) “Core Affect and the Psychological Construction of Emotion,” *Psychological Review*, *110*(1), 145–172.

---- (2005) “Emotion in Human Consciousness is Built on Core Affect,” *Journal of Consciousness Studies*, *12*(8-9), 26-42.

---- (2009) “Emotion, Core Affect, and Psychological Construction,” *Cognition and Emotion*, *23*(7), 1259-1283.

---- (2014) “My Psychological Constructionist Perspective, With a Focus on Conscious Affective Experience,” in L. F. Barrett and J. A. Russell (eds.), *The Psychological Construction of Emotion*, New York: Guilford.

Russell, J. A., Bachorowski, J. and Fernández-Dols, J. M. (2003) “Facial and Vocal Expressions of Emotion,” *Annual Review of Psychology, 54*, 329–349.

Saarimäki, H., Gotsopoulos, A., Jääskeläinen, I. P., Lampinen, J., Vuilleumier, P., Hari, R., Sams, M. and Nummenmaa, L. (2016) “Discrete Neural Signatures of Basic Emotions,” *Cerebral Cortex, 26*(6), 2563-2573. https://doi.org/10.1093/cercor/bhv086

Sauter, D. A. and Russell, J. A. (this volume) “What Do Nonverbal Expressions Tell Us About Emotion,” in A. Scarantino (ed.), *Routlegde Handbook of Emotion Theory*, London: Routledge.

Scarantino, A. (2012) “How to Define Emotions Scientifically,” *Emotion Review, 4*(4), 358–368.

Schachter, S. (1964) “The Interaction of Cognitive and Physiological Determinants of Emotional State,” in L. Berkowitz (ed.), *Advances in Experimental Social Psychology*. Vol. 1, Academic Press. https://doi.org/10.1016/S0065-2601(08)60048-9

Schachter, S. and Singer, J. E. (1962) “Cognitive, Social, and Physiological Determinants of an Emotional State,” *Psychological Review*, *69*(5), 379–399.

Scherer, K. R. (2009) “Emotions are Emergent Processes: They Require a Dynamic Computational Architecture,” *Philosophical Transactions of the Royal Society: B, 364,* 3459–3474.

Scherer, K. R. and Moors, A. (2019) “The Emotion Process: Event Appraisal and Component Differentiation,” *Annual Review of Psychology*, *70*, 719-745.

Scherer, K. R., Mortillaro, M., Rotondi, I., Sergi, I. and Trznadel, S. (2018) “Appraisal-Driven Facial Actions as Building Blocks for Emotion Inference,” *Journal of Personality and Social Psychology*, 114(3), 358–379.

Shariff, A. F. and Tracy, J. L. (2011) “Emotion Expressions: On Signals, Symbols, and Spandrels—A response to Barrett (2011),” *Current Directions in Psychological Science*, *20*(6), 407-408.

Shiota, M. N. (this volume) “Basic and Discrete Emotion Theories,” in A. Scarantino (ed.), *Routlegde Handbook of Emotion Theory*, London: Routledge.

Smith, C. A. and Kirby, L. D. (2001) “Toward Delivering on the Promise of Appraisal Theory,” in K. R. Scherer, A. Schorr and T. Johnstone (eds.), *Appraisal Processes in Emotion: Theory, Methods, Research,* Oxford: Oxford University Press.

Storbeck, J., Robinson, M. D. and McCourt, M. E. (2006) “Semantic Processing Precedes Affect Retrieval: The Neurological Case for Cognitive Primacy in Visual Processing,” *Review of General Psychology*, *10*(1), 41-55.

Strack, F., Martin, L. L. and Stepper, S. (1988) “Inhibiting and Facilitating Conditions of the Human Smile: A Nonobtrusive Test of the Facial Feedback Hypothesis,” *Journal of Personality and Social Psychology*, *54*(5), 768–777

Swartz, N. (1997) "*Definitions, Dictionnaries, and Meanings.* Retrieved from http://www.sfu.ca/~swartz/definitions.htm

Teachman, B. A., Clerkin, E. M., Cunningham, W. A., Dreyer-Oren, S. and Werntz, A. (2019) "Implicit Cognition and Psychopathology: Looking Back and Looking Forward," *Annual Review of Clinical Psychology*, *15*, 123-148.

Tracy, J. L. and Randles, D. (2011) "Four Models of Basic Emotions: A Review of Ekman and Cordaro, Izard, Levenson, and Panksepp and Watt," *Emotion Review*, *3*(4), 397-405.

Tracy, J. L., Shariff, A. F. and Cheng, J. T. (2010) "A Naturalist's View of Pride," *Emotion Review*, *2*(2), 163–177.

Tracy, J. L., Shariff, A. F. and Zhao, W. and Henrich, J. (2013) "Cross-Cultural Evidence That the Nonverbal Expression of Pride Is an Automatic Status Signal," *Journal of Experimental Psychology: General, 142*(1), 163–180.

Vervliet, B. and Boddez, Y. (2020) "Memories of 100 Years of Human Fear Conditioning Research and Expectations for Its Future," *Behaviour Research and Therapy, 135*, 103732.

Wagenmakers, E. J., Beek, T., Dijkhoff, L., Gronau, Q. F., Acosta, A., Adams Jr, R. B., ...and Zwaan, R. A. (2016) "Registered Replication Report: Strack, Martin and Stepper (1988)," *Perspectives on Psychological Science*, *11*(6), 917-928.

Wilson-Mendenhall, C. D., Barrett, L. F., Simmons, W. K. and Barsalou, L. W. (2011) "Grounding Emotion in Situated Conceptualization," *Neuropsychologia*, *49*(5), 1105-1127.

[1] Preparation of this chapter was supported by the Research Foundation - Flanders (FWO) (G022313N and G073317N) and the Research Fund of KU Leuven (C14/17/047).

[2] Note that network theories are one-factor theories: Emotions are caused by the network or script, but this script is not used as a source of information to be combined with core affect.

[3] It may also be recalled that some theorists have abondoned the ambition to provide a classic intensional definition and have turned to a cluster-type definition instead in which conditions are not strictly necessary but rather typical. Thus, several individual theorists would in fact tone down the definitions that I attribute to their theory.

[4] Note that some theorists have proposed treating control precedence (Frijda 1986, 2007) and synchronicity (Scherer 2000) as scientific criteria for demarcation of the set of emotions. In an earlier section, however, I treated control precedence as part of the working definition of emotion, and the same might be done for synchronicity.