Review Article

Brain and Neuroscience Advances

Brain and Neuroscience Advances
Volume 2: 1–8

DOI: 10.1177/2398212818812628
journals.sagepub.com/home/bna
SAGE

Perspectives from affective science on understanding the nature of emotion

Elaine Fox

Abstract

Emotions are at the heart of how we understand the human mind and of our relationships within the social world. Yet, there is still no scientific consensus on the fundamental nature of emotion. A central quest within the discipline of affective science is to develop an in-depth understanding of emotions, moods, and feelings and how they are embodied within the brain (affective neuroscience). This article provides a brief overview of the scientific study of emotion with a particular emphasis on psychological and neuroscientific perspectives. Following a selective snapshot of past and present research in this field, some current challenges and controversies in affective science are highlighted.

Keywords

Emotion, mood, affect, affective neuroscience, emotion science

Received: 12 February 2018

Contemporary neuroscience and psychological science has witnessed an explosion of research and theoretical developments on emotion, addressing topics as diverse (and important) as what emotions are, where they come from, how they are consciously experienced, and how they are implemented in the brain and body. Yet, there is still little consensus within the field of affective science as to the fundamental nature of emotion. Two broad perspectives can be identified. First, many researchers assume that emotions are biologically given action plans that help humans and other animals to navigate the complexities of the world and can be thought of as ‘decoupled reflexes’ (Adolphs and Anderson, 2018). Following Barrett (2006a), this can be called a ‘natural-kind’ view of emotion. A very different perspective, however, hypothesises that emotions are primarily social constructions that emerge from a dynamic brain organisation alongside a highly developed conceptual system that helps to make sense of incoming sensory information. From this perspective, emotions are not reactions to sensory events, but rather they are conceptual constructions of the world (Barrett, 2017a). This can be called the ‘conceptual construction’ view of emotion.

Addressing the implications of these contrasting perspectives allows affective scientists to consider (or re-consider) the fundamental nature of what it is that affective science needs to explain (see Adolphs (2017a, 2017b) and Barrett (2017b, 2017c) for a recent debate on these contrasting viewpoints). This is an exciting time for affective science with an abundance of ongoing empirical and theoretical developments. Indeed, some emotion theorists (Barrett, 2017b) have claimed that contemporary discoveries in neuroscience have brought us to the brink of a paradigm shift in terms of understanding, not only how the brain works but also how emotions are to be understood. In this article, I will attempt to provide a flavour of past and present research in affective science, and to introduce readers to the current challenges facing this dynamic field.

The study of emotion has a long history

For generations, a dominant theme in Western philosophy was that emotions, or ‘passions’ as they were called, were base and destructive and actively opposed the loftier processes of ‘reason’ (Solomon, 1993). Reason was seen as the ultimate human virtue and the emotions were considered to be antithetical to this, often serving to undermine the pursuit of a rational way of life. This notion goes back at least to Plato (2003 [375 BC]) and has led to a lingering distrust of emotions in Western thought. Even Charles Darwin, who wrote a seminal book on emotions (Darwin, 2009 [1872]), reflected this general distrust of our emotional life and implied that emotions were the vestiges of our evolution from lower life forms. Darwin writes, ‘our descent then, is the origin of our evil passions’ (cited in Gruber and Barrett (1974: 289)).

Contemporary science espouses a very different view of emotions, and the natural-kind view of emotion, which has been dominant in affective science until recently, has emphasised the crucial

Department of Experimental Psychology, University of Oxford, Oxford, UK

Corresponding author:
Elaine Fox, Department of Experimental Psychology, University of Oxford, New Radcliffe House, 49 Walton Street, Oxford OX2 6AE, UK.
Email: elaine.fox@psy.ox.ac.uk

Table 1. Some proposed distinctions among affective phenomena that have been studied under the rubric of 'affect' within affective science and affective neuroscience.

English term	Typical duration	Common elicitor	Function
Emotions	Seconds to minutes	Specific object or situation	To bias actions
Moods	Hours, days, weeks	General internal and external environment/milieu	To bias cognition
Feelings	Variable, but probably relatively short-term	Activation of internal milieu (e.g. interoceptive system)	To bring to conscious awareness
Emotion concept	Variable, but probably very long-term	General environment and semantic system	To inform rational discussion and decision-making

role that emotions play in tuning us in to the most salient events going on around us; influencing our decision-making, magnifying our memory for important past events, and colouring our perception of the future. Emotions from this perspective help to signal, both internally and externally, our deepest motivations and priorities. Emotions prompt adaptive actions in those experiencing the emotion while the perception of emotional expressions in others can elicit adaptive behaviour on the part of the perceiver. In this way, 'emotions are controlled and shaped by these social situations; and they, in turn, serve to shape social situations' (Niedenthal et al., 2006: 3).

While emotions and their associated feelings, such as sadness, fear, happiness, disgust, pride, gratitude, and curiosity, are all part of everyday life, they can also become destructive in a wide range of affective disorders such as anxiety and depression. It is understandable, then, that the attempt to understand emotional responsiveness has been a central objective for psychological science and for neuroscience. In spite of all of this scholarly activity across many disciplines, however, progress in developing a scientific understanding of emotion has been relatively slow in comparison to other domains such as perception, attention, and memory. There are several potential reasons for this and just some of these are heightened in this article.

I will not attempt to present an extensive review of affective science, as several comprehensive overviews are available (e.g. Adolphs and Anderson, 2018; Barrett et al., 2016; Davidson et al., 2003; Fox, 2008; Gross, 2007; Keltner et al., 2013; Niedenthal et al., 2006; Panksepp, 1998; Pessoa, 2013). Instead, I will highlight some key challenges facing affective science and outline some potential ways forward for theoretical and empirical development.

Challenges facing a science of emotion

Progress in developing a comprehensive understanding of emotion and emotional disorders has been hampered in my view by three broad issues:

1. There are ongoing disagreements as to how emotion should be defined. This seems important to resolve.
2. Affective science is a broad multidisciplinary field, and the phenomena of interest are investigated at many different levels of analysis (e.g. from recording of single cells to asking people how they feel) with a variety of different methodologies. This problem is not unrelated to definitional problems and can lead to confusion and difficulties in integrating across different sub-disciplines.
3. Affective scientists often hold strong pre-existing assumptions about the nature of emotion (e.g. 'emotions are conceptual constructions' or 'emotions are natural-kinds'), and this colours the perception and interpretation of scientific data. The influence of the observer/scientist is a problem for all of science, of course, but it has led to particularly difficult problems in affective science in terms of agreeing on exactly what it is that we are measuring.

How should we define what we mean by different affective phenomena?

The scientific investigation of emotion has been hampered by the fact that many distinct phenomena are often studied under the same linguistic rubric. Several different words, such as emotion, mood, or feeling, have been used interchangeably by researchers, who often work in different sub-disciplines of the field, to refer to the same basic phenomena. Conversely, many research programmes investigate the same affective phenomenon, but use quite different terminology. This has led to an inevitable degree of confusion.

There have, of course, been several attempts to establish some agreed definitions, but an agreed terminology has proven difficult to establish. In my view, a useful distinction that can be made, which has found some degree of consensus, is to separate those components of affect that correspond to the English terms 'emotions', 'moods', 'feelings', and 'concepts' (Fox, 2008). As shown in Table 1, the term 'emotion' is reserved within this framework, to refer to a relatively short-lived (seconds to minutes) state of body and mind/brain that is a reaction to a specific object or situation with a primary function of biasing action tendencies. This is the meaning that has been utilised traditionally by 'emotions as natural-kinds' approaches (e.g. Izard, 2007; Panksepp, 1998), but as will be discussed later, may not be inconsistent with approaches that take an 'emotions as conceptual constructions' approach (e.g. Barrett, 2006a). In contrast, 'mood states' are considered to be longer lasting, are typically not linked with specific events, and tend to bias cognitions rather than actions. It also seems to be important to separate the notions of 'emotions' and 'moods' from the conscious experience (feelings) of these affective states. Within this framework, affect is reserved as an overarching term used to refer to responses related to emotions, moods, and the regulation of these states (see Fox (2008), Chapter 2, for further discussion). More recently, it has also been argued that the ability to think about emotions (i.e. concepts of emotion) without necessarily experiencing these affective states (Adolphs, 2017a) should be distinguished. Table 1 outlines some commonly used terminologies that have been used within affective science and affective neuroscience alongside the typical time frame of these phenomena with a suggestion of their primary functions and common elicitors. These are inevitably speculative and many key questions remain unanswered (see Adolphs and Anderson (2018), Barrett (2017b), and Fox (2008) for further discussion).

LeDoux (2012) has considered the complications that arise for a neuroscience of emotion when trying to interpret research on 'emotion' in the absence of an agreed understanding of what is meant by emotion. In a proposed comprehensive framework, he points out that many behavioural phenomena occur when an organism detects and responds to opportunities and challenges in everyday life. These manifestations reflect fundamental functions and neural circuits that are related to survival, and the recasting of these survival circuits as 'emotions' and 'feelings' has led to tremendous confusion in the understanding of affective processes. LeDoux (2012) does not attempt to define emotion, but argues that a broader framework for investigating critical phenomena related to survival functions will free researchers to think about affective processes without having to worry about an exact definition of what they mean by the term emotion. This evolutionary perspective does provide a useful framework in which to investigate emotion but also conflates proximal goals with more distal goals (Adolphs, 2017a), which can reduce the explanatory power of the concept of emotion to understand behaviour.

Of course it seems reasonable to assume that the entire affective system serves survival as a primary goal or driver, but there are also many proximal goals in life (find food, find a job, find a partner, and avoid immediate dangers) that are facilitated by affective phenomena, so some more detailed understanding (and definitions) of emotion would seem to be useful.

Some implications of using different methodologies in affective science

Moving beyond definitional problems, interpreting research on affective phenomena is further complicated by the plethora of research methodologies that have been and are being utilised to investigate the various phenomena of interest. Take the emotion of *fear* as a prototypical example. Much of what we know about fear comes from electrodes surgically implanted in the brains of rats that measure individual neurones when the animal is exposed to a fearful stimulus (e.g. a predator), or using classical fear conditioning where animals are trained to associate certain sensory stimuli with aversive outcomes (Blanchard and Blanchard, 1972). This large and systematic body of work has led to the widespread assumption that a sub-cortical structure – the amygdala – is central for the perception of fear as well as for the orchestration of fear responses (LeDoux, 2000).

In human research, fear-conditioning paradigms have also highlighted the role of the amygdala in the development of changes in patterns of neural reactivity to salient events (Calder et al., 2001; Critchley et al., 2002; LeDoux, 2000). Fear conditioning studies have further revealed how learned fear associations can result in different patterns of attentional biases towards certain environmental cues and how individual differences (e.g. in anxiety levels) can influence the development of these associations (e.g. Fox et al., 2012). Yet, other research traditions use neuropsychological techniques in order to assess emotion processing in patients with various forms of brain damage (especially in relation to the amygdala and related structures: Calder et al. (2001)) while research using functional magnetic resonance imaging (fMRI) with healthy participants with intact brains has also begun to map out the neural networks that are involved in fear perception and emotion processing more generally (Lindquist et al., 2012; Vuilleumier, 2005).

Yet, other research traditions investigate the cognitive and neural aspects of affective disorders, such as anxiety and depression (Fox, 2008; Mathews and MacLeod, 2005) while some affective scientists focus on the nature of consciously felt affective experiences (Barrett and Russell, 1999). It is no surprise, then, that integrating the many facts that are known about emotions, moods, and feelings from across these disparate research traditions represents a real challenge to the scientific understanding of affect.

The problem of the observer in affective science

As pointed out by Barrett (2017c), a scientist's deepest held assumptions about the nature of the phenomena under study acts as a guide to what is a 'signal' and what is 'noise' within the data. These pre-existing assumptions determine not only what type of methodology is considered most appropriate but also how different data sets should be interpreted. While the often unseen influence of the observer is an issue that all of the sciences have to acknowledge, it has led to some fundamental assumptions within affective science that are often taken as given, rather than as hypotheses to be tested (see Barrett (2006a, 2017c) for further discussion).

For instance, two long-standing hypotheses within the scientific investigation of emotions, that are sometimes assumed to be statements of fact, rather than hypotheses to be tested, have recently been challenged robustly by the conceptual construction approach to emotion (Barrett, 2006a, 2017a, 2017b). The first relates to the notion that there are biologically given emotion circuits within the brain and the second refers to the assumption that emotions represent the flexible coordination of different physiological and behavioural components in the service of harnessing action plans to deal with opportunities and threats.

In the following sections, I will briefly review these assumptions before considering ways in which a range of empirical findings and theoretical assumptions might be reconciled.

Are there innate emotion circuits within the brain?

A common hypothesis that has been influenced heavily by neuroscience proposes that there are innate emotion circuits within the brain. This 'emotions as natural-kinds' perspective suggests that nature has endowed us with a set of 'basic' emotions, or emotional systems, that are universally recognisable and experienced, and that are triggered by distinct neural circuits (e.g. Ekman, 1992; Izard, 2007; Panksepp, 1998, 2000). Thus, emotions such as anger, fear, sadness, disgust, and happiness, which are easily recognisable, are often considered to be basic in the sense of being 'natural-kinds' that exist in nature and are independent of our perception of them. Panksepp (1998), for instance, has summarised decades of behavioural neuroscience research primarily with rodents and has concluded that there is evidence for the existence of seven primary emotional systems – SEEKING, RAGE, FEAR, LUST, CARE, PANIC, PLAY – that are ancient in evolutionary terms and operate in similar ways across species.

The veracity of the existence of a set of discrete and basic emotions has been questioned by research from a very different perspective that assumes that language plays a significant role in shaping our experience of emotion (Russell and Barrett, 1999). Originating from research on the subjective perception of emotional feelings, this research tradition has demonstrated that the experience of emotion is best described as varying along two dimensions that relate to arousal or activation (low to high intensity) on one hand and valence (unpleasant to pleasant) on the other. An emotion such as 'sadness', for instance, would be described as being towards the unpleasant end of the valence dimension and relatively low on the arousal dimension, while 'fear' would be similar on the valence dimension (unpleasant) but much higher in terms of its arousing qualities (Russell, 1980). In the 'conceptual act theory' (CAT) of emotion, Barrett has built upon this dimensional approach to propose that what we perceive and experience as discrete emotions (fear, sadness, and happiness), rather than being biologically given, are actually conceptual constructions that emerge from our categorisation of a more basic psychological process that she calls 'core affect' (Barrett, 2006b; Barrett and Russell, 1999).

The essence of the CAT is that our background core affect (or what we might consider to be more general mood states – see Table 1) is experienced along the general dimensions of valence and arousal, and that these states are then categorised into discrete emotion categories based on a cognitive appraisal of the current context. This psychological constructionist approach, which more recently has been recast as the 'theory of constructed emotion' (Barrett, 2017a, 2017b) assumes that emotional episodes are constructed from more basic psychological operations – that are not specific to emotion – by means of the brain making sense of and interpreting the sensory information that comes from the environment and from internal signals (Barrett, 2014).

Evidence for this approach has come from an extensive meta-analysis of neuroimaging evidence that found little evidence that discrete emotions are localised to distinct brain regions instead revealing that a range of interacting brain regions involved in multiple affective and non-affective processes are active during emotional experiences across a range of discrete emotion categories (Lindquist et al., 2012). This led Lindquist et al. (2012) to claim support for a psychological constructionist view of the mind. Not everyone agrees with this interpretation as discussed in the wide range of opinions expressed in the various commentaries following the target article (Lindquist et al., 2012). Moreover, as several authors point out, fMRI data do not have the resolution required to identify distinct emotion-specific brain regions even if they did exist (e.g. LeDoux, 2012). Nevertheless, this debate raises important topics within affective science and has sparked a new wave of neuroimaging and other researches into the fundamental nature of human emotions.

Do emotions involve coordinated neural, cognitive, and behavioural components?

Often related to the hypothesis of discrete emotions, a common assumption is that emotions involve the temporary coordination of changes among physiological, psychological, and behavioural components that are important for survival (e.g. LeDoux, 2012; Scherer, 2001). The hypothesis is that emotions are made up of several components that become momentarily synchronised in response to a major life challenge, or opportunity, that is highly relevant to a person's goals and aspirations. These components include feelings, actions, physiological changes, facial expressions, and cognitive appraisals. The 'component process' model of emotion assumes that these five components of an emotion are linked to underlying organismic sub-systems, each of which is associated with a set of core functions (Scherer, 2001). From this perspective, an emotion is considered to be 'an episode of interrelated, synchronized changes in the states of all or most of the five organismic subsystems in response to the evaluation of an external or internal stimulus event as relevant to major concerns of the organism' (Scherer, 2005: 697).

The inclusion of cognitive appraisal as an integral part of emotion is controversial since many theorists argue that the evaluative frameworks that people use to make sense of events in relation to personal well-being play a crucial role in the elicitation of emotional experiences (see Ellsworth and Scherer (2003) for overview). Magda Arnold (1960), for instance, proposed that organisms are constantly appraising their surroundings in relation to their personal well-being, determining whether significant stimuli are present or absent, potentially harmful or beneficial, and whether they are easy or difficult to approach and avoid. Lazarus (1966) distinguished between 'primary appraisal' which refers to the implications of an environmental event for one's well-being and 'secondary appraisal' which refers to a judgement about one's ability to cope with the situation at hand. Scherer's (2001) componential appraisal approach builds on these earlier theoretical frameworks and proposes that emotion episodes progress in a highly reactive way to environmental events that are driven by *parallel* appraisal processes that sequentially evaluate those unfolding events in terms of novelty, intrinsic pleasantness, and relevance for a person's goals. These appraisal processes are also considered to evaluate compatibility with social and personal norms as well as one's general ability to cope.

The question of whether cognitive appraisals might precede or operate in parallel to emotional responses fell out of a vigorous historical debate surrounding the question of whether 'emotion' and 'cognition' are independent processes that can operate independently of each other, which reached its zenith in the 1980s (see Lazarus (1982, 1984) and Zajonc (1980, 1984)). Lazarus and others argued that cognition always preceded emotions, whereas Zajonc (1980) argued for the primacy, or at least the independence, of emotional responses from cognitive processes. This debate reflected the clear separation between emotion and cognition that had lingered since Plato's contrast between the 'passions' and 'reason'.

There has been a dramatic shift in this perspective in contemporary affective science, however, with the realisation that cognitive and affective processes are so tightly integrated at both cognitive and neural levels that it is impossible to separate them. This contemporary view has been driven to a large extent by the investigation of affective processes at the neural level. The question of how emotion is embodied in the brain has a long history (Dalgleish, 2004), and the history of affective neuroscience (Davidson and Sutton, 1995) has demonstrated that multiple areas of the brain are involved in affective processing effectively de-bunking the notion that the 'limbic system' is the seat of emotions (LeDoux, 2000). Affective neuroscience has reframed our understanding of affective and cognitive processes in terms of how multiple interacting sub-systems underpin emotional responding rather than a framework in which 'affect' is simply pitted against 'cognition' (Lindquist et al., 2012; Ochsner and Gross, 2008). The architectural features of the brain allow for massively parallel processing at multiple levels. Extensive connections and communication take place between the visual cortex and the amygdala, for instance, so that visual processing is contextually informed by signals occurring across a wide range of brain regions as well as the amygdala (Pessoa and Adolphs, 2010). This implies that vision is always 'affective vision' even at the level of visual cortex (Pessoa, 2012). This brings us full circle from the 1940s and 1950s when the 'New Look' movement in perception proposed that most cognitions and perceptions are infused by affective significance (Bruner and Postman, 1949). Neuroscience techniques have helped to rediscover this framework and begun to delineate the multiple interacting sub-systems within the brain that underpin our emotional life.

Just as affective and cognitive processes are integrated seamlessly in everyday behaviour and experience, it seems to be the case at the neural level with no clear segregation of affective and cognitive processes. Thus, Davidson et al. (2003) conclude,

> It is simply not possible to identify regions of the brain devoted exclusively to affect or exclusively to cognition. This fact should dispel claims about their independence and help to foster a more nuanced appreciation of the ways in which affect and cognition interact. (p. 5)

Indeed, the contemporary consensus in neuroscience that the brain operates as a massive network whose modus operandi is by means of developing accurate predictions of the world outside of itself and uses 'prediction errors' to update its internal model of the external world (Friston, 2018) has led some emotion theorists to claim that emotions (or at least consciously felt experiences) are constructed by the brain in exactly the same way as other semantic concepts, such as money (Barrett, 2006a, 2017a, 2017b).

Future theoretical and empirical directions

The time now seems right to attempt to bring together the many empirical and theoretical insights from a wide range of research programmes in affective science in order to begin building a truly systematic science of emotion. Because of definitional problems with our core concepts: *emotions, moods, feelings, and so on*, alongside the multitude of methods that are commonplace in affective science, this is not an easy task. There are two broad areas, however, which seem especially important and where there appears to be real potential to make substantive progress. The first is the possibility of integrating the many insights that have been discovered in independent lines of research that have emerged as *emotions as natural-kinds* or 'discrete emotions' approaches (e.g. Izard, 2007; Panksepp, 1998) on one hand, in contrast to approaches that have taken an 'emotions as conceptual constructions' or 'dimensional' approaches (e.g. Barrett and Russell, 1999; Russell, 1980) on the other.

The second (and not unrelated) divergence in affective science relates to theoretical approaches that take a broad *functional* perspective (e.g. Adolphs, 2017b; Adolphs and Anderson, 2018; LeDoux, 2012) and those that take a *psychological constructionist* perspective (e.g. Barrett, 2017a, 2017b); it may not be possible to bring these approaches together. Instead, it may be the case that we are on the brink of a paradigm shift in our understanding of the brain, which will set the stage to transform our understanding of emotion as well as the way in which we conduct affective science (Barrett, 2017b). Following a brief consideration of whether discrete emotion and dimensional approaches can be integrated in some way, I will conclude with some thoughts on the developing theoretical landscape in affective science and neuroscience.

Can 'natural-kind' and 'conceptual construction' approaches to emotion be reconciled?

While discrete emotions and dimensional approaches have often been pitted against each other in affective science, their integration would seem to be necessary in order to explain the entirety of emotional expression and experience. Fox (2008) has pointed out that these perspectives have emerged from very different research traditions with those taking an 'emotions as natural-kinds' perspective typically utilising evidence from neuroscience studies in animals (e.g. Panksepp, 1998) and humans (e.g. Calder et al., 2001; Davidson and Sutton, 1995), or assessing whether different emotional states are associated with distinct neural or cognitive reactions to proposed eliciting stimuli, such as facial expressions (e.g. Ekman, 1992; Fox, 2002). In contrast, those taking 'emotions as conceptual constructions' approaches, have typically utilised evidence from the analysis of subjective experience of emotions or moods in humans usually by means of self-report (e.g. Barrett and Russell, 1999; Russell, 1980). However, there is also evidence from neuroimaging studies with healthy human participants that has been argued to support a dimensional approach (Lindquist et al., 2012).

It is perhaps unsurprising that these very different data sets and methodologies have led to very different perspectives on the potential shape of the affective system. Fox (2008) proposed a potential unifying framework to see whether dimensional and discrete approaches might be fruitfully combined (see Figure 1 for an updated version of this framework).

The essence of the proposed framework is to take the seven primary emotional systems, as identified by Panksepp (1998) on the basis of decades of behavioural neuroscience research, as exemplars of highly adaptive and narrowly focused ranges of responses that are activated rapidly and automatically in response to an affective stimulus or situation. Each of these primary emotion systems is assumed to mobilise action tendencies as well as having an impact on the background mood or core affect of the individual. My central hypothesis is, however, that these primary

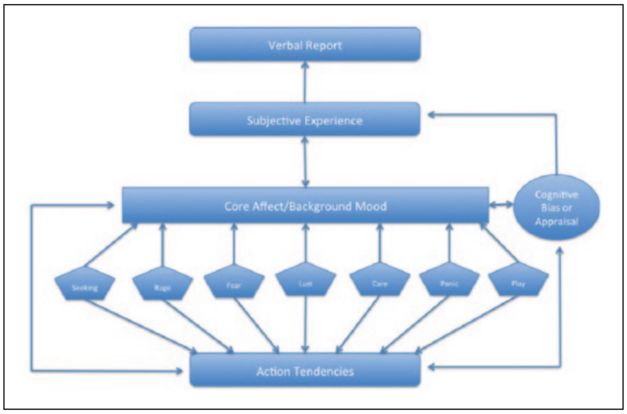

Figure 1. A framework for integrating discrete (basic) emotion approaches and dimensional approaches to the understanding of emotions and moods.
Adapted from Fox (2008: Figure 11.7, p. 372).

emotional systems are not *directly* accessible to conscious awareness. Instead, when activated, the hypothesis is that they lead to a perturbation in the background mood state (or core affective state) of the individual. Above a certain threshold, these perturbations will be experienced subjectively, but on the basis of a long research tradition (Russell, 1980) these feeling states will only be experienced along the general dimensions of arousal and valence. In this way, 'what we consciously experience and can report are broad affective categories, but underlying these are separate and ancient discrete emotion systems' (Fox, 2008: 355).

LeDoux's (2012) focus on survival circuits is similar in proposing that feelings may reflect states of consciousness when a 'global organismic state' – by which I assume he means something similar to a mood or core affective state – is represented in a cognitive workspace maintained in neocortical regions (e.g. Shallice, 1988). The idea is that various survival circuits relating to threat detection, feeding behaviour, and so on are unlikely to be directly associated with specific feelings, but nevertheless can influence subjective feeling states indirectly. These perspectives (Fox, 2008; LeDoux, 2012) imply that multiple levels of analysis will be essential for the progression of affective science. While commonplace in neuroscience (e.g. Albright et al., 2000) in spite of some attempts (e.g. Power and Dalgleish, 1997), multilevel models are still underdeveloped in the study of affective phenomena.

Should functional approaches to emotion be replaced by a psychological construction approach?

A recent debate between Lisa Feldman Barrett (2017b, 2017c) and Ralph Adolphs (2017a, 2017b) has neatly drawn the lines between traditional functional approaches to developing a scientific understanding of emotion and psychological constructionist approaches. These different perspectives provide a very different view of what a science of emotion might look like. Barrett (2017b) is clear that, in her view, affective science is on the brink of a major paradigm shift due to new developments in neuroscience that will transform our conception of how to study emotion scientifically. In brief, her proposal is that the appropriate place to start is with the structure and function of the brain and, from there, deduce what the biological basis of emotions might be. Adolphs (2017b) strongly disagrees with this arguing that this approach is 'squarely impossible' because he thinks that affective scientists must instead begin with the observation of behaviour and 'derive your categories' of emotion from there.

These seem to be very different questions. Barrett (2017b) is asking essentially how emotions are implemented in the brain, and to answer this, it seems a reasonable step to start with how the brain works. Once you do this, then predictive coding (Friston, 2018) and various neural mechanisms that underpin brain function, such as neural degeneracy, come in to play and begin to shape how we might think emotions are implemented in the brain. In contrast, Adolphs (2017a, 2017b) raises the rather different question of what are the categories of emotion? Different categories or dimensions of emotion are only of real interest, of course, if they have different functions, and so starting with behaviour makes sense from this perspective.

Like LeDoux (2012), Barrett (2017b) takes a long view and argues that the ultimate function of the brain is to regulate the internal milieu of the body so that an organism can 'grow, survive, and reproduce' (p. 3). This process of 'allostasis' therefore dictates everything that happens in the brain. Put simply, the proposal is that the brain creates a series of *concepts* so that incoming sensory information can be categorised in a meaningful way and therefore guides actions in a useful way. The definition of 'concept' that Barrett (2017b) uses, however, is very broad and seems to refer to a whole brain representation of the external environment (based on incoming sensory information and past experience) the main function of which is to *predict* what is about to happen, what is the best way to deal with the predicted events, and what are the implications of the predicted events for allostasis. The brain codes what actually happens and computes if this fits well with what was predicted. Following the principles of predictive coding (Friston, 2018), if there is a good match, the prediction becomes a perception or an experience whereas if not, a 'prediction error' is recorded so that this experience can be used to construct a concept so that future sensory information can be categorised efficiently and action plans can be rapidly implemented. From this perspective, a concept is essentially a prediction signal (Barrett, 2017b).

To recap, Barrett's (2017a, 2017b) proposal is that when the brain's internal model creates an emotion concept, the eventual categorisation of sensations results in a meaningful instance of emotion. This implies that emotion concepts are no different to other semantic concepts, such as money, in the sense that there are various different objects that have served as currency over the ages that have no particular physical similarities. In the same way, she argues, 'emotion categories don't have distinct, dedicated neural essences' (Barrett, 2017b: 13). The central assumption therefore is that 'emotions are constructions of the world, not reactions to it' (Barrett, 2017b: 16).

As outlined in Table 1, it seems to me to be useful to keep notions of emotion concepts – in the sense of a semantic concept – separate from the experience of a full-blown emotion episode. It seems that humans are well able to think and talk about emotions without necessarily experiencing an emotional episode in the sense of a temporary synchronisation of behavioural, subjective, and physiological components in relation to some event in the immediate environment (e.g. Scherer, 2005). However, in the context of Barrett's constructed emotion theory, a concept is redefined as a collection of whole body-brain representations that predicts what is about to happen. This broad redefinition of concept seems to lose much of the explanatory power of what it means to experience an emotion (see Adolphs (2017a, 2017b) for a similar point). For example, if the function of an emotion concept is to enable the categorisation of sensory information using synchronised body and brain states that occur in a specific context to predict what is likely to happen next and to mobilise action plans, then the definition of 'concept' in Barrett's terms seems very close to what other affective scientists – coming from a discrete emotions perspective – would call 'affective systems' (LeDoux, 2012; Panksepp, 1998) or 'emotions' (Izard, 2007; Scherer, 2005) or 'emotion states' (Adolphs, 2017a).

Following on from this, Adolphs (2017b) admits some confusion as to the relevance of predictive coding and the general workings of the brain as no-where in Barrett's (2017b) article could he find 'any criteria for what counts as an emotion' (p. 33). Perhaps, however, this is exactly Barrett's point. If I understand it correctly, the idea is that the brain works this way and constructs all sorts of perceptions and memories and experiences some of which we happen to call 'emotions'. But, the key is that they have no reality or exclusivity that sets them apart from other concepts.

The functional approach (Adolphs and Anderson, 2018) provides a contrasting view in proposing that in order to understand what an emotion is we have to begin with a *functional* definition of emotions. From this perspective, a functional account is the type of description that is necessary in order to judge whether a physiological or psychological state is an emotion or not. If it is an emotion, then a functional account is also needed to determine what specific emotion is being experienced. The suggestion that emotions have proximal functions as well as the more distal function of survival seems a compelling one (Adolphs, 2017b) to me, and highlights the practical and theoretical usefulness of functional approaches. However, Barrett (2017b) may be correct when she says 'functionalism cannot save the classical view of emotion'.

The jury is still out with regard to which of these approaches will stand the test of time, new data, and theoretical developments. However, this debate has re-vitalized affective science and it will be fascinating to see whether the concept of functionalist, discrete emotions will eventually be abandoned to reveal a new landscape of psychological constructionism.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author was funded by the European Research Council (ERC) Advanced Investigator Award under the European Union's Seventh Framework Programme (FP7/2007–2013)/ERC grant agreement no: 324176. The funding sources had no involvement in the writing of the report, or in the decision to submit the article for publication.

References

Adolphs R (2017a) How should neuroscience study emotions? By distinguishing emotion states, concepts, and experiences. *Social Cognitive and Affective Neuroscience* 12(1): 24–31.

Adolphs R (2017b) Reply to Barrett: Affective neuroscience needs objective criteria for emotions. *Social Cognitive and Affective Neuroscience* 12(1): 32–33.

Adolphs R and Anderson DJ (2018) *The Neuroscience of Emotion: A New Synthesis*. Princeton, NJ: Princeton University Press.

Albright TD, Jessell TM, Kandel ER, et al. (2000) Neural science: A century of progress and the mysteries that remain. *Neuron* 100 (Special Issue): 1–55.

Arnold MB (1960) *Emotion and Personality*. New York: Columbia University Press.

Barrett LF (2006a) Are emotions natural kinds? *Perspectives on Psychological Science* 1(1): 28–58.

Barrett LF (2006b) Solving the emotion paradox: Categorization and the experience of emotion. *Personality and Social Psychology Review* 10(1): 20–46.

Barrett LF (2014) The conceptual act theory: A précis. *Emotion Review* 6(4): 292–297.

Barrett LF (2017a) *How Emotions Are Made: The Secret Life of the Brain*. New York: Houghton Mifflin Harcourt.

Barrett LF (2017b) The theory of constructed emotion: An active inference account of interoception and categorization. *Social Cognitive and Affective Neuroscience* 12(1): 1–23.

Barrett LF (2017c) Functionalism cannot save the classical view of emotion. *Social Cognitive and Affective Neuroscience* 12(1): 34–36.

Barrett LF and Russell JA (1999) The structure of current affect: Controversies and emerging consensus. *Current Directions in Psychological Science* 8(1): 10–14.

Barrett LF, Lewis M and Haviland-Jones JM (eds) (2016) *Handbook of Emotions* (4th edn). New York: Guilford Press.

Blanchard DC and Blanchard RJ (1972) Innate and conditioned reactions to threat in rats with amygdaloid lesions. *Journal of Comparative and Physiological Psychology* 81(2): 281–290.

Bruner JS and Postman L (1949) On the perception of incongruity: A paradigm. *Journal of Personality* 18(2): 206–223.

Calder AJ, Lawrence AD and Young AW (2001) Neuropsychology of fear and loathing. *Nature Reviews Neuroscience* 2(5): 352–363.

Critchley HD, Mathias CJ and Dolan RJ (2002) Fear conditioning in humans: The influence of awareness and autonomic arousal on functional neuroanatomy. *Neuron* 33(4): 653–663.

Dalgleish T (2004) The emotional brain. *Nature Reviews Neuroscience* 5: 583–589.

Darwin C (2009 [1872]) *The Expression of the Emotions in Man and Animals*. London: Penguin Classics.

Davidson RJ and Sutton SK (1995) Affective neuroscience: The emergence of a discipline. *Current Opinion in Neurobiology* 5(2): 217–224.

Davidson RJ, Scherer KR and Goldsmith HH (2003) *Handbook of Affective Sciences*. New York: Oxford University Press.

Ekman P (1992) An argument for basic emotions. *Cognition and Emotion* 6(3–4): 169–200.

Ellsworth PC and Scherer KR (2003) Appraisal processes in emotion. In: Davidson RJ, Scherer KR and Goldsmith HH (eds) *Handbook of Affective Sciences*. New York: Oxford University Press, pp. 572–595.

Fox E (2002) Processing emotional facial expressions: The role of anxiety and awareness. *Cognitive, Affective, & Behavioral Neuroscience* 2(1): 52–63.

Fox E (2008) *Emotion Science: Cognitive and Neuroscientific Approaches to Understanding Human Emotions*. New York: Palgrave Macmillan.

Fox E, Yates A and Ashwin C (2012) Trait anxiety and perceptual load as determinants of emotion processing in a fear conditioning paradigm. *Emotion* 12(2): 236–249.

Friston K (2018) Does predictive coding have a future? *Nature Neuroscience* 21(8): 1019–1021.

Gross JJ (2007) *Handbook of Emotion Regulation*. New York, NY, US: Guilford Press.

Gruber HE and Barrett PH (1974) *Darwin on Man: A Psychological Study of Scientific Creativity* (together with Darwin's early and unpublished notebooks). New York: Dutton.

Izard CE (2007) Basic emotions, natural kinds, emotion schemas, and a new paradigm. *Perspectives on Psychological Science* 2(3): 260–280.

Keltner D, Oatley K and Jenkins JM (2013) *Understanding Emotions* (3rd edn). London: Wiley.

Lazarus RS (1966) *Psychological Stress and the Coping Process*. New York: McGraw-Hill.

Lazarus RS (1982) Thoughts on the relations between emotion and cognition. *American Psychologist* 37(9): 1019–1024.

Lazarus RS (1984) On the primacy of cognition. *American Psychologist* 39(2): 124–129.

LeDoux JE (2000) Emotion circuits in the brain. *Annual Review of Neuroscience* 23: 155–184.

LeDoux JE (2012) Rethinking the emotional brain. *Neuron* 73(4): 653–676.

Lindquist KA, Wager TD, Kober H, et al. (2012) The brain basis of emotion: A meta-analytic review. *Behavioral and Brain Sciences* 35(3): 121–143.

Mathews AM and MacLeod C (2005) Cognitive vulnerability to emotional disorders. *Annual Review of Clinical Psychology* 1: 167–195.

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

Ochsner KN and Gross JJ (2008) Cognitive emotion regulation: Insights from social cognitive and affective neuroscience. *Current Directions in Psychological Science* 17(2): 153–158.

Panksepp J (1998) *Affective Neuroscience: The Foundations of Human and Animal Emotions*. New York: Oxford University Press.

Pessoa L (2012) Beyond brain regions: Network perspective of cognition-emotion interactions. *Behavioral and Brain Sciences* 35(3): 158–159.

Pessoa L (2013) *The Cognitive-Emotional Brain: From Interactions to Integration*. Cambridge, MA: MIT Press.

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): 773–783.

Plato (2003 [375 BC]) *The Republic*. London: Penguin Classics.

Power MJ and Dalgleish T (1997) *Cognition and Emotion: From Order to Disorder*. Hove: Psychology Press.

Russell JA (1980) A circumplex model of affect. *Journal of Personality and Social Psychology* 39(6): 1161–1178.

Russell JA and Barrett LF (1999) Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. *Journal of Personality and Social Psychology* 76(5): 805–819.

Scherer KR (2001) Appraisal considered as a process of multi-level sequential checking. In: Scherer KR, Schorr A and Johnstone T (eds) *Appraisal Processes in Emotion: Theory, Methods, Research*. New York: Oxford University Press, pp. 92–120.

Scherer KR (2005) What are emotions? And how can they be measured? *Social Science Information* 44(4): 693–727.

Shallice T (1988) Information processing models of consciousness. In: Marcel A and Bisiach E (eds) *Consciousness in Contemporary Science*. Oxford: Oxford University Press, pp. 305–333.

Solomon RC (1993) The philosophy of emotions. In: Lewis M and Haviland JM (eds) *Handbook of Emotions*. New York: Guildford Press, pp. 3–15.

Vuilleumier P (2005) How brains beware: Neural mechanisms of emotional attention. *Trends in Cognitive Sciences* 9(12): 585–594.

Zajonc RB (1980) Feeling and thinking: Preferences need no inferences. *American Psychologist* 35(2): 151–176.

Zajonc RB (1984) On the primacy of affect. *American Psychologist* 39(2): 117–124.