



# Piecing together the puzzle of emotional consciousness

Tahnée Engelen, Rocco Mennella

## ► To cite this version:

Tahnée Engelen, Rocco Mennella. Piecing together the puzzle of emotional consciousness. *Neuroscience of Consciousness*, 2023, 2023 (1), pp.104306. 10.1093/nc/niad005 . hal-04357299

**HAL Id: hal-04357299**

**<https://hal.science/hal-04357299>**

Submitted on 2 Feb 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Piecing together the puzzle of emotional consciousness

Tahnée Engelen<sup>1</sup>, Rocco Mennella<sup>1,2</sup>

March 2023

**1** Cognitive and Computational Neuroscience Laboratory (LNC2), Inserm U960, Department of Cognitive Studies, École Normale Supérieure, PSL University, 29 rue d’Ulm, 75005, Paris, France

**2** Laboratory on the Interactions between Cognition, Action, and Emotion (LICAÉ) – Paris Nanterre University, 200 avenue de La République, 92001 Nanterre Cedex, France.

Corresponding authors: Tahnée Engelen & Rocco Mennella

Corresponding authors e-mails: rmennella@parisnanterre.fr & tahnée.engelen@ens.fr

Corresponding authors telephone: +33144322638

Corresponding authors fax number: +33144322642

---

**Abstract.** The search for neural correlates of emotional consciousness has gained momentum in the last decades. Nonetheless, disagreements concerning the mechanisms that determine the experiential qualities of emotional consciousness – the what is it like to feel an emotion - as well as on their neural correlates have far reaching consequence on how researchers study and measure emotion, sometimes leading to seemingly irresolvable impasses. The current paper lays out in a balanced way the viewpoint of both cognitive and precognitive approaches to emotional consciousness, on the basis of commonalities and differences between the claims of some relevant theories of emotions. We examine the sufficiency of the existing evidence in support of the proposed theories of emotional consciousness, by going through the methodological specificity of the study of emotional consciousness and its unique challenges, highlighting what can and cannot be imported by advances on research on perceptual consciousness. We propose that there are three key experimental contrasts which are each equally necessary in the search for the neural correlates of emotional consciousness, each contrast alone coming with its own limitations. We conclude by acknowledging some of the most promising avenues in the field which may help go beyond current limitations and collaboratively piece together the puzzle of emotional consciousness.

**Keywords:** Emotion, Consciousness, Theories of Emotional Consciousness, Neural Correlates of Consciousness

## 1. Background and scopes

Understanding the phenomenal dimension of consciousness, the “what it is like” to have a specific conscious experience (Nagel, 1974), has long been considered a “hard” scientific problem (Chalmers, 1995). In the last decades, neuroscientists proposed an increasing number of theories of consciousness (see Seth & Bayne, 2022 for a review), to try to close (or at least reduce) the “explanatory gap” that we intuitively feel (but see Dennett, 2019) between conscious experience and its neural determinants, the so called Neural Correlates of Consciousness (NCC; Koch et al., 2016). In recent years, the development of NCC theories has run in parallel with, and sometimes inspired, a

particularly intense debate in the field of affective neuroscience around what emotions are and how to study them (Adolphs et al., 2019; D. J. Anderson & Adolphs, 2014; Barrett, 2017c, 2017b; Berridge, 2018; Fanselow & Pennington, 2018; LeDoux, 2019, 2020b, 2021; LeDoux & Brown, 2017; Mobbs et al., 2019; Panksepp et al., 2017; Taschereau-Dumouchel et al., 2022). In fact, most of the controversies around how to define and measure emotion stem from a disagreement concerning the mechanisms that determine the experiential qualities of emotional consciousness – the what is it like to feel an emotion - as well as on their neural correlates.

Despite the uncertainty around definitions and measures of emotion (Dukes et al., 2021), several

authors have brought forward theories of the neural correlates of emotional consciousness (from now on NCeC), which - not surprisingly - are often in contradiction with one another. In our opinion, the strength of the conviction by which, for instance, some theories define core concepts (e.g., “emotion”) is at present only partially justified by empirical data, mostly due to the unique methodological challenges in the study of the NCeC, which we will detail in the following sections. As previously pointed out by others (Pessoa, 2019), tight definitions can slow down scientific progress by hampering novel research ideas, which violate theoretical boundaries, rather than pushing them forward. This can create “tunnel vision” (e.g., Paré & Quirk, 2017), resulting in difficulty in studying the construct of interest in its full complexity, as well as in making comparisons between theories. Finally, this can lead to aprioristically valuing or devaluing specific measures (e.g., self-reports, physiology, behavior) or fields (e.g., human vs. animal research) in providing insightful information concerning the NCeC (for a discussion see e.g., Panksepp et al., 2017). By no means do we question the necessity of theory building for proper confirmatory hypothesis testing, as it is evident from the contrastive approach that we propose in the second part of the present work. Nonetheless, we may be missing crucial knowledge about the elements that are needed for deriving hypotheses from theories, namely concept definition, agreement on valid and reliable measures, definition of the relationships between concepts, boundary conditions, auxiliary assumptions and statistical predictions (see Scheel et al., 2021), which sometimes leads to seemingly irresolvable impasses. The present paper aims at reflecting upon possible strategies to build this knowledge, in order to strengthen theories of NCeC and move the debate forward. After reviewing the existing main approaches to emotional consciousness, we will try to challenge some of their fundamental assumptions, often by referring to advancements in research on the neural correlates of perceptual consciousness. The main aim of the present paper is to outline what we think are the theoretical and methodological specificities to the study of the NCeC. We argue that some of these limitations are unique to the study of the NCeC, that is, they are in part not shared with the study of other contents of consciousness. In this spirit, we will discuss the main experimental contrasts to be used in the search for NCeC, as

well as their interpretative value and limitations. In the final section we will reflect upon possible strategies to strengthen NCeC theories.

## 1.1. The contemporary debate on the NCeC

The present section does not mean to provide a full picture of the existing theories of emotion (for comprehensive reviews see Lange et al., 2020; Moors, 2022). We will instead focus on the place of consciousness in what are, in our view, some of the most influential recent perspectives. For the purpose of this paper, theoretical propositions on emotional consciousness will be divided in two approaches, which we call cognitive vs. precognitive approaches to NCeC. The cognitive vs. precognitive distinction refers here, respectively, to whether emotional consciousness is proposed to require cognitive processing (e.g., working memory, WM) to arise, or to whether it is thought to precede it. Other dichotomies have previously been put forward, such as cognitive vs. perceptual (Block, 2019), higher- vs. first-order (Ledoux & Brown, 2017) or access vs. phenomenal consciousness theories (Block, 1995, 2005), but we felt that the cognitive vs. precognitive distinction is more generalizable to the different sub-theories within each category, as it will be clear in the next sections. Within each approach, we will highlight important differences between theoretical perspectives, without aiming at presenting the full complexity of each author’s theory of emotion. Of note, the topic of the interaction between affect and consciousness, using a comparative science perspective, has also recently been the object of an excellent review, to which we address the interested reader (Paul et al., 2020).

### 1.1.1. Cognitive approach to emotional consciousness

Overall, emotion theories which we include into the cognitive approach to NCeC propose that, while humans and other animals share evolutionary conserved subcortical circuits to respond to threat, face aggression or seek rewards (LeDoux & Daw, 2018), the activation of none of these circuits alone is sufficient to cause a subjective emotional experience as humans conceive it. Theories within this approach agree that, in order for the individual to experience emotion, additional top-down pro-

cesses supported by neocortical associative areas are needed (e.g., Barrett, 2017c; Ledoux & Brown, 2017; Seth, 2013; Smith & Lane, 2015). The cognitive approach does not deny the existence of behavioral and visceral reactions to salient environmental stimuli, but do not support a reliable and specific causal relationship between these reactions and an instance of what we call emotion (Barrett, 2017c; Ledoux & Brown, 2017). Therefore, in this view, the subcortical circuits do not qualify as NCeC.

The precise way through which neocortical areas sustain emotional consciousness varies from one cognitive theory to another. Among the most influential ones, Ledoux’s higher-order theory of emotional consciousness (HOTEC; e.g., LeDoux, 2019, 2020a, 2020b, 2021; Ledoux & Brown, 2017; LeDoux & Hofmann, 2018, 2018; LeDoux & Pine, 2016; Taschereau-Dumouchel et al., 2022), combined existing higher-order theories of consciousness (e.g., R. Brown, 2015; Rosenthal, 2005) with previous work on fear and threat responses (for a summary see LeDoux, 2012). For the HOTEC, for an emotional experience to arise, higher-order representations of the stimulus, as opposed to first-order – sensorial or conceptual ones – need to enter working memory, via the brain’s “general networks of cognition”, which includes the lateral and medial prefrontal cortex, insular cortex and posterior parietal cortex. For instance, for experiencing fear, a threatening stimulus (e.g., a snake) needs first to be represented in visual areas (first-order representation) and integrated with long-term semantic memories - stored in medial temporal and other cortical areas - in a common unconscious WM representation (the first higher-order representation; HOR). This unconscious HOR can further be re-represented in WM, including autobiographic (self-related) memories (self-HOROR), becoming conscious. Specifically, for this self-HOROR to be an emotion (emotional self-HOROR), the consequences of the stimulus-evoked activation of the defensive circuits, including cortical arousal, behavioral and physiological reactions, need to be included in the representation in WM as well (Ledoux & Brown, 2017). The HOTEC proposes that our “emotion schemas”, i.e., what we know about emotions, which are part of the unconscious HORs, refine with experience, rendering the emotional experience more and more differentiated throughout development. Overall, the theory implies that emotions coincide with subjective

feelings (i.e., conscious emotion experiences) and therefore, since emotions can never be unconscious, individuals cannot be mistaken about the emotion that they are feeling (Ledoux & Brown, 2017). Hence, introspection and subjective reports, despite not being a one-to-one readout of the experience, are the gold standard for studying emotion (LeDoux & Hofmann, 2018).

Within the cognitive approach, most authors agree in identifying WM as a pivotal process to maintain feelings active over short time periods, in order to guide decisional and goal-directed behavior (Mikels & Reuter-Lorenz, 2019). Nonetheless, not every author agrees that higher-order representations in WM are needed for emotional consciousness to arise. For instance, Smith and Lane (2015, 2016), by combining Prinz’s perceptual model of emotional consciousness (Prinz, 2004, 2008) with cognitive theories of consciousness based on a global neuronal workspace (GNW) model (Baars, 2005; Dehaene & Naccache, 2001), proposed that the content of first-order representations of our own bodily reactions to the emotional situation can directly become conscious. This happens when it is globally broadcasted across a network of fronto-parietal regions, which renders the first-order representations available in WM, to guide goal-directed behavior. More specifically, emotional episodes are proposed to involve a hierarchical, and iterative, sequence of appraisals of salient stimuli/situations’ representations. The hierarchy runs upstream in the brain from automatic appraisals of basic characteristics (i.e., novelty and concern relevance), which primarily involve subcortical areas (amygdala, hippocampus), to more sophisticated ones (i.e., goal congruence, agency, value compatibility and affect meaning), subtended by temporal and prefrontal regions. This hierarchy of appraisals triggers and continuously refines bodily reactions (together with cognitive changes), which, as in Damasio’s “as-if loops” (Damasio, 1994), can sometimes only be represented and not embodied. In this view, emotions depend on our own perception of the abovementioned bodily changes, again in a hierarchical fashion, from the perception of discrete body features (stage 1; brainstem), to whole body patterns - which produce phenomenologically distinguishable bodily feelings (stage 2; anterior insula) - and finally to emotion concepts (stage 3; rACC), which categorize bodily feelings in an integrative manner (e.g., different bodily feelings can activate the same

emotion concept). Unconscious emotion representations in stages 2/3 are thought to compete to gain access to consciousness and, if selected via the global broadcasting mechanism in fronto-parietal areas, are rendered available to guide deliberative thoughts and action. In line with HOTEK, GNW theory of emotional consciousness proposes that the NCC are cortical, largely overlap with NCC of other contents of consciousness (Smith, Lane, et al., 2018, 2018), and rely on shared cognitive mechanisms (WM, attention). However, differently from HOTEK, GNW theory proposes that emotion representations can both be conscious and unconscious, and that our own conscious and reportable emotional experience can sometimes not be aligned with our behavioral and physiological emotional reactions, depending on how the probabilistic competition between active emotion concepts is resolved (e.g., one can act angry without feeling angry; Smith, Killgore, et al., 2018).

The focus on embodiment and on the probabilistic attribution of emotional concepts to bodily changes in specific situations is also at the core of cognitive theories centered around the notion of predictive coding (Barrett, 2017b, 2017c; Seth, 2013; Seth & Critchley, 2013). In this view, the brain is specifically conceived as a Bayesian prediction machine (Clark, 2013; Seth, 2015), which continuously runs and updates an internal model of the body in the world, in order to perform allostasis, that is, the regulation of the body based on metabolic cost and benefits (Barrett, 2017c; see also Damasio, 1994, 1998, 1999; Parvizi & Damasio, 2001). Internal models contain a) sensory predictions concerning the most probable next stimulation, b) motor predictions relative to the most appropriate actions to take and 3) visceromotor predictions regarding anticipated consequences for allostasis. Interoception of internal sensations (Craig, 2009) is thought to produce the lower dimensional feelings of affect (i.e., so called core affect) - valence and arousal - that are basic features of consciousness (Barrett, 2017c). Crucially, although core affects are probably shared in humans and other animals (Barrett et al., 2007; Barrett, 2017b), they are not emotions, nor are they specific to emotional episodes. While receiving sensory information both through exteroception and interoception, the brain computes a prediction error between the model and the actual state of the body in the world. The internal model that has the best fit, i.e., minimizes the prediction error, constitutes our

present perception, namely our experience. From this constructionist point of view, emotional experiences (i.e., fear) are concepts (internal models like any other), built upon experience, that best explain the present state of the body in the world and direct action (Barrett, 2017c). With regard to the neural correlates, the same reported emotion might depend on the activation of different neural correlates under different contexts, as it has been suggested that there is a many-to-many mapping between emotion categories and the combinations of somato-visceral/cognitive reactions and situational patterns they map to (Barrett, 2017c). Nonetheless, some authors insist on the importance of the salience network for running internal model of the body in the world, centered on the anterior insula and on the ACC (Seth, 2013), while others propose a more complex interaction among brain intrinsic networks (Barrett, 2017c; Barrett & Satpute, 2013).

HOTEK, GNW and predictive processing theories agree that emotional concepts are learned progressively through each individual's development, ultimately allowing for the attribution of meaning to our experiences in the world (Barrett, 2006, 2017c; Lane et al., 2015; Ledoux & Brown, 2017). This implies that each person has their own unique emotional experiences, as the way humans mentally represent, perceive, recognize and express emotions is profoundly shaped by life-history, socio-cultural influences and language (Barrett et al., 2019; Crivelli & Fridlund, 2018, 2019; R. E. Jack et al., 2012; Jackson et al., 2019). Individuals differ in their ability to conceptualize and understand their own affective states and responses – i.e., “emotional awareness” (Lane & Smith, 2021). Multiple possible neurocomputational mechanisms (Smith, Lane, et al., 2019; Smith, Parr, et al., 2019) explain why some individuals might, for instance, be able to feel and report valenced perceptual experiences of physiological changes (e.g., the distress associated to an increase in heart rate) but fail to be aware of and report a specific emotion category (e.g., fear). The ability to “put feelings into words”, reporting separate and detailed emotional experiences – so called “emotion differentiation” or “emotional granularity” - has proven adaptive for both psychological and social well-being and contributes to the heterogeneity of emotional experiences and reports (Kashdan et al., 2015). GNW theory in particular does not exclude that perceptually-based appraisals, pre-conceptual

ones, can trigger bodily changes and simple felt emotions in the absence of further cognitive processing, or via subcortical stimulation (Smith & Lane, 2015, 2016). Nonetheless, it is safe to say that cognitive theories converge in proposing that, while it is not impossible that other animals experience some form of affect, this is hardly comparable to the complex emotional experience available to humans and will hardly be directly measurable (Barrett, 2017b; LeDoux, 2021; LeDoux & Pine, 2016). In fact, it has been proposed that unjustified inference of subjective emotion from objective behavioral and physiological measures has led animal studies to produce scarce results in developing new pharmacological treatments for mental disorders in humans, such as depression and anxiety (LeDoux & Pine, 2016; Taschereau-Dumouchel et al., 2022).

### 1.1.2. Precognitive approach to emotional consciousness

The precognitive approach to NCeC claims either that emotions can be conscious in a way that is inaccessible to introspective scrutiny (pre-conscious), or that they can remain fully unconscious, in both cases having core neural determinants in specific evolutionary old brain structures, most of which are subcortical (Adolphs & Anderson, 2018; Berridge, 2018; Fanselow & Pennington, 2018; Izard, 2007b). Indeed, some authors within this approach align with theories identifying two types of consciousness (e.g., Block, 1995; Damasio, 1994, 1998; Merker, 2007): a “phenomenal” (or “core”) consciousness, which corresponds to the experience, the “what it is like” to be in a particular state from our unique first-person perspective, and an “access” (or “extended”) consciousness, which makes contents of experience available for further cognitive elaboration (e.g., subjective report). As two types of consciousness exist, two different NCCs are postulated (Block, 2005). Notably, it has been proposed that phenomenal consciousness emerged relatively early in evolution, as a core mechanism of integration of bodily signals and environmental objects, to guide action, and therefore strongly depend on the activation of subcortical brain structures, notably brainstem structures, such as the superior colliculus (SP) and its connections with the thalamus (Merker, 2007, 2013; Parvizi & Damasio, 2001; Shine, 2022). Participating to the ongoing de-

bate around the phenomenal vs. access distinction (Block, 2007; Naccache, 2018), several authors within the precognitive approach explicitly adhere to this distinction (Izard, 2007a, 2007b; Lieberman, 2019; Panksepp, 1998b, 2007; Panksepp et al., 2017), proposing that additional subcortical structures specifically determine the content of emotional phenomenal experience (e.g., the periaqueductal grey, PAG; Panksepp, 1998b, 1998a); other authors propose instead that emotions can either be conscious or unconscious (and not phenomenally pre-conscious), still largely attributing an important role to subcortical structure for unconscious emotions, for instance to the SC and the pulvinar (Celegnin et al., 2015; Méndez et al., 2022), the amygdala (AMY; D. J. Anderson & Adolphs, 2014; Fanselow, 2018) and the striatum (STR; Winkielman & Berridge, 2004).

Despite these differences concerning the existence of pre-conscious vs. fully unconscious emotional states, the underlying theory of emotion is largely shared between authors within the precognitive approach, and therefore more easily summarized compared to the cognitive approach. Overall, this approach proposes that the conscious representation that allows humans to produce verbal reports about what they feel (e.g., “I feel angry”) has no special status in defining what an emotion is. In fact, reportable subjective feelings in humans are only one of the components of a functionally organized set of behavioral, physiological and cognitive responses to environmental challenges, caused by “central generators” or “central states” of emotions (D. J. Anderson & Adolphs, 2014; Fanselow, 2018). Such central generators would (at least in part) map onto dedicated neural systems (Adolphs, 2013), mostly subcortical (Panksepp, 2007), which have been shaped by evolution. Accordingly, these ancient circuits control the execution of typical physiological and behavioral responses, such as freezing (Fanselow, 1994; Fanselow & Lester, 1988), “wanting” (Berridge et al., 1989; Berridge & Valenstein, 1991; Treit & Berridge, 1990) or RAGE (Panksepp, 1998a), which have been conserved since they have proven useful to adapt to environmental challenges. This set of physiological and behavioral responses is therefore a measurable and objective indicator of the activation of the central emotion generator. Since, for instance, direct stimulation of dedicated emotional neural circuits has reinforcing effects on behavior (i.e., animals actively try to pro-

long or terminate the stimulation) these circuits are inferred to be responsible for a phenomenal form of emotional consciousness (Panksepp, 1998a, 1998b), which is more differentiated than a simple core affect experience of something good or bad (as in Barrett, 2017b, 2017c). Other authors argue instead that, since such reinforcing effects on behavior, as well as typical emotion-related neural and physiological activations, can be elicited even in response to emotional stimuli which are not consciously detected and in the absence of changes in subjective reports, emotions can be fully unconscious (Berridge, 2018; Celeghin et al., 2015; Winkielman et al., 2005; Winkielman & Berridge, 2004). Depending on the author, the functional properties of evolutionarily conserved emotional circuits are thought to be partially (Adolphs, 2017) or almost entirely (Panksepp, 1998a) conserved between humans and other animals. The precognitive approach largely acknowledges that humans can further elaborate, reappraise and access to core emotional reactions to report them verbally. Some authors even made the distinction between “basic emotions” - the evolutionary conserved emotional responses - and the associations between such emotions and cognitive responses, acquired through learning, called “emotion schemas” (Ekman, 1992; Izard, 2007a). In this view, animal studies, as well as research on infants (Izard, 1991), are often thought to represent the gold standard for understanding the emotion primitives, non-contaminated by subsequent cognitive elaboration (Berridge, 2018; Fanselow & Pennington, 2018; Panksepp et al., 2017). Concerning animal research, this also stems from the possibility of directly stimulating/interfering with the central brain state with invasive techniques and measure effects on behavior, as well as provoking strong emotional states, which would be unethical in humans (Panksepp et al., 2017). Overall, it is expected that animal studies will be crucial (and for the precognitive approach they already have been) for gaining a better understanding of psychopathology and how to treat it (Berridge, 2018; Fanselow, 2018; Fanselow & Pennington, 2018; Panksepp et al., 2017).

## 1.2. Implications of the disagreements between approaches to study the Neural Correlates of Emotional Consciousness

The two approaches have some important points of disagreement. The primary theoretical point on the nature of emotional consciousness is whether the experience of emotion requires (or not) cognitive elaboration to take place. The cognitive approach proposes that it does, either in the form of a hierarchy of higher order representations in WM (e.g., Ledoux & Brown, 2017), or via the competition of first-order representations for global broadcasting in WM (e.g., Smith & Lane, 2015), or through the top-down categorization of bodily states in the world via emotion concepts (e.g., Barrett, 2017c). The precognitive approach proposes that cognition is not needed, either because the phenomenal, pre-conscious, experience is automatically elicited by the activation of central emotion generators in the brain and differs from the ability to reflect upon this experience (access consciousness) (e.g., Ekman, 1992; Izard, 2007b; Panksepp, 2007), or because emotions can be fully unconscious, when the stimulation of these circuits is not strong enough to produce conscious emotional experiences (D. J. Anderson & Adolphs, 2014; Fanselow, 2018; Winkielman & Berridge, 2004). Notably, this distinction can also be seen from the perspective of whether emotional consciousness and its neural correlates are largely shared with other contents of consciousness and mostly cortical (cognitive approach) or are specific to emotion and mostly subcortical (precognitive approach). Overall, such disagreements have far reaching consequences concerning the study of emotion (for recent discussions, see Mobbs et al., 2019; Panksepp et al., 2017; Paul et al., 2020), such as on 1) the neuroimaging techniques which have the capacity of providing evidence for the neural correlates of emotion (non-invasive techniques, as E/MEG, fMRI, NIRS, TMS, TDCS, TACS, which have no or limited access to subcortical structures, vs. invasive techniques, such as iEEG, intracranial stimulation, lesion studies, which can access these structures); 2) the measures that best capture what we mean by emotion (subjective reports vs. behavioral/physiological responses); 3) the pertinence of developmental studies, studies on patients with extended cognitive deficits and animal studies; 4) the potential for developing treatment

for mental disorders, such as depression and anxiety. Due to their extensive implications, the consequences of building theories of emotional consciousness on preliminary evidence can lead to particularly disturbing closed-ended impasses. This is well exemplified when comparing strong statements by authors on both sides, which are in explicit contradiction. For instance, for LeDoux and Hofmann (2018, p. 67), "the most direct way to assess conscious emotional feelings is through verbal self-report". In the very same year, Adolphs and Anderson stated that "A science of emotion should, in the first instance, use behavior, cognition, and neurobiology in its vocabulary. It should not be based on self-report of feelings in people" (Adolphs & Anderson, 2018, p. 51). Of note, not all authors that adopt a cognitive vs. precognitive approach to NCeC would fully commit to these statements, and we think that many bridges between approaches exist. To give an example, some of the theories which we assigned to either the cognitive approach (e.g., GNW; Smith & Lane, 2016) or to the precognitive (e.g., Celeghein et al., 2015; Winkielman & Berridge, 2004), which disagree on the centrality of cognition in making an emotion conscious, agree on the other hand on the existence of both conscious and unconscious emotions, and do not fully disregard some measures over others. Here we propose possible ways to cross bridges even further, by 1) identifying the methodological specificities of the study of NCeC compared to other NCC, 2) highlighting limitations to some of the claims of existing theories of NCeC, and by 3) exploring methodological solutions for testing specific hypotheses to build theories on a more solid ground in the future.

## **2. Are times mature to build theories of emotional consciousness and its neural correlates?**

### **2.1 Are separate theories of emotional consciousness and its neural correlates needed?**

Let us start with a global reflection concerning theories of emotional consciousness. A principle of parsimony would suggest not to bother building separate theories of emotional consciousness and

its neural correlates, unless evidence supports the notion that consciously experiencing an emotion differs in some way from other forms of experience, such as seeing red, feeling a fatigued muscle after exercise, or feeling thirsty. Intuition-wise, emotional phenomenal experience might seem to some of us qualitatively different from other conscious experiences, simply because watching our children making their first steps feels different from watching anybody else walking. Following this intuition makes the "hard problem" of consciousness (Chalmers, 1995) even harder when it comes to emotional experience, as there seems to be a 'something extra' to emotional consciousness that needs to be explained. As we have seen, across approaches, several authors provide explanation for this "something extra", arguing for example that it is the activation of a number of so-called "limbic" areas and connections, together with its cognitive, physiological and behavioral consequences, that is specific to the emergence of the conscious experience of emotion, irrespective of the differences on whether these areas fully qualify as NCeC (D. J. Anderson & Adolphs, 2014; Berridge, 2018; Fanselow, 2018; LeDoux, 2020a; Winkielman & Berridge, 2004). Other authors insist more on the activation of brain structures that allow for our own perception of changes in bodily states in a given situation as the unique mechanism that separate emotional from other forms of consciousness (e.g., Seth, 2013; Smith & Lane, 2015). Nonetheless, it is noteworthy that other theorists explicitly reject both the idea that there are dedicated emotion circuits in the brain and that feelings coming from the perception of bodily changes in the world are specific to emotion (Barrett, 2017b), aligning with theories that renounced at distinguishing emotional from other forms of experience at the mechanistic level (Moors, 2022; Russell, 2003), thus apparently questioning the necessity of separate theories of NCeC.

Another argument which is brought in favor of a special relationship between emotion and consciousness is that emotions are often thought to be deeply connected with unconscious aspects of our mental life. As we have briefly introduced, some research showed that, even when not consciously perceived, emotional stimuli can elicit physiological responses and bias behavior, similarly (but not identically) as would be observed when the stimulus would be reported as consciously perceived. In a phenomenon called affective blindsight, cor-



tically blind patients can identify the emotional expressions of faces above chance level but are incapable of doing the same for facial attributes unrelated to the emotional expression, such as identity (Rossion et al., 2000), potentially arguing for a special status for non-conscious emotional perception (Tamietto & de Gelder, 2010). In healthy participants, such findings are typically supported by adapting psychophysical paradigms commonly used in the study of visual consciousness, such as backward masking, binocular rivalry or continuous flash suppression (Kim & Blake, 2005). Results from these paradigms, which manipulate conscious access to emotional stimuli, show a ‘preferential access to awareness’ of emotional stimuli. For instance, there is a dominant viewing time for fearful faces in binocular rivalry (Amting et al., 2010) and times at which fearful faces break through binocular suppression are shortened (Yang et al., 2007). Once again, it is noteworthy that these findings are not unchallenged. Methodologically, it has been shown that differences between paradigms in the way stimulus’ awareness is suppressed influence emotion priming effects (Faivre et al., 2012). Furthermore, some have shown that, when stimulus’ awareness is not inferred by stimulus’ duration but it is based on actual subjects’ report, affective categorization of emotional stimuli was not better than chance in the absence of stimulus’ awareness (Lähteenmäki et al., 2015) and amygdala’s activation did not differ for presentation of fearful and neutral faces (Pessoa et al., 2006). By considering reported awareness, it has also been shown that only some physiological systems respond to unaware emotional vs. neutral stimuli in a continuous flash suppressing paradigm (Tooley et al., 2017). Other authors more generally argued that when fully controlling for stimulus awareness, neither behavioral nor physiological responses to emotional vs. neutral stimuli are observed (for a review, see Tsikandilakis et al., 2021). Mixed findings and opposing conclusions concerning unconscious emotional processing (Mertens & Engelhard, 2020; Rohr & Wentura, 2021) indicate that these mechanisms are still partially undetermined. Definitive evidence either in support for or against the necessity of building separate theories of emotional consciousness is still lacking, as the debate on the “something extra” to emotional consciousness, as well as on the specificity of unconscious emotional processing, wages on. However, and most importantly for the scopes of the present pa-

per, we will argue that irrespective on one’s opinion on the matter, the study of emotional consciousness comes with unique methodological challenges, which we will further elaborate in the following sections. More than the proven necessity of a separate theory of NCeC, it is the specificity of the methodological challenges to the study of emotional consciousness that justifies, in our opinion, that this research gets its own empirical and theoretical attention, before it can be clustered together with existing theories of consciousness.

## 2.2 Can content-specific NCeC be separated from neural prerequisites and consequences?

Across both approaches, a number of theories of emotional consciousness were predominantly built on evidence coming from research on single emotions, mostly on fear (e.g., D. J. Anderson & Adolphs, 2014; Celeghin et al., 2015; Fanselow, 2018; Ledoux & Brown, 2017), despite this not being the case for all theories (Barrett, 2017b; Izard, 2007b; Panksepp, 2007; Seth & Critchley, 2013; Smith & Lane, 2015; Winkielman & Berridge, 2004). This emphasis on fear research is understandable, as it derives from the robustness and reliability across species of fear/threat-related paradigms, such as fear/threat conditioning (e.g., Büchel & Dolan, 2000; Delgado et al., 2006; Fullana et al., 2020). Nonetheless, this solution comes with an important limitation, namely the possibility to make a distinction between the content-specific NCC and the full NCC (Koch, 2004; Koch et al., 2016). Content-specific NCC are defined as the neural substrates of specific phenomenal characteristic within an experience, e.g., the unique experience of seeing a face. Thus, the content-specific NCC differ from the so-called full NCC, i.e., the neural substrates of consciousness experience in its entirety, irrespective of the specific content of experience, meaning the combination of content-specific NCC for all possible contents of experience (Koch et al., 2016). To isolate content-specific NCC, typically, in the perceptual consciousness literature, contrastive approaches are used, for instance comparing brain activation when consciously perceiving a face vs. not consciously perceiving it, relying on the experimental paradigms introduced in the previous section. Importantly, a known shortcoming of such a contrastive approach is that it not only reveals

content-specific NCC for perceiving a face, but also all neural substrates preceding and following conscious perception, the so-called neural prerequisites and neural consequences of consciousness (Aru et al., 2012; de Graaf et al., 2012).

On the one hand, the neural prerequisites of consciousness refer to the mechanisms that are necessary for the conscious experience to arise, as for instance, in some circumstances, the fact of directing attention toward the stimulus, but are not the determinants of the content of the phenomenal experience (de Graaf et al., 2012). On the other hand, neural consequences are the after-effects associated to a given phenomenal experience, for instance when an episodic memory automatically comes to mind after consciously perceiving an object, like the Proustian madeleine (de Graaf et al., 2012). One proposed way of separating the different types of NCC relates to the notions of content-invariance and content-specificity. In more details, if a neural substrate is involved in the emergence of two distinct phenomenal experiences (content-invariance), it cannot explain the subjective difference between the two (content-specificity), and thus is more likely to be a prerequisite or a consequence than a content-specific NCC (de Graaf et al., 2012).

We argue that these notions directly apply to the search of NCeC. For instance, it is quite undebated that global affect dimensions such as arousal, valence and action tendencies participate to different extents to the emergence of each and every one of our emotional experiences (Lang & Bradley, 2010), therefore, applying the abovementioned logic, qualifying as neural prerequisites/consequences of emotional experience, rather than as content-specific NCeC. This is in line with what is proposed by predictive coding theories within the cognitive approach (Barrett, 2017b, 2017c; Seth, 2013; Seth & Critchley, 2013), which indeed deny the existence of content-specific NCC of what we call “emotions”, as emotions are learned concepts which vary across individuals and cultures, and therefore show idiosyncratic brain activation in similar emotional circumstances (« In those theories, variability is assumed to be the norm, rather than a nuisance to be explained after the fact »; Barrett, 2017a, p. 9). But even if we refer to theories that admit the existence of specific NCeC, the focus on fear over other emotions might have involuntarily produced a bias over the importance of specific neural structures over others. As an ex-

ample, the amygdala and its subcortical/cortical connections are at the core of a number of theories in both approaches, either as an important central emotion generator (Adolphs & Anderson, 2018; D. J. Anderson & Adolphs, 2014; Celeghin et al., 2015; Fanselow, 2018; Fanselow & Pennington, 2018), responsible for both conscious and unconscious fear, or at least as a necessary determinant for qualifying the conscious experience as “emotional” (see « emotional self-HOROR »; Ledoux & Brown, 2017). Not surprisingly, and in our opinion partially due to the fear-centered lenses and to the methodological difficulty in distinguishing content-specific NCC from prerequisites/consequences, it is debated whether this amygdala-centered network in humans is really fear/threat specific (Burra et al., 2019; McFadyen et al., 2019; Méndez-Bértolo et al., 2016), or it generally encodes affect dimensions such as arousal (Lin et al., 2020), action relevance (Guex et al., 2020) or stimulus valence (Kragel et al., 2021).

Partially supporting the confusion between content-specific NCeC and its prerequisites/consequences, research focused on other phenomenal emotional experiences, such as disgust, have found quite remarkably different neural correlates (e.g., the ventral striatum and the insula (Berridge, 2018; Chapman & Anderson, 2012)). Importantly, a similar confound applies for any research focusing on only one content of experience. Back to disgust and to the involvement of insula and generally interoceptive cortices (Chapman & Anderson, 2012), many authors agree that interoception is central in a wide range of emotional experiences (e.g., Pavuluri et al., 2015; Zaki et al., 2012), as we draw upon signals coming from our body to understand how we feel, and might for the very same reason as before not fully qualify as NCeC. It has to be noted, to avoid oversimplification, that the idea that the entire range of emotions that we are capable of experiencing would map onto one single brain area or network is refuted by both cognitive and precognitive approaches and irrespective of whether the authors believe emotions to be discrete or continuous in nature. For example, neither authors opposing (Clark-Polner et al., 2016) nor supporting (Saarimäki et al., 2016, 2018) the existence of discrete neural signatures of emotions would argue that all kind of emotional experience can be mapped onto a single generalized network. It is therefore at odds that approaches to emotional consciousness sometimes

overlook the necessity of fully integrating research on different contents of emotional consciousness before theory building. These premises made, how to exactly define what constitutes an emotional content is far from being resolved as, for instance, the debate on whether emotions are continuous or discrete in nature is still lively (Barrett, 2017a; Celeghin et al., 2017). The way we define an emotional content can directly impact our methods and conclusions we draw from results. For instance, it has been shown that assigning experimental stimuli to specific emotional categories has an impact on the performance of supervised machine learning algorithms that look for their brain correlates, and that the same categories are sometimes not retrieved with non-supervised algorithms (Azari et al., 2020). The need for clarity in content-definitions in the search for NCeC, which we will develop further throughout the paper, calls for even more caution when constructing theories of emotional consciousness.

### 2.3. Are the neural correlates of emotional consciousness cortical or subcortical?

When it comes to the specific claims on NCeC, the cognitive approach argues that, in order for a conscious emotional experience to exist, domain-general cognitive mechanisms, such as WM and attention, need to be recruited. In support of this claim, subcortical activations are proposed to be insufficient for conscious emotional experiences to arise and cortical activations are thought to be needed, namely activation in a prefrontal-parietal network (Smith & Lane, 2015), or more precisely in specific regions of the prefrontal cortex (Ledoux & Brown, 2017; Seth, 2013), which has indeed been found to correlate with self-reported emotions (Williams et al., 2006). Criticisms concerning the centrality of the prefrontal cortex in conscious experience have already been raised in the domain of perceptual consciousness, resulting from findings obtained with so-called no-report paradigms (Tsuchiya et al., 2015). As the name indicates, these paradigms require no subjective report from participants and rely on the idea that conscious contents can be inferred from physiological and behavioral changes, within specific experimental conditions. For instance, by inferring the perceived direction of competing moving stimuli under binocular rivalry from eye movements, rather than from

subjective reports, it has been shown that frontal activations that were present during perceptual transition in the report task were absent in the no-report variant of the task (Frassle et al., 2014). While the debate on the necessity of prefrontal cortex in perceptual consciousness continues (Block, 2019; Boly et al., 2017; Northoff & Lamme, 2020; Odegaard et al., 2017) emotional theories of consciousness claiming for a role of the prefrontal cortex face on this point the same methodological challenges as their non-emotional counterparts.

However, when it comes to the involvement of the prefrontal cortex in the NCeC, we argue that an additional methodological challenge exists. Namely, it is not clear yet how a no report paradigm could be adapted to an emotional context. In the emotion domain, even if the participant is aware of a fearful face or of an attacking snake, one cannot unambiguously assume that they are also experiencing a corresponding emotion. A direct example of how no-report paradigms might be challenging to adapt to emotional stimuli comes from work by Vetter et al. (2019). In their study, the authors presented both angry and fearful faces using a continuous flash suppression paradigm. They showed that, even in the absence of awareness of the face stimulus, eye-movements were still influenced in a seemingly goal-directed fashion, notably deviating toward the fearful face and away from the angry one (Vetter et al., 2019). These findings, which align with results for spontaneous approach/avoidance decision to task-irrelevant emotional faces (Grèzes et al., 2021; Mennella et al., 2020, 2022; Vilarem et al., 2020), show how emotion information can drive adaptive behavior in the absence of conscious experience, challenging a systematic inference of emotional states from behavior. Overall, in light of the possibility that the PFC might not be necessary to the phenomenal experience, but to the capacity to report it, we suggest that rigorous evidence in the support of the necessity of prefrontal areas to form conscious emotional experiences is lacking, in particular because no emotional counterpart to the no-report paradigm has been conceived at present. It is important to point out that the sufficiency of subcortical structures for the emergence of emotional consciousness, as proposed by the precognitive approach, is equally debated. Possibly due to the abovementioned fear-related bias, a longstanding question in the domain has been whether conscious fear can or cannot be experienced with-

out intact amygdalae. The precognitive approach typically refers to lesion studies in which bilateral amygdala damage impairs the recognition of fearful faces, the conscious experience of fear, and fear-related avoidance behaviors (Adolphs et al., 1994; Feinstein et al., 2011), as well as amygdala’s stimulation studies in humans that elicited in some cases conscious experiences of fear and anxiety (e.g., Inman et al., 2020; Lanteaume et al., 2007). On the other hand, authors within the cognitive approach more often refer to studies in which patients with amygdala damage were still able to experience conscious fear (A. K. Anderson & Phelps, 2002; Feinstein et al., 2013, 2016). Mixed findings are not surprising however, for both methodological and theoretical reasons. Methodologically, irrespective of the value of patient and lesion data in providing causal insight into brain functioning (Adolphs, 2016; Vaidya et al., 2019), such experiments come with limitations. Congenital lesions often result in plastic restructuring of the brain (Wieloch & Nikolich, 2006) and stimulation studies can be difficult to generalize, as different stimulation parameters (such as polarity and intensity) create varying electric fields in different patients (Selimbeyoglu, 2010). Theoretically, and more relevant for the present discussion, fear of exteroceptive threats (e.g., a predator) and of interoceptive ones (e.g., hypercapnic states due to exaggerated CO<sub>2</sub> levels), despite falling under the same “fear” label, are now known to rely on partially dissociable brain mechanisms and amygdala shows a different involvement in the two types of fear (for a recent review, see Feinstein et al., 2022), which likely explain some of the mixed findings. This exemplifies once more how ambiguity and disagreement in the definition of emotion concepts (e.g., fear) might artificially lead to conflicting and seemingly irreconcilable interpretations of the same research results.

As we have argued throughout this section, we believe that, due to 1) the lack of strong evidence against or in favor of the specificity of emotional vs. non-emotional consciousness, 2) the present difficulty in disambiguating content-specific NCeC from neural prerequisites/consequences and 3) unresolved issues substantiating either the predominantly cortical or subcortical nature of NCeC, there is no clear need of committing to either cognitive or precognitive approaches to emotional consciousness, nor to accept (or discard) one theory over another, among the proposed ones. This should, in our opinion, free researchers from some

of the abovementioned impasses which might lead to the exclusion of specific research fields, methods or techniques in the search for NCeC, pushing them to collaboratively refine concepts and find agreements on methods (see section 4). In the next section we discuss three fundamental experimental contrasts which are commonly employed in the search for NCeC, focusing on what can and cannot be inferred from each of them. We hope that this discussion can contribute to building a shared methodological ground for future theory building and research.

### 3. 3. Experimental contrasts and what can(not) be inferred

We propose that there are three main experimental contrasts which are crucial for hypothesis testing on the NCeC. These are widely employed in the literature but what we can conclude from them often remains implicit. We here directly point out the interpretive value of each contrast, the extent to which it can eventually be adapted from the perceptual to the emotional consciousness field, and its specificities (and shortcomings) to the study of the NCeC. We argue that, given the limitations of any of the contrasts taken alone, the combination each of them is crucial to further our understanding of emotional consciousness and its neural correlates. Of note, for simplicity, our discussion focuses on the study of emotional consciousness in response to external stimuli, but we by no mean disregard the fact that emotion responses and subjective experience might be caused by fully internal factors, both cognitive (e.g., thoughts, memories) and physiological (e.g., hormonal changes, physical fatigue, inflammation).

#### 3.1 Contrast 1: Subliminal versus supraliminal perception of the emotional stimulus

The contrast between sublimally and supraliminally perceived stimuli is widely used in studies of both perceptual and emotional consciousness (Mitchell & Greening, 2012), which we already introduced in previous sections. In paradigms such as masking or binocular rivalry, an emotional stimulus can alternate between being perceived by the participant (supraliminal) or not (subliminal), on a trial-by-trial basis. The report of the partici-

pant in these cases, if any is recorded, is whether they perceived the stimulus (in cases of masking or continuous flash suppression), or which stimulus they saw (in case of binocular rivalry, meaning that the non-reported stimulus was suppressed from awareness). These two types of trials are then contrasted to infer which areas were additionally activated by the conscious perception of the stimulus (de Graaf et al., 2012; Kim & Blake, 2005). As highlighted above, this contrast alone is limited in how much insight it can give us into the NCeC, in two ways. First, whatever neural correlates emerge from this contrast might not be content-specific and could rather be a reflection of a prerequisite or a consequence of emotional consciousness (Aru et al., 2012; de Graaf et al., 2012). Second, and more importantly, the subliminal versus supraliminal perception of the stimulus, albeit emotional, does not provide information about whether a corresponding emotion was felt. On the one hand, awareness of the stimulus and the emotion can be dissociated from one another, as not all emotional stimuli consciously perceived provoke subjective emotional experiences. On the other hand, as previously discussed, unconsciously perceived emotional stimuli might influence emotional perception, responses and decisions (Celeghein et al., 2015; Tamietto & de Gelder, 2010), including judgements of stimulus’ valence (E. Anderson et al., 2012), while possibly not conscious emotional feelings (Winkielman et al., 2005; Winkielman & Gogolushko, 2018). This aspect is also intimately connected with the notion of the respective timescales of perceptual vs. emotional experiences, which adds another level of complexity. On the one hand, it is still debated whether perceptual conscious experience of visual stimuli correlates with “early” neural activity, around 200 ms after stimulus presentation, i.e., the Visual Awareness Negativity (VAN) ERP component, or with “late” ERP components, such as the P3 and late positivity (LP) (for a recent review, see Förster et al., 2020). On the other hand, consciousness of emotional stimuli (emotional faces) might relate in a different and specific way to the early and late ERP components and their neural substrates, as compared to neutral stimuli (e.g., Sun et al., 2023). Crucially, this whole literature, which we only briefly mention here, moderately informs regarding the timescale of the emergence of the subjective emotional experience, which seems to us much less investigated. This might be due to

multiple factors, not last the fact that timing of the conscious emotional experience after discrete stimuli is inevitably confounded with the time and the neural activity needed to report it. While some studies did investigate the neural correlates of the transition between emotional vs. neutral conscious states, via repeated emotional stimuli presentation, the time needed for the instantiation of the emotional experience, as well as its specificity compared to mood changes for instance (Eldar et al., 2021), deserves further attention. Overall, these considerations point out some of the unique limitations to the use of the “Subliminal versus supraliminal perception” contrast in the search for NCeC. Therefore, the addition of the following contrasts is rendered necessary.

### **3.2 Contrast 2: Self-reported experienced (felt) vs not experienced (unfelt/different) emotion**

To be able to reveal content-specific correlates of emotional consciousness, conditions in which a stimulus elicit a reportable emotional experience need to be contrasted with conditions in which a stimulus elicits either an alternative or absence of reportable emotional experience. Unlike the previous contrast, this relies on supraliminal presentation, while varying, for example, stimulus’ content. This is a founding method in affective neuroscience, which makes use of countless databases of stimuli of different nature (e.g., pictures, sounds, imagery scripts) which are validated for their capacity to elicit, on average, different emotion experiences in the perceiver. Let us examine the limitations inherent to this contrast for the study of the NCeC. First, if neural activation is compared following a stimulus inducing one conscious emotional experience (e.g., fear) vs. another (e.g., sadness), one can learn about different correlates of the two emotions but cannot conclude much about what determines their conscious experience (an “unfelt” control condition is lacking). Second, if neural activation is compared following a stimulus inducing one conscious emotional experience (e.g., fear) vs. a neutral state, here the neural differences will contain the activity needed for the conscious emotion, but it will be confounded by the difference between stimuli, both sensorial and affective ones, such as arousal and valence (see Gasper et al., 2019 for further discussion on the use of neutral states as baseline in contrasts).

Third, for its interpretation, this contrast relies on self-report, through which participants somehow declare whether they consciously experienced an emotion and which one. Interestingly, a recent large-scale survey among researchers in the field of consciousness revealed that, although generally aware of their possible biases in measuring the content of experience, as well as of the above-mentioned influence of self-reporting on neural activation, researchers overall declared subjective reports to be their preferred method to measure consciousness (Francken et al., 2022). It is true that, whereas indeed accuracy of self-report based on metacognition abilities has been shown to be poor (Nisbett & Wilson, 1977), both consciousness and emotional consciousness are first-person experiences, and might be quite accurately accessible through introspection (Overgaard & Sandberg, 2012). Nonetheless, previous research in emotion has highlighted that validity of self-report is influenced by the way scales are constructed (e.g., dimensional vs. discrete), by where the focus of introspective attention is placed, or by the amount of elapsed time before the experience is captured (A. I. Jack & Roepstorff, 2003; Overgaard & Sandberg, 2012; Robinson & Clore, 2002). Furthermore, for the study of emotional consciousness, an additional level of difficulty exists, as a given stimulus might evoke a feeling of fear, but also feelings of anger, panic, sadness, or no feeling at all. Different emotional states can co-exist at the same time, a known phenomenon called dialecticism (Bagozzi et al., 1999; Lindquist & Barrett, 2008). Therefore, even if participants were immediately asked to indicate how fearful they felt on a continuous scale, we might be missing a great deal of relevant conscious experiences. A participant might even be primed to believe that what they had felt was fear, by framing the report measure in a particular way. Previous work has already shown that the mere act of asking a participant to report on their feelings can change both physiological and neural responses to a particular stimulus or task (Creswell et al., 2007; Kassam & Mendes, 2013; Lieberman et al., 2007). Recent research advancements are putting forward exciting new solutions to try and overcome some of these limitations. For instance, the use of continuous naturalistic stimuli, such as movies, have the power to profoundly modify experience (Kovarski et al., 2022), including emotion (Saarimäki, 2021), which can be measured over time on multiple di-

mensions, without loss of complexity. Continuous behavioral reports of affective (e.g., Smirnov et al., 2019) or discrete emotional states (e.g., Hudson et al., 2020) from the participant can be used, but this would come with the neural confound associated to self-report. To circumvent this, rating from independent samples can be collected, or automated emotion-feature extraction can be relied upon (Kragel et al., 2019). While each and every one of these methods comes with limitations (for a detailed review, see Saarimäki, 2021), these advancements nicely demonstrate where the additional challenge lies when studying emotional consciousness, as well as possible future solutions. The presented limitations for inference relative to the second contrast bring us to the third and final contrast.

### 3.3. Contrast 3: Presence/absence of a behavioral/physiological response

As a third point, conditions in which a stimulus elicits a given behavioral/physiological response need to be contrasted with conditions in which a stimulus elicits either an alternative or no behavioral/physiological responses. As presented in the first section, the precognitive approach supports the idea that behavioral and physiological responses can be a readout of emotional experience in animals and in humans which cannot verbally report experience (e.g., infants, patients) and, for some authors, represent the sole readout of phenomenal (as opposed to access) consciousness (Panksepp et al., 2017). As far as behavior is concerned, models have been proposed, for instance, based on the perceived distance to the threat, to infer from animals’ defensive behavior (e.g., freezing vs. vigorous escape attempts) their corresponding emotional experiences (e.g., fear vs. panic, respectively) (The Predatory Imminence Continuum Model; Fanselow et al., 2019; Fanselow & Lester, 1988; Mobbs et al., 2020). While the tight link between emotion and action is undebated, possible limitations to this kind of model are brought by converging research on defensive behavior in animals (e.g., Evans et al., 2019; Vale et al., 2017) and humans (Mennella et al., 2020; Rotteveel & Phaf, 2004, 2004; Schlund et al., 2016; Vilarem et al., 2020). Such research supports the notion that defensive behavior are flexible and not stereotyped in response to threatening stimuli, resulting from a complex, and still partially

unexplored interaction of stimulus-driven reactions (e.g., reflexes and automated reactions) and rapid - sometimes unconscious - goal-directed responses (LeDoux & Daw, 2018; Mendl & Paul, 2020; Moors et al., 2017). This is corroborated by the fact that goal-directed approach/avoidance behaviors to emotional stimuli can be elicited in the absence of a reportable strategy for action (Mennella et al., 2022), and in the absence of stimulus’ awareness (Vetter et al., 2019). A similar discourse applies to physiological responses, which are indeed typically well correlated to subjective reports of emotion at the group level (Friedman et al., 2014; Taschereau-Dumouchel et al., 2020), but not always to a great extent at the individual level: early studies on the relationship between behavioral ratings and physiological responses showed that a correlation between zygomatic and corrugator muscle responses with the rated experienced valence reached significance in roughly 50% of the participants, and the correlations between skin conductance responses and experienced arousal reached significance in around 30% of the sample (Lang et al., 1993). Such findings, together with the fact that, as discussed above, consciously undetected emotional stimuli can elicit physiological responses, have raised the question of whether subjective experience and physiological responses are subtended by the same brain mechanisms. Using multivoxel pattern analysis, Taschereau-Dumouchel and colleagues (2020) indeed showed that, despite a significant correlation between reported fear and skin conductance responses at the group level, some areas were differentially involved in the prediction of the two measures. Other recent work demonstrated that when modeling what contributes to subjective ratings that participants give in response to affective images, both physiological responses and neural (interoceptive) markers explain unique parts of the variance observed in self reports (Engelen et al., 2023). This means that although physiological responses did make a significant unique contribution to self-reports, they did not account for all of the variance observed in such ratings. Overall, there is evidence that specific behavior and physiological responses can be more or less correlated with subjectively reported emotional experiences, and individual differences, as well as the intensity of the emotional situation, have been found to influence the strength of this correlation, which is named emotion coherence (Mauss et al., 2005). It is noteworthy that

increased emotional coherence has been related to well-being (C. L. Brown et al., 2020), which aligns with the fact that emotion-related disorders, such as anxiety and depression, are undoubtedly “mental” disorders (Taschereau-Dumouchel et al., 2022), but also behavioral and physiological ones (American Psychiatric Association, 2013). This is the at the foundation of the necessity of the third contrast, as separate and complementary to the first two, without disregarding one or the other as not pertinent for the search of NCeC.

## Conclusions and future directions

We hope to have defended the idea that researchers should feel unconstrained with respect to the existing approaches to emotional consciousness, which, despite their value in pushing toward the formation of integrated and meaningful pictures of the existing research, at present rely on preliminary findings, which do not allow to arbitrate unambiguously between them. We insist that this state of things is not necessarily specific to the NCeC literature (Oberauer & Lewandowsky, 2019). Across many domains in psychology and neuroscience emerges the prevalence of a strong hypothethico-deductive method, which focuses on scientific progress as the repeated empirical test of hypotheses entailed by theories, while putting less attention upon the use of systematic and collective methods to developing theories in the first place (Borsboom et al., 2021). Another possibility is to conceive the relationship between phenomena, data and theories as a circular one, in which, before getting to the formulation of a complex theory, several steps are taken, starting with the identification of relevant phenomena and going through the initial formulation of preliminary explanatory models (prototheories), involving a small set of general principles that putatively explain the phenomena of interest (Borsboom et al., 2021). Theory construction in this sense builds on a range of non-confirmatory activities, including descriptive research and exploratory experimentation, which are crucial to reach consensus on concept definition and measures’ validity, among the elements which are necessary for theory building (Scheel et al., 2021). Back to the specific complexity of the study of NCeC, as we have detailed, methods cannot always be directly translated from the field of

perceptual consciousness to the study of NCeC. For example, no-report paradigms which have led to new insights in the search for the NCCs, cannot at present lend themselves in the same way for studying emotional consciousness. Likewise, the use of subjective reports in emotional consciousness might need further methodological elaboration for the study of NCeC as compared to NCC, due to, among other things, the dialectical nature of emotions. Lastly, behavioral and physiological responses play a possibly privileged role in the search for NCeC, as compared to NCC, as emotional coherence varies profoundly across individuals and situations.

Throughout the paper, we presented what we think are the necessary advancements to arrive to the separation of content-specific NCeC. These pertain to 1) how an emotional content of consciousness is defined and to 2) how the experimental contrasts must be combined, to avoiding confounding factors to the identification of the NCeC. Concerning concept definition, cooperative work in the form of adversarial collaborations is a way in which the energy of different labs (even adhering to different theories) can be combined. An adversarial collaboration is a multi-lab effort in which one prediction stemming from a particular theory is empirically tested. Such a collaboration has already been suggested (Seth & Bayne, 2022) and set up in the field of consciousness (Melloni et al., 2021), where predictions stemming from GNW and integrated information theory will be directly compared against one another. Most interestingly, an adversarial collaboration has been successfully applied in the field of emotions as well, to investigate how facial mimicry and voluntary facial activation can induce feelings of happiness (Coles et al., 2022). Aside from adversarial collaborations, data-driven approaches are particularly promising as they might play a part in the definition of emotion ontologies (R. E. Jack et al., 2018), thereby assisting in coming to a consensus in how to define emotional contents, in a more agnostic manner. For example, research into mapping the semantic space of emotional experience using several self-report measures in response to a wide array of emotionally evocative videos was able to identify 27 distinct categories of emotional experiences (Cowen & Keltner, 2017). Data-driven approaches have likewise been used to define the characteristics that define dynamic facial and bodily movements (de Gelder & Poyo Solanas, 2021; R. E. Jack

et al., 2014), as well as bodily sensation maps associated with subjective feelings (Nummenmaa et al., 2014).

Concerning the proposed contrasts, individual efforts focused on combining the three experimental contrasts for uncovering the NCeC can possibly benefit from some methodologies newly introduced to the study of emotion, such as computational models of behavior (Roberts & Hutcherson, 2019). These methods might help giving a more mechanistic account into the computational involvement of distinct brain areas, and possibly in the search for NCeC. For instance, recent work combining high field fMRI, continuous flash suppression, and drift diffusion modeling showed how the faster breaking through flash suppression of fearful faces is due to more rapid perceptual evidence accumulation, associated to activity in frontoparietal regions, occipital lobe, and amygdala. Activity in other areas, such as the insula and posterior cingulate cortex was rather correlated with a lower decision boundary (Kahlan, 2022). Finally, synthesis of existing data would make the comparison of all the different contrasts (which are unlikely to be addressed in one single study) feasible. Large scale meta-analyses have become more and more achievable thanks to the diffusion of open science practices, with the added advantage of increasing statistical power and determining consistency of effects (Yarkoni et al., 2010). Again, in the field of emotional consciousness such an approach has already shown to be informative, for instance to summarize evidence about physiological responses to subliminal negative affective stimuli (van der Ploeg et al., 2017).

To draw a final parallel, in the field of perceptual consciousness there is likewise still little agreement concerning the best theory of consciousness, but at least there seems to be a tendency toward a consensus on which methods should be employed to study the NCC (Francken et al., 2022). Acknowledging the necessity and strengths of all different methods and paradigms at our disposal to study the NCeC would be a major step forward, as the search for the NCeC is likely to stagnate by being exclusionary.

## Acknowledgements

T.E. is supported by funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innova-



tion program (grant agreement No 670325, Advanced grant BRAVIUS). R.M. was benefited from a Gretty Mirdal Junior Fellowship at the Paris Institute for Advanced Study (France) and was supported by Fondation de France (grant number 00100076). Both authors benefited from the Agence Nationale de la Recherche grant ANR-17-EURE-0017 (EUR FrontCog). The authors would like to thank Julie Grèzes, Catherine Tallon-Baudry, Tom de Graaf, Minye Zhan, Klara Kovarski, Morgan Beaurenaut, Clémence Alméras and Victor Chung for their support and valuable feedback on the manuscript.

## References

- Adolphs, R. (2013). The Biology of Fear. *Current Biology*, 23(2), R79-R93. <https://doi.org/10.1016/j.cub.2012.11.055>
- Adolphs, R. (2016). Human Lesion Studies in the 21st Century. *Neuron*, 90(6), 1151-1153. <https://doi.org/10.1016/j.neuron.2016.05.014>
- Adolphs, R. (2017). How should neuroscience study emotions? by distinguishing emotion states, concepts, and experiences. *Social Cognitive and Affective Neuroscience*, 12(1), 24-31. <https://doi.org/10.1093/scan/nsw153>
- Adolphs, R., & Anderson, D. J. (2018). *The Neuroscience of Emotion*. Princeton University Press. <https://doi.org/10.2307/j.ctvc77b1j>
- Adolphs, R., Mlodinow, L., Barrett, L. F. (2019). What is an emotion? *Current Biology*, 29(20), R1060-R1064. <https://doi.org/10.1016/j.cub.2019.09.008>
- Adolphs, R., Tranel, D., Damasio, H., & Damasio, A. (1994). Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. *Nature*, 372(6507), 669-672. <https://doi.org/10.1038/372669a0>
- American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders (Fifth Edition)*. American Psychiatric Association. <https://doi.org/10.1176/appi.books.9780890425596>
- Amting, J. M., Greening, S. G., & Mitchell, D. G. V. (2010). Multiple Mechanisms of Consciousness[U+202F]: The Neural Correlates of Emotional Awareness. *The Journal of Neuroscience*, 30(30), 10039-10047. <https://doi.org/10.1523/JNEUROSCI.6434-09.2010>
- Anderson, A. K., & Phelps, E. A. (2002). Is the Human Amygdala Critical for the Subjective Experience of Emotion? Evidence of Intact Dispositional Affect in Patients with Amygdala Lesions. *Journal of Cognitive Neuroscience*, 14(5), 709-720. <https://doi.org/10.1162/08989290260138618>
- Anderson, D. J., & Adolphs, R. (2014). A framework for studying emotions across species. *Cell*, 157(1), 187-200. <https://doi.org/10.1016/j.cell.2014.03.003>
- Anderson, E., Siegel, E., White, D., & Barrett, L. F. (2012). Out of Sight but Not Out of Mind[U+202F]: Unseen Affective Faces Influence Evaluations and Social Impressions. *Emotion (Washington, D.C.)*, 12(6), 1210-1221. <https://doi.org/10.1037/a0027514>
- Aru, J., Bachmann, T., Singer, W., & Melloni, L. (2012). Distilling the neural correlates of consciousness. *Neuroscience Biobehavioral Reviews*, 36(2), 737-746. <https://doi.org/10.1016/j.neubiorev.2011.12.003>
- Azari, B., Westlin, C., Satpute, A. B., Hutchinson, J. B., Kragel, P. A., Hoemann, K., Khan, Z., Wormwood, J. B., Quigley, K. S., Erdogmus, D., Dy, J., Brooks, D. H., & Barrett, L. F. (2020). Comparing supervised and unsupervised approaches to emotion categorization in the human brain, body, and subjective experience. *Scientific Reports*, 10(1), 20284. <https://doi.org/10.1038/s41598-020-77117-8>
- Baars, B. J. (2005). Global workspace theory of consciousness[U+202F]: Toward a cognitive neuroscience of human experience. In S. Laureys (Ed.), *Progress in Brain Research* (Vol. 150, p. 45-53). Elsevier. [https://doi.org/10.1016/S0079-6123\(05\)50004-9](https://doi.org/10.1016/S0079-6123(05)50004-9)
- Bagozzi, R. P., Wong, N., & Yi, Y. (1999). The Role of Culture and Gender in the Relationship between Positive and Negative Affect. *Cognition & Emotion*, 13(6), 641-672. <https://doi.org/10.1080/026999399379023>
- Barrett, L. F. (2006). Are Emotions Natural Kinds? Perspectives on Psychological Science, 1(1), 28-58. <https://doi.org/10.1111/j.1745-6916.2006.00003.x>
- Barrett, L. F. (2017a). Categories and Their Role in the Science of Emotion. *Psychological inquiry*, 28(1), 20-26. <https://doi.org/10.1080/1047840X.2017.1261581>
- Barrett, L. F. (2017b). How emotions are made[U+202F]: The secret life of the brain. Pan Macmillan. Barrett, L. F. (2017c). The theory of constructed emotion[U+202F]: An active inference account of interoception and categorization. *Social cognitive and affective neuroscience*, 12(1), 1-23. <https://doi.org/10.1093/scan/nsw154>
- Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019). Emotional Expressions Reconsidered[U+202F]: Challenges to Inferring Emotion From Human Facial Movements. *Psychological Science in the Public Interest*, 20(1), 1-68. <https://doi.org/10.1177/1529100619832930>
- Barrett, L. F., Lindquist, K. A., Bliss-Moreau, E., Duncan, S., Gendron, M., Mize, J., & Brennan, L. (2007). Of Mice and Men[U+202F]: Natural Kinds of Emotions in the Mammalian Brain[U+202F]? A Response to Panksepp and Izard. *Perspectives on Psychological Science*, 2(3), 297-312. <https://doi.org/10.1111/j.1745-6916.2007.00046.x>
- Barrett, L. F., & Satpute, A. B. (2013). Large-scale brain networks in affective and social neuroscience[U+202F]: Towards an integrative functional architecture of the brain. *Current Opinion in Neurobiology*, 23(3), 361-372. <https://doi.org/10.1016/j.conb.2012.12.012>
- Berridge, K. C. (2018). Evolving concepts of emotion and motivation. *Frontiers in Psychology*, 9(SEP), 1-20. <https://doi.org/10.3389/fpsyg.2018.01647>
- Berridge, K. C., & Valenstein, E. S. (1991). What Psychological Process Mediates Feeding Evoked by Electrical

- Stimulation of the Lateral Hypothalamus? *Behavioral Neuroscience*. <https://doi.org/10.1037/0735-7044.105.1.3>
- Berridge, K. C., Venier, I. L., & Robinson, T. E. (1989). Taste Reactivity Analysis of 6-Hydroxydopamine-Induced Aphagia[U+202F]: Implications for Arousal and Anhedonia Hypotheses of Dopamine Function. *Behavioral Neuroscience*. <https://doi.org/10.1037/0735-7044.103.1.36>
- Block, N. (1995). On a confusion about a function of consciousness. *Behavioral and Brain Sciences*, 18(2), 227-247. <https://doi.org/10.1017/S0140525X00038188>
- Block, N. (2005). Two neural correlates of consciousness. *Trends in Cognitive Sciences*, 9(2), 46-52. <https://doi.org/10.1016/j.tics.2004.12.006>
- Block, N. (2007). Consciousness, accessibility, and the mesh between psychology and neuroscience. *Behavioral and Brain Sciences*, 30(5-6), 481-499. <https://doi.org/10.1017/S0140525X07002786>
- Block, N. (2019). What Is Wrong with the No-Report Paradigm and How to Fix It. *Trends in Cognitive Sciences*, 23(12), 1003-1013. <https://doi.org/10.1016/j.tics.2019.10.001>
- Boly, M., Massimini, M., Tsuchiya, N., Postle, B. R., Koch, C., & Tononi, G. (2017). Are the Neural Correlates of Consciousness in the Front or in the Back of the Cerebral Cortex? *Clinical and Neuroimaging Evidence*. *The Journal of Neuroscience*, 37(40), 9603-9613. <https://doi.org/10.1523/JNEUROSCI.3218-16.2017>
- Borsboom, D., van der Maas, H. L. J., Dalege, J., Kievit, R. A., & Haig, B. D. (2021). Theory Construction Methodology[U+202F]: A Practical Framework for Building Theories in Psychology. *Perspectives on Psychological Science*, 16(4), 756-766. <https://doi.org/10.1177/1745691620969647>
- Brown, C. L., Van Doren, N., Ford, B. Q., Mauss, I. B., Sze, J. W., & Levenson, R. W. (2020). Coherence between subjective experience and physiology in emotion[U+202F]: Individual differences and implications for well-being. *Emotion*, 20(5), 818-829. <https://doi.org/10.1037/emo0000579>
- Brown, R. (2015). The HOROR theory of phenomenal consciousness. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 172(7), 1783-1794.
- Büchel, C., & Dolan, R. J. (2000). Classical fear conditioning in functional neuroimaging. *Current Opinion in Neurobiology*, 10(2), 219-223. [https://doi.org/10.1016/s0959-4388\(00\)00078-7](https://doi.org/10.1016/s0959-4388(00)00078-7)
- Burra, N., Hervais-Adelman, A., Celeghin, A., de Gelder, B., & Pegna, A. J. (2019). Affective blindsight relies on low spatial frequencies. *Neuropsychologia*, 128, 44-49. <https://doi.org/10.1016/j.neuropsychologia.2017.10.009>
- Celeghin, A., de Gelder, B., & Tamietto, M. (2015). From affective blindsight to emotional consciousness. *Consciousness and Cognition*, 36, 414-425. <https://doi.org/10.1016/j.concog.2015.05.007>
- Adolphs, R. (2013). The Biology of Fear. *Current Biology*, 23(2), R79-R93. <https://doi.org/10.1016/j.cub.2012.11.055>
- Adolphs, R. (2016). Human Lesion Studies in the 21st Century. *Neuron*, 90(6), 1151-1153. <https://doi.org/10.1016/j.neuron.2016.05.014>
- Adolphs, R. (2017). How should neuroscience study emotions[U+202F]? By distinguishing emotion states, concepts, and experiences. *Social Cognitive and Affective Neuroscience*, 12(1), 24-31. <https://doi.org/10.1093/scan/nsw153>
- Adolphs, R., & Anderson, D. J. (2018). *The Neuroscience of Emotion*. Princeton University Press. <https://doi.org/10.2307/j.ctvc77b1j>
- Adolphs, R., Mlodinow, L., & Barrett, L. F. (2019). What is an emotion? *Current Biology*, 29(20), R1060-R1064. <https://doi.org/10.1016/j.cub.2019.09.008>
- Adolphs, R., Tranel, D., Damasio, H., & Damasio, A. (1994). Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. *Nature*, 372(6507), 669-672. <https://doi.org/10.1038/372669a0>
- American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders (Fifth Edition)*. American Psychiatric Association. <https://doi.org/10.1176/appi.books.9780890425596>
- Amting, J. M., Greening, S. G., & Mitchell, D. G. V. (2010). Multiple Mechanisms of Consciousness[U+202F]: The Neural Correlates of Emotional Awareness. *The Journal of Neuroscience*, 30(30), 10039-10047. <https://doi.org/10.1523/JNEUROSCI.6434-09.2010>
- Anderson, A. K., & Phelps, E. A. (2002). Is the Human Amygdala Critical for the Subjective Experience of Emotion? Evidence of Intact Dispositional Affect in Patients with Amygdala Lesions. *Journal of Cognitive Neuroscience*, 14(5), 709-720. <https://doi.org/10.1162/08989290260138618>
- Anderson, D. J., & Adolphs, R. (2014). A framework for studying emotions across species. *Cell*, 157(1), 187-200. <https://doi.org/10.1016/j.cell.2014.03.003>
- Anderson, E., Siegel, E., White, D., & Barrett, L. F. (2012). Out of Sight but Not Out of Mind[U+202F]: Unseen Affective Faces Influence Evaluations and Social Impressions. *Emotion (Washington, D.C.)*, 12(6), 1210-1221. <https://doi.org/10.1037/a0027514>
- Aru, J., Bachmann, T., Singer, W., & Melloni, L. (2012). Distilling the neural correlates of consciousness. *Neuroscience Biobehavioral Reviews*, 36(2), 737-746. <https://doi.org/10.1016/j.neubiorev.2011.12.003>
- Azari, B., Westlin, C., Satpute, A. B., Hutchinson, J. B., Kragel, P. A., Hoemann, K., Khan, Z., Wormwood, J. B., Quigley, K. S., Erdogmus, D., Dy, J., Brooks, D. H., & Barrett, L. F. (2020). Comparing supervised and unsupervised approaches to emotion categorization in the human brain, body, and subjective experience. *Scientific Reports*, 10(1), 20284. <https://doi.org/10.1038/s41598-020-77117-8>
- Baars, B. J. (2005). Global workspace theory of consciousness[U+202F]: Toward a cognitive neuroscience of human experience. In S. Laureys (Éd.), *Progress in Brain Research* (Vol. 150, p. 45-53). Elsevier. [https://doi.org/10.1016/S0079-6123\(05\)50004-9](https://doi.org/10.1016/S0079-6123(05)50004-9)
- Bagozzi, R. P., Wong, N., & Yi, Y. (1999). The Role of Culture and Gender in the Relationship between Positive and Negative Affect. *Cognition & Emotion*, 13(6), 641-672. <https://doi.org/10.1080/026999399379023>
- Barrett, L. F. (2006). Are Emotions Natural Kinds?

- Perspectives on Psychological Science, 1(1), 28-58. <https://doi.org/10.1111/j.1745-6916.2006.00003.x>
- Barrett, L. F. (2017a). Categories and Their Role in the Science of Emotion. *Psychological inquiry*, 28(1), 20-26. <https://doi.org/10.1080/1047840X.2017.1261581>
- Barrett, L. F. (2017b). How emotions are made: The secret life of the brain. Pan Macmillan.
- Barrett, L. F. (2017c). The theory of constructed emotion[U+202F]: An active inference account of interoception and categorization. *Social cognitive and affective neuroscience*, 12(1), 1-23. <https://doi.org/10.1093/scan/nsw154>
- Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019). Emotional Expressions Reconsidered[U+202F]: Challenges to Inferring Emotion From Human Facial Movements. *Psychological Science in the Public Interest*, 20(1), 1-68. <https://doi.org/10.1177/1529100619832930>
- Barrett, L. F., Lindquist, K. A., Bliss-Moreau, E., Duncan, S., Gendron, M., Mize, J., & Brennan, L. (2007). Of Mice and Men[U+202F]: Natural Kinds of Emotions in the Mammalian Brain[U+202F]? A Response to Panksepp and Izard. *Perspectives on Psychological Science*, 2(3), 297-312. <https://doi.org/10.1111/j.1745-6916.2007.00046.x>
- Barrett, L. F., & Satpute, A. B. (2013). Large-scale brain networks in affective and social neuroscience[U+202F]: Towards an integrative functional architecture of the brain. *Current Opinion in Neurobiology*, 23(3), 361-372. <https://doi.org/10.1016/j.conb.2012.12.012>
- Berridge, K. C. (2018). Evolving concepts of emotion and motivation. *Frontiers in Psychology*, 9(SEP), 1-20. <https://doi.org/10.3389/fpsyg.2018.01647>
- Berridge, K. C., & Valenstein, E. S. (1991). What Psychological Process Mediates Feeding Evoked by Electrical Stimulation of the Lateral Hypothalamus? *Behavioral Neuroscience*. <https://doi.org/10.1037/0735-7044.105.1.3>
- Berridge, K. C., Venier, I. L., & Robinson, T. E. (1989). Taste Reactivity Analysis of 6-Hydroxydopamine-Induced Aphagia[U+202F]: Implications for Arousal and Anhedonia Hypotheses of Dopamine Function. *Behavioral Neuroscience*. <https://doi.org/10.1037/0735-7044.103.1.36>
- Block, N. (1995). On a confusion about a function of consciousness. *Behavioral and Brain Sciences*, 18(2), 227-247. <https://doi.org/10.1017/S0140525X00038188>
- Block, N. (2005). Two neural correlates of consciousness. *Trends in Cognitive Sciences*, 9(2), 46-52. <https://doi.org/10.1016/j.tics.2004.12.006>
- Block, N. (2007). Consciousness, accessibility, and the mesh between psychology and neuroscience. *Behavioral and Brain Sciences*, 30(5-6), 481-499. <https://doi.org/10.1017/S0140525X07002786>
- Block, N. (2019). What Is Wrong with the No-Report Paradigm and How to Fix It. *Trends in Cognitive Sciences*, 23(12), 1003-1013. <https://doi.org/10.1016/j.tics.2019.10.001>
- Boly, M., Massimini, M., Tsuchiya, N., Postle, B. R., Koch, C., & Tononi, G. (2017). Are the Neural Correlates of Consciousness in the Front or in the Back of the Cerebral Cortex? Clinical and Neuroimaging Evidence. *The Journal of Neuroscience*, 37(40), 9603-9613. <https://doi.org/10.1523/JNEUROSCI.3218-16.2017>
- Borsboom, D., van der Maas, H. L. J., Dalege, J., Kievit, R. A., & Haig, B. D. (2021). Theory Construction Methodology[U+202F]: A Practical Framework for Building Theories in Psychology. *Perspectives on Psychological Science*, 16(4), 756-766. <https://doi.org/10.1177/1745691620969647>
- Brown, C. L., Van Doren, N., Ford, B. Q., Mauss, I. B., Sze, J. W., & Levenson, R. W. (2020). Coherence between subjective experience and physiology in emotion[U+202F]: Individual differences and implications for well-being. *Emotion*, 20(5), 818-829. <https://doi.org/10.1037/emo0000579>
- Brown, R. (2015). The HOROR theory of phenomenal consciousness. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 172(7), 1783-1794.
- Büchel, C., & Dolan, R. J. (2000). Classical fear conditioning in functional neuroimaging. *Current Opinion in Neurobiology*, 10(2), 219-223. [https://doi.org/10.1016/s0959-4388\(00\)00078-7](https://doi.org/10.1016/s0959-4388(00)00078-7)
- Burra, N., Hervais-Adelman, A., Celeghin, A., de Gelder, B., & Pegna, A. J. (2019). Affective blindsight relies on low spatial frequencies. *Neuropsychologia*, 128, 44-49. <https://doi.org/10.1016/j.neuropsychologia.2017.10.009>
- Celeghin, A., de Gelder, B., & Tamietto, M. (2015). From affective blindsight to emotional consciousness. *Consciousness and Cognition*, 36, 414-425. <https://doi.org/10.1016/j.concog.2015.05.007>
- Celeghin, A., Diano, M., Bagnis, A., Viola, M., & Tamietto, M. (2017). Basic Emotions in Human Neuroscience[U+202F]: Neuroimaging and Beyond. *Frontiers in Psychology*, 8. <https://www.frontiersin.org/articles/10.3389/fpsyg.2017.01432>
- Chalmers, D. (1995). Facing Up to the Problem of Consciousness. *Journal of Consciousness Studies*, 2(3), 200-219.
- Chapman, H. A., & Anderson, A. K. (2012). Understanding disgust. *Annals of the New York Academy of Sciences*, 1251(1), 62-76. [doi.org/10.1111/j.1749-6632.2011.06369.x](https://doi.org/10.1111/j.1749-6632.2011.06369.x)
- Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36(3), 181-204. <https://doi.org/10.1017/S0140525X12000477>
- Clark-Polner, E., Johnson, T. D., & Barrett, L. F. (2016). Multivoxel Pattern Analysis Does Not Provide Evidence to Support the Existence of Basic Emotions. *Cerebral Cortex*, bhw028. <https://doi.org/10.1093/cercor/bhw028>
- Coles, N. A., March, D. S., Marmolejo-Ramos, F., Larsen, J. T., Arinze, N. C., Ndukaihe, I. L. G., Willis, M. L., Foroni, F., Reggev, N., Mokady, A., Forscher, P. S., Hunter, J. F., Kaminski, G., Yüvrük, E., Kapucu, A., Nagy, T., Hajdu, N., Tejada, J., Freitag, R. M. K., ... Liuzza, M. T. (2022). A multi-lab test of the facial feedback hypothesis by the Many Smiles Collaboration. *Nature Human Behaviour*, 6(12), Art. 12. [doi.org/10.1038/s41562-022-01458-9](https://doi.org/10.1038/s41562-022-01458-9)
- Cowen, A. S., & Keltner, D. (2017). Self-report captures 27 distinct categories of emotion bridged

- by continuous gradients. *Proceedings of the National Academy of Sciences*, 114(38), E7900-E7909. <https://doi.org/10.1073/pnas.1702247114>
- Craig, A. D. (2009). How do you feel — now? The anterior insula and human awareness. *Nature Reviews Neuroscience*, 10(1), Art. 1. <https://doi.org/10.1038/nrn2555>
- Creswell, J. D., Way, B. M., Eisenberger, N. I., & Lieberman, M. D. (2007). Neural correlates of dispositional mindfulness during affect labeling. *Psychosomatic Medicine*, 69(6), 560-565. <https://doi.org/10.1097/PSY.0b013e3180f6171f>
- Crivelli, C., & Fridlund, A. J. (2018). Facial Displays Are Tools for Social Influence. *Trends in Cognitive Sciences*, 22(5), 388-399. <https://doi.org/10.1016/j.tics.2018.02.006>
- Crivelli, C., & Fridlund, A. J. (2019). Inside-Out [U+202F]: From Basic Emotions Theory to the Behavioral Ecology View. In *Journal of Nonverbal Behavior* (Vol. 43, Numéro 2). Springer US. [doi.org/10.1007/s10919-019-00294-2](https://doi.org/10.1007/s10919-019-00294-2)
- Damasio, A. R. (1994). *Descartes' Error* [U+202F]: Emotion, Reason, and the Human Brain. (Putnam).
- Damasio, A. R. (1998). Investigating the biology of consciousness. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 353(1377), 1879. <https://doi.org/10.1098/rstb.1998.0339>
- Damasio, A. R. (1999). The Feeling of What Happens [U+202F]: Body and Emotion in the Making of Consciousness. Harcourt Brace and Co.
- de Gelder, B., & Poyo Solanas, M. (2021). A computational neuroethology perspective on body and expression perception. *Trends in Cognitive Sciences*, 1-13. <https://doi.org/10.1016/j.tics.2021.05.010>
- de Graaf, T. A., Hsieh, P.-J., & Sack, A. T. (2012). The « correlates » in neural correlates of consciousness. *Neuroscience and Biobehavioral Reviews*, 36(1), 191-197. <https://doi.org/10.1016/j.neubiorev.2011.05.012>
- Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness [U+202F]: Basic evidence and a workspace framework. *Cognition*, 79(1), 1-37. [https://doi.org/10.1016/S0010-0277\(00\)00123-2](https://doi.org/10.1016/S0010-0277(00)00123-2)
- Delgado, M. R., Olsson, A., & Phelps, E. A. (2006). Extending animal models of fear conditioning to humans. *Biological Psychology*, 73(1), 39-48. <https://doi.org/10.1016/j.biopsycho.2006.01.006>
- Dennett, D. C. (2019). Welcome to Strong Illusionism. *Journal of Consciousness Studies*, 26(9-10), 48-58. Dukes,
- D., Abrams, K., Adolphs, R., Ahmed, M. E., Beatty, A., Berridge, K. C., Broomhall, S., Brosch, T., Campos, J. J., Clay, Z., Clément, F., Cunningham, W. A., Damasio, A., Damasio, H., D'Arms, J., Davidson, J. W., de Gelder, B., Deonna, J., de Sousa, R., ... Sander, D. (2021). The rise of affectivism. *Nature Human Behaviour*, 5(7), Art. 7. <https://doi.org/10.1038/s41562-021-01130-8>
- Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, 6(3-4), 169-200. <https://doi.org/10.1080/02699939208411068>
- Eldar, E., Pessiglione, M., & van Dillen, L. (2021). Positive affect as a computational mechanism. *Current Opinion in Behavioral Sciences*, 39, 52-57. <https://doi.org/10.1016/j.cobeha.2021.01.007>
- Engelen, T., Buot, A., Grèzes, J., & Tallon-Baudry, C. (2023). Whose emotion is it? Perspective matters to understand brain-body interactions in emotions. *NeuroImage*, 268, 119867. <https://doi.org/10.1016/j.neuroimage.2023.119867>
- Evans, D. A., Stempel, A. V., Vale, R., & Branco, T. (2019). Cognitive Control of Escape Behaviour. *Trends in Cognitive Sciences*, 23(4), 334-348. <https://doi.org/10.1016/j.tics.2019.01.012>
- Faivre, N., Berthet, V., & Kouider, S. (2012). Non-conscious influences from emotional faces [U+202F]: A comparison of visual crowding, masking, and continuous flash suppression. *Frontiers in Psychology*, 3, 129. <https://doi.org/10.3389/fpsyg.2012.00129>
- Fanselow, M. S. (1994). Neural organization of the defensive behavior system responsible for fear. *Psychonomic bulletin & review*, 1(4), 429-438. <https://doi.org/10.3758/BF03210947>
- Fanselow, M. S. (2018). Emotion, motivation and function. *Current Opinion in Behavioral Sciences*, 19, 105-109. <https://doi.org/10.1016/j.cobeha.2017.12.013>
- Fanselow, M. S., Hoffman, A. N., & Zhuravka, I. (2019). Timing and the transition between modes in the defensive behavior system. *Behavioural Processes*, 166, 103890. <https://doi.org/10.1016/j.beproc.2019.103890>
- Fanselow, M. S., & Lester, L. S. (1988). A functional behavioristic approach to aversively motivated behavior [U+202F]: Predatory imminence as a determinant of the topography of defensive behavior. *Evolution and Learning*, January 1988, 185-212. Fanselow, M. S., &
- Pennington, Z. T. (2018). A return to the psychiatric dark ages with a two-system framework for fear. *Behaviour Research and Therapy*, 100(November 2017), 24-29. <https://doi.org/10.1016/j.brat.2017.10.012>
- Feinstein, J. S., Adolphs, R., Damasio, A., & Tranel, D. (2011). The Human Amygdala and the Induction and Experience of Fear. *Current Biology*, 21(1), 34-38. <https://doi.org/10.1016/j.cub.2010.11.042>
- Feinstein, J. S., Adolphs, R., & Tranel, D. (2016). A tale of survival from the world of patient S. M. In *Living without an amygdala* (p. 1-38). The Guilford Press. Feinstein, J. S., Buzza, C., Hurlemann, R., Follmer, R. L., Dahdaleh, N. S., Coryell, W. H., Welsh, M. J., Tranel, D., & Wemmie, J. A. (2013). Fear and panic in humans with bilateral amygdala damage. *Nature Neuroscience*, 16(3), 270-272. <https://doi.org/10.1038/nn.3323>
- Feinstein, J. S., Gould, D., & Khalsa, S. S. (2022). Amygdala-driven apnea and the chemoreceptive origin of anxiety. *Biological Psychology*, 170, 108305. <https://doi.org/10.1016/j.biopsycho.2022.108305>
- Förster, J., Koivisto, M., & Revonsuo, A. (2020). ERP and MEG correlates of visual consciousness [U+202F]: The second decade. *Consciousness and Cognition*, 80, 102917. <https://doi.org/10.1016/j.concog.2020.102917>
- Francken, J. C., Beerendonk, L., Molenaar, D., Fahrenfort, J. J., Kiverstein, J. D., Seth, A. K., & van Gaal, S. (2022). An academic survey on theoretical foundations,

- common assumptions and the current state of consciousness science. *Neuroscience of Consciousness*, 2022(1), niac011. <https://doi.org/10.1093/nc/niac011>
- Friedman, B. H., Stephens, C. L., & Thayer, J. F. (2014). Redundancy analysis of autonomic and self-reported, responses to induced emotions. *Biological Psychology*, 98, 19-28. <https://doi.org/10.1016/j.biopsycho.2013.12.006>
- Fullana, M. A., Dunsmoor, J. E., Schruers, K. R. J., Savage, H. S., Bach, D. R., & Harrison, B. J. (2020). Human fear conditioning[U+202F]: From neuroscience to the clinic. *Behaviour Research and Therapy*, 124, 103528. <https://doi.org/10.1016/j.brat.2019.103528>
- Gasper, K., Spencer, L. A., & Hu, D. (2019). Does Neutral Affect Exist? How Challenging Three Beliefs About Neutral Affect Can Advance Affective Research. *Frontiers in Psychology*, 10. <https://www.frontiersin.org/articles/10.3389/fpsyg.2019.02476>
- Grèzes, J., Erblang, M., Vilarem, E., Quiquempoix, M., Van Beers, P., Guillard, M., Sauvet, F., Mennella, R., & Rabat, A. (2021). Impact of total sleep deprivation and related mood changes on approach-avoidance decisions to threat-related facial displays. *Sleep*, zsab186. <https://doi.org/10.1093/sleep/zsab186>
- Guex, R., Méndez-Bértolo, C., Moratti, S., Strange, B. A., Spinelli, L., Murray, R. J., Sander, D., Seeck, M., Vuilleumier, P., & Domínguez-Borràs, J. (2020). Temporal dynamics of amygdala response to emotion and action-relevance. *Scientific Reports*, 10(1), 11138. <https://doi.org/10.1038/s41598-020-67862-1>
- Hudson, M., Seppälä, K., Putkinen, V., Sun, L., Gleason, E., Karjalainen, T., Karlsson, H. K., Hirvonen, J., & Nummenmaa, L. (2020). Dissociable neural systems for unconditioned acute and sustained fear. *NeuroImage*, 216, 116522. <https://doi.org/10.1016/j.neuroimage.2020.116522>
- Inman, C. S., Bijanki, K. R., Bass, D. I., Gross, R. E., Hamann, S., & Willie, J. T. (2020). Human amygdala stimulation effects on emotion physiology and emotional experience. *Neuropsychologia*, 145, 106722. <https://doi.org/10.1016/j.neuropsychologia.2018.03.019>
- Izard, C. E. (1991). *The Psychology of Emotions*. Springer US. <https://doi.org/10.1007/978-1-4899-0615-1>
- Izard, C. E. (2007a). Basic Emotions, Natural Kinds, Emotion Schemas, and a New Paradigm. *Perspectives on Psychological Science*, 2(3), 260-280. <https://doi.org/10.1111/j.1745-6916.2007.00044.x>
- Izard, C. E. (2007b). Levels of emotion and levels of consciousness. *Behavioral and Brain Sciences*, 30(1), 96-98. <https://doi.org/10.1017/S0140525X07001045>
- Jack, A. I., & Roepstorff, A. (2003). Why Trust the Subject? *Journal of Consciousness Studies*, 10, v-xx.
- Jack, R. E., Crivelli, C., & Wheatley, T. (2018). Data-Driven Methods to Diversify Knowledge of Human Psychology. *Trends in Cognitive Sciences*, 22(1), 1-5. <https://doi.org/10.1016/j.tics.2017.10.002>
- Jack, R. E., Garrod, O. G. B., & Schyns, P. G. (2014). Dynamic Facial Expressions of Emotion Transmit an Evolving Hierarchy of Signals over Time. *Current Biology*, 24(2), 187-192. <https://doi.org/10.1016/j.cub.2013.11.064>
- Jack, R. E., Garrod, O. G. B., Yu, H., Caldara, R., & Schyns, P. G. (2012). Facial expressions of emotion are not culturally universal. *Proceedings of the National Academy of Sciences of the United States of America*, 109(19), 7241-7244. <https://doi.org/10.1073/pnas.1200155109>
- Jackson, J. C., Watts, J., Henry, T. R., List, J., Forkel, R., Mucha, P. J., Greenhill, S. J., Gray, R. D., & Lindquist, K. A. (2019). Emotion semantics show both cultural variation and universal structure. *Science (New York, N.Y.)*, 366(6472), 1517-1522. <https://doi.org/10.1126/science.aaw8160>
- Kashdan, T. B., Barrett, L. F., & McKnight, P. E. (2015). Unpacking Emotion Differentiation[U+202F]: Transforming Unpleasant Experience by Perceiving Distinctions in Negativity. *Current Directions in Psychological Science*, 24(1), 10-16. <https://doi.org/10.1177/0963721414550708>
- Kassam, K. S., & Mendes, W. B. (2013). The Effects of Measuring Emotion[U+202F]: Physiological Reactions to Emotional Situations Depend on whether Someone Is Asking. *PLoS ONE*, 8(6), e64959. <https://doi.org/10.1371/journal.pone.0064959>
- Kim, C.-Y., & Blake, R. (2005). Psychophysical magic[U+202F]: Rendering the visible 'invisible'. *Trends in Cognitive Sciences*, 9(8), 381-388. <https://doi.org/10.1016/j.tics.2005.06.012>
- Koch, C. (2004). *The Quest for Consciousness[U+202F]: A Neurobiological Approach*. Roberts and Company.
- Koch, C., Massimini, M., Boly, M., & Tononi, G. (2016). Neural correlates of consciousness[U+202F]: Progress and problems. *Nature Reviews Neuroscience*, 17(5), 307-321. <https://doi.org/10.1038/nrn.2016.22>
- Kovarski, K., Dos Reis, J., Chevais, C., Hamel, A., Makowski, D., & Sperduti, M. (2022). Movie editing influences spectators' time perception. *Scientific Reports*, 12(1), Art. 1. <https://doi.org/10.1038/s41598-022-23992-2>
- Kragel, P. A., Čeko, M., Theriault, J., Chen, D., Satpute, A. B., Wald, L. W., Lindquist, M. A., Feldman Barrett, L., & Wager, T. D. (2021). A human colliculus-pulvinar-amygdala pathway encodes negative emotion. *Neuron*, 109(15), 2404-2412.e5. <https://doi.org/10.1016/j.neuron.2021.06.001>
- Kragel, P. A., Reddan, M. C., LaBar, K. S., & Wager, T. D. (2019). Emotion schemas are embedded in the human visual system. *Science Advances*, 5(7), eaaw4358. <https://doi.org/10.1126/sciadv.aaw4358>
- Lähteenmäki, M., Hyönä, J., Koivisto, M., & Nummenmaa, L. (2015). Affective processing requires awareness. *Journal of Experimental Psychology: General*, 144(2), 339-365. <https://doi.org/10.1037/xge0000040>
- Lane, R. D., & Smith, R. (2021). Levels of Emotional Awareness[U+202F]: Theory and Measurement of a Socio-Emotional Skill. *Journal of Intelligence*, 9(3), 42. <https://doi.org/10.3390/jintelligence9030042>
- Lane, R. D., Weihs, K. L., Herring, A., Hishaw, A., & Smith, R. (2015). Affective agnosia[U+202F]: Expansion of the alexithymia construct and a new opportunity to integrate and extend Freud's legacy. *Neuroscience and Biobehavioral Reviews*, 55, 594-611. <https://doi.org/10.1016/j.neubiorev.2015.06.007>
- Lang, P. J., & Bradley, M. M. (2010). Emotion and the

- motivational brain. *Biological Psychology*, 84(3), 437-450. <https://doi.org/10.1016/j.biopsycho.2009.10.007>
- Lang, P. J., Greenwald, M. K., Bradley, M. M., & Hamm, A. O. (1993). Looking at pictures[U+202F]: Affective, facial, visceral, and behavioral reactions. *Psychophysiology*, 30(3), 261-273. doi.org/10.1111/j.1469-8986.1993.tb03352.x
- Lange, J., Dalege, J., Borsboom, D., van Kleef, G. A., & Fischer, A. H. (2020). Toward an Integrative Psychometric Model of Emotions. *Perspectives on Psychological Science*, 15(2), 444-468. <https://doi.org/10.1177/1745691619895057>
- Lanteaume, L., Khalfa, S., Régis, J., Marquis, P., Chauvel, P., & Bartolomei, F. (2007). Emotion Induction After Direct Intracerebral Stimulations of Human Amygdala. *Cerebral Cortex*, 17(6), 1307-1313. <https://doi.org/10.1093/cercor/bhl041>
- LeDoux, J. E. (2012). Rethinking the Emotional Brain. *Neuron*, 73(4), 653-676. <https://doi.org/10.1016/j.neuron.2012.02.004>
- LeDoux, J. E. (2019). The Deep History of Ourselves[U+202F]: How ancient microbes became conscious brains. Books on Tape.
- LeDoux, J. E. (2020a). How does the non-conscious become conscious? *Current Biology*, 30(5), R196-R199. <https://doi.org/10.1016/j.cub.2020.01.033>
- LeDoux, J. E. (2020b). Thoughtful feelings. *Current Biology*, 30(11), R619-R623. <https://doi.org/10.1016/j.cub.2020.04.012>
- LeDoux, J. E. (2021). What emotions might be like in other animals. *Current Biology*, 31(13), R824-R829. <https://doi.org/10.1016/j.cub.2021.05.005>
- Ledoux, J. E., & Brown, R. (2017). A higher-order theory of emotional consciousness. *Proceedings of the National Academy of Sciences of the United States of America*, 114(10), E2016-E2025. <https://doi.org/10.1073/pnas.1619316114>
- LeDoux, J. E., & Daw, N. D. (2018). Surviving threats[U+202F]: Neural circuit and computational implications of a new taxonomy of defensive behaviour. *Nature Reviews Neuroscience*, 19(5), 269-282. <https://doi.org/10.1038/nrn.2018.22>
- LeDoux, J. E., & Hofmann, S. G. (2018). The subjective experience of emotion[U+202F]: A fearful view. *Current Opinion in Behavioral Sciences*, 19(December), 67-72. <https://doi.org/10.1016/j.cobeha.2017.09.011>
- LeDoux, J. E., & Pine, D. S. (2016). Using neuroscience to help understand fear and anxiety[U+202F]: A two-system framework. *American Journal of Psychiatry*, 173(11), 1083-1093. <https://doi.org/10.1176/appi.ajp.2016.16030353>
- Lieberman, M. D. (2019). Boo[U+202F]! The consciousness problem in emotion. *Cognition and Emotion*, 33(1), 24-30. <https://doi.org/10.1080/02699931.2018.1515726>
- Lieberman, M. D., Eisenberger, N. I., Crockett, M. J., Tom, S. M., Pfeifer, J. H., & Way, B. M. (2007). Putting Feelings Into Words. *Psychological Science*, 18(5), 421-428. <https://doi.org/10.1111/j.1467-9280.2007.01916.x>
- Lin, H., Müller-Bardorff, M., Gathmann, B., Brieke, J., Mothes-Lasch, M., Bruchmann, M., Miltner, W. H. R., & Straube, T. (2020). Stimulus arousal drives amygdalar responses to emotional expressions across sensory modalities. *Scientific Reports*, 10(1), Art. 1. <https://doi.org/10.1038/s41598-020-58839-1>
- Lindquist, K. A., & Barrett, L. F. (2008). Emotional complexity. In *Handbook of emotions*, 3rd ed (p. 513-530). The Guilford Press. Mauss, I. B., Levenson, R. W., McCarter, L., Wilhelm, F. H., & Gross, J. J. (2005). The Tie That Binds? Coherence Among Emotion Experience, Behavior, and Physiology. *Emotion*, 5(2), 175-190. <https://doi.org/10.1037/1528-3542.5.2.175>
- McFadyen, J., Mattingley, J. B., & Garrido, M. I. (2019). An afferent white matter pathway from the pulvinar to the amygdala facilitates fear recognition. *ELife*, 8, e40766. <https://doi.org/10.7554/eLife.40766>
- Melloni, L., Mudrik, L., Pitts, M., & Koch, C. (2021). Making the hard problem of consciousness easier. *Science*, 372(6545), 911-912. <https://doi.org/10.1126/science.abj3259>
- Méndez, C. A., Celeghin, A., Diano, M., Orsenigo, D., Ocak, B., & Tamietto, M. (2022). A deep neural network model of the primate superior colliculus for emotion recognition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 377(1863), 20210512. <https://doi.org/10.1098/rstb.2021.0512>
- Méndez-Bértolo, C., Moratti, S., Toledano, R., Lopez-Sosa, F., Martínez-Alvarez, R., Mah, Y. H., Vuilleumier, P., Gil-Nagel, A., & Strange, B. A. (2016). A fast pathway for fear in human amygdala. *Nature Neuroscience*, 19(8), Art. 8. <https://doi.org/10.1038/nn.4324>
- Mendl, M., & Paul, E. S. (2020). Animal affect and decision-making. *Neuroscience & Biobehavioral Reviews*, 112, 144-163. <https://doi.org/10.1016/j.neubiorev.2020.01.025>
- Mennella, R., Bavard, S., Mentec, I., & Grèzes, J. (2022). Spontaneous instrumental avoidance learning in social contexts. *Scientific Reports*, 12(1), Art. 1. <https://doi.org/10.1038/s41598-022-22334-6>
- Mennella, R., Vilarem, E., & Grèzes, J. (2020). Rapid approach-avoidance responses to emotional displays reflect value-based decisions[U+202F]: Neural evidence from an EEG study. *NeuroImage*, 222, 117253. <https://doi.org/10.1016/j.neuroimage.2020.117253>
- Merker, B. (2007). Consciousness without a cerebral cortex[U+202F]: A challenge for neuroscience and medicine. *Behavioral and Brain Sciences*, 30(1), 63-81. <https://doi.org/10.1017/S0140525X07000891>
- Merker, B. (2013). The efference cascade, consciousness, and its self[U+202F]: Naturalizing the first person pivot of action control. *Frontiers in Psychology*, 4. <https://www.frontiersin.org/articles/10.3389/fpsyg.2013.00501>
- Mertens, G., & Engelhard, I. M. (2020). A systematic review and meta-analysis of the evidence for unaware fear conditioning. *Neuroscience and Biobehavioral Reviews*, 108, 254-268. <https://doi.org/10.1016/j.neubiorev.2019.11.012>
- Mikels, J. A., & Reuter-Lorenz, P. A. (2019). Affective Working Memory[U+202F]: An Integrative Psychological Construct. *Perspectives on psychological science[U+202F]: a journal of the Association for Psychological Science*, 14(4),

543-559. <https://doi.org/10.1177/1745691619837597>

Mitchell, D. G. V., & Greening, S. G. (2012). Conscious Perception of Emotional Stimuli[U+202F]: Brain Mechanisms. *The Neuroscientist*, 18(4), 386-398. <https://doi.org/10.1177/1073858411416515>

Mobbs, D., Adolphs, R., Fanselow, M. S., Barrett, L. F., LeDoux, J. E., Ressler, K., & Tye, K. M. (2019). Viewpoints[U+202F]: Approaches to defining and investigating fear. *Nature Neuroscience*, 22(8), 1205-1216. <https://doi.org/10.1038/s41593-019-0456-6>

Mobbs, D., Headley, D. B., Ding, W., & Dayan, P. (2020). Space, Time, and Fear[U+202F]: Survival Computations along Defensive Circuits. *Trends in Cognitive Sciences*, 24(3), 228-241. <https://doi.org/10.1016/j.tics.2019.12.016>

Moors, A. (2022). Demystifying Emotions[U+202F]: A Typology of Theories in Psychology and Philosophy (1re éd.). Cambridge University Press. <https://doi.org/10.1017/9781107588882>

Moors, A., Boddez, Y., & De Houwer, J. (2017). The Power of Goal-Directed Processes in the Causation of Emotional and Other Actions. *Emotion Review*, 9(4), 310-318. <https://doi.org/10.1177/1754073916669595>

Naccache, L. (2018). Why and how access consciousness can account for phenomenal consciousness. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 373(1755), 20170357. <https://doi.org/10.1098/rstb.2017.0357>

Nagel, T. (1974). What Is It Like to Be a Bat? *The Philosophical Review*. <https://doi.org/10.2307/2183914>

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know[U+202F]: Verbal reports on mental processes. *Psychological Review*, 84(3), 231-259. <https://doi.org/10.1037/0033-295X.84.3.231>

Northoff, G., & Lamme, V. (2020). Neural signs and mechanisms of consciousness[U+202F]: Is there a potential convergence of theories of consciousness in sight? *Neuroscience & Biobehavioral Reviews*, 118, 568-587. <https://doi.org/10.1016/j.neubiorev.2020.07.019>

Nummenmaa, L., Glerean, E., Hari, R., & Hietanen, J. K. (2014). Bodily maps of emotions. *Proceedings of the National Academy of Sciences of the United States of America*, 111(2), 646-651. <https://doi.org/10.1073/pnas.1321664111>

Oberauer, K., & Lewandowsky, S. (2019). Addressing the theory crisis in psychology. *Psychonomic Bulletin & Review*, 26(5), 1596-1618. [doi.org/10.3758/s13423-019-01645-2](https://doi.org/10.3758/s13423-019-01645-2)

Odegaard, B., Knight, R. T., & Lau, H. (2017). Should a Few Null Findings Falsify Prefrontal Theories of Conscious Perception? *The Journal of Neuroscience*, 37(40), 9593-9602. <https://doi.org/10.1523/JNEUROSCI.3217-16.2017>

Overgaard, M., & Sandberg, K. (2012). Kinds of access[U+202F]: Different methods for report reveal different kinds of metacognitive access. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1594), 1287-1296. <https://doi.org/10.1098/rstb.2011.0425>

Panksepp, J. (1998a). Affective neuroscience[U+202F]:

The foundations of human and animal emotions. In *New York[U+202F]: Oxford University Press*. Panksepp,

J. (1998b). The periconscious substrates of consciousness[U+202F]: Affective states and the evolutionary origins of the self. *Journal of Consciousness Studies*, 5, 566-582. Panksepp, J. (2007). Emotional feelings originate below

the neocortex[U+202F]: Toward a neurobiology of the soul. In *Behavioral and Brain Sciences* (Vol. 30, Numéro 1, p. 101-103). <https://doi.org/10.1017/S0140525X07001094>

Panksepp, J., Lane, R. D., Solms, M., & Smith, R. (2017). Reconciling cognitive and affective neuroscience perspectives on the brain basis of emotional experience. *Neuroscience and Biobehavioral Reviews*, 76(July), 187-215. <https://doi.org/10.1016/j.neubiorev.2016.09.010>

Paré, D., & Quirk, G. J. (2017). When scientific paradigms lead to tunnel vision[U+202F]: Lessons from the study of fear. *Npj Science of Learning*, 2(1), 6. <https://doi.org/10.1038/s41539-017-0007-4>

Parvizi, J., & Damasio, A. R. (2001). Consciousness and the brainstem. *Cognition*, 79(1), 135-160. [https://doi.org/10.1016/S0010-0277\(00\)00127-X](https://doi.org/10.1016/S0010-0277(00)00127-X)

Paul, E. S., Sher, S., Tamietto, M., Winkielman, P., & Mendl, M. T. (2020). Towards a comparative science of emotion[U+202F]: Affect and consciousness in humans and animals. *Neuroscience & Biobehavioral Reviews*, 108, 749-770. <https://doi.org/10.1016/j.neubiorev.2019.11.014>

Pavuluri, M., May, A., & 1 Pediatric Mood Disorders Program and Pediatric Brain Research and Intervention Center, Department of Psychiatry, College of Medicine, University of Illinois at Chicago, Chicago, IL 60608, USA; (2015). I Feel, Therefore, I am[U+202F]: The Insula and Its Role in Human Emotion, Cognition and the Sensory-Motor System. *AIMS Neuroscience*, 2(1), 18-27. <https://doi.org/10.3934/Neuroscience.2015.1.18>

Pessoa, L., Japee, S., Sturman, D., & Ungerleider, L. G. (2006). Target Visibility and Visual Awareness Modulate Amygdala Responses to Fearful Faces. *Cerebral Cortex*, 16(3), 366-375. <https://doi.org/10.1093/cercor/bhi115>

Prinz, J. J. (2004). Gut Reactions[U+202F]: A Perceptual Theory of Emotion. Oxford University Press. Prinz, J. J.

(Éd.). (2008). Précis of « Gut Reactions ». *Philosophy and Phenomenological Research*, 76(3), 707-711. Roberts,

I. D., & Hutcherson, C. A. (2019). Affect and decision making[U+202F]: Insights and predictions from computational models. *Trends in cognitive sciences*, 23(7), 602-614. <https://doi.org/10.1016/j.tics.2019.04.005>

Robinson, M. D., & Clore, G. L. (2002). Episodic and semantic knowledge in emotional self-report[U+202F]: Evidence for two judgment processes. *Journal of Personality and Social Psychology*, 83(1), 198-215. <https://doi.org/10.1037/0022-3514.83.1.198>

Rohr, M., & Wentura, D. (2021). Degree and Complexity of Non-conscious Emotional Information Processing—A Review of Masked Priming Studies. *Frontiers in Human Neuroscience*, 15, 689369. <https://doi.org/10.3389/fnhum.2021.689369>

Rosenthal, D. M. (2005). Consciousness and Mind. Oxford University Press UK. Rossion, B., de Gelder, B.,

Pourtois, G., Guérit, J.-M., & Weiskrantz, L. (2000).

- Early extrastriate activity without primary visual cortex in humans. *Neuroscience Letters*, 279(1), 25-28. [https://doi.org/10.1016/S0304-3940\(99\)00926-X](https://doi.org/10.1016/S0304-3940(99)00926-X)
- Rottevel, M., & Phaf, R. H. (2004). Automatic Affective Evaluation Does Not Automatically Predispose for Arm Flexion and Extension. *Emotion* (Washington, D.C.), 4, 156-172. <https://doi.org/10.1037/1528-3542.4.2.156>
- Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, 110(1), 145-172. <https://doi.org/10.1037/0033-295X.110.1.145>
- Saarimäki, H. (2021). Naturalistic Stimuli in Affective Neuroimaging[U+202F]: A Review. *Frontiers in Human Neuroscience*, 15, 675068. <https://doi.org/10.3389/fnhum.2021.675068>
- Saarimäki, H., Ejtehadian, L. F., Glerean, E., Jääskeläinen, I. P., Vuilleumier, P., Sams, M., & Nummenmaa, L. (2018). Distributed affective space represents multiple emotion categories across the human brain. *Social Cognitive and Affective Neuroscience*, 13(5), 471-482. <https://doi.org/10.1093/scan/nsy018>
- Saarimäki, H., Gotsopoulos, A., Jääskeläinen, I. P., Lampinen, J., Vuilleumier, P., Hari, R., Sams, M., & Nummenmaa, L. (2016). Discrete Neural Signatures of Basic Emotions. *Cerebral Cortex*, 26(6), 2563-2573. <https://doi.org/10.1093/cercor/bhv086>
- Scheel, A. M., Tiokhin, L., Isager, P. M., & Lakens, D. (2021). Why Hypothesis Testers Should Spend Less Time Testing Hypotheses. *Perspectives on Psychological Science*, 16(4), 744-755. <https://doi.org/10.1177/1745691620966795>
- Schlund, M. W., Brewer, A. T., Magee, S. K., Richman, D. M., Solomon, S., Ludlum, M., & Dymond, S. (2016). The tipping point[U+202F]: Value differences and parallel dorsal-ventral frontal circuits gating human approach-avoidance behavior. *NeuroImage*, 136, 94-105. <https://doi.org/10.1016/j.neuroimage.2016.04.070>
- Selimbeyoglu, A. (2010). Electrical stimulation of the human brain[U+202F]: Perceptual and behavioral phenomena reported in the old and new literature. *Frontiers in Human Neuroscience*. <https://doi.org/10.3389/fnhum.2010.00046>
- Seth, A. K. (2013). Interoceptive inference, emotion, and the embodied self. *Trends in Cognitive Sciences*, 17(11), 565-573. <https://doi.org/10.1016/j.tics.2013.09.007>
- Seth, A. K. (2015). The Cybernetic Bayesian BrainThe Cybernetic Bayesian Brain[U+202F]: From Interoceptive Inference to Sensorimotor Contingencies. *Open MIND*. <https://doi.org/10.15502/9783958570108>
- Seth, A. K., & Bayne, T. (2022). Theories of consciousness. *Nature Reviews Neuroscience*, 23(7), Art. 7. <https://doi.org/10.1038/s41583-022-00587-4>
- Seth, A. K., & Critchley, H. D. (2013). Extending predictive processing to the body[U+202F]: Emotion as interoceptive inference. *Behavioral and Brain Sciences*, 36(3), 227-228. <https://doi.org/10.1017/S0140525X12002270>
- Shine, J. M. (2022). Adaptively navigating affordance landscapes[U+202F]: How interactions between the superior colliculus and thalamus coordinate complex, adaptive behaviour. *Neuroscience & Biobehavioral Reviews*, 143, 104921. <https://doi.org/10.1016/j.neubiorev.2022.104921>
- Smirnov, D., Saarimäki, H., Glerean, E., Hari, R., Sams, M., & Nummenmaa, L. (2019). Emotions amplify speaker-listener neural alignment. *Human Brain Mapping*, 40(16), 4777-4788. <https://doi.org/10.1002/hbm.24736>
- Smith, R., Killgore, W. D. S., & Lane, R. D. (2018). The structure of emotional experience and its relation to trait emotional awareness[U+202F]: A theoretical review. *Emotion*, 18(5), 670-692. <https://doi.org/10.1037/emo0000376>
- Smith, R., & Lane, R. D. (2015). The neural basis of one's own conscious and unconscious emotional states. *Neuroscience & Biobehavioral Reviews*, 57, 1-29. <https://doi.org/10.1016/j.neubiorev.2015.08.003>
- Smith, R., & Lane, R. D. (2016). Unconscious emotion[U+202F]: A cognitive neuroscientific perspective. *Neuroscience & Biobehavioral Reviews*, 69, 216-238. <https://doi.org/10.1016/j.neubiorev.2016.08.013>
- Smith, R., Lane, R. D., Alkozei, A., Bao, J., Smith, C., Sanova, A., Nettles, M., & Killgore, W. D. S. (2018). The role of medial prefrontal cortex in the working memory maintenance of one's own emotional responses. *Scientific Reports*, 8(1), Art. 1. <https://doi.org/10.1038/s41598-018-21896-8>
- Smith, R., Lane, R. D., Parr, T., & Friston, K. J. (2019). Neurocomputational mechanisms underlying emotional awareness[U+202F]: Insights afforded by deep active inference and their potential clinical relevance. *Neuroscience & Biobehavioral Reviews*, 107, 473-491. <https://doi.org/10.1016/j.neubiorev.2019.09.002>
- Smith, R., Parr, T., & Friston, K. J. (2019). Simulating Emotions[U+202F]: An Active Inference Model of Emotional State Inference and Emotion Concept Learning. *Frontiers in Psychology*, 10. <https://www.frontiersin.org/articles/10.3389/fpsyg.2019.02844>
- Sun, B., Zeng, X., Chen, X., Zhao, J., & Fu, S. (2023). Neural correlates of conscious processing of emotional faces[U+202F]: Evidence from event-related potentials. *Neuropsychologia*, 108478. <https://doi.org/10.1016/j.neuropsychologia.2023.108478>
- Tamietto, M., & de Gelder, B. (2010). Neural bases of the non-conscious perception of emotional signals. *Nature Reviews Neuroscience*, 11(10), 697-709. <https://doi.org/10.1038/nrn2889>
- Taschereau-Dumouchel, V., Kawato, M., & Lau, H. (2020). Multivoxel pattern analysis reveals dissociations between subjective fear and its physiological correlates. *Molecular Psychiatry*, 25(10), 2342-2354. <https://doi.org/10.1038/s41380-019-0520-3>
- Taschereau-Dumouchel, V., Michel, M., Lau, H., Hofmann, S. G., & LeDoux, J. E. (2022). Putting the “mental” back in “mental disorders” [U+202F]: A perspective from research on fear and anxiety. *Molecular Psychiatry*. <https://doi.org/10.1038/s41380-021-01395-5>
- Tooley, M. D., Carmel, D., Chapman, A., & Grimshaw, G. M. (2017). Dissociating the physiological components of unconscious emotional responses. *Neuroscience of Consciousness*, 2017(1), nix021. <https://doi.org/10.1093/nc/nix021>
- Treit, D., & Berridge, K. C. (1990). A comparison of benzodiazepine, serotonin, and dopamine agents in the taste-reactivity paradigm. *Pharmacology, Biochemistry and Behavior*. [https://doi.org/10.1016/0091-3057\(90\)90011-6](https://doi.org/10.1016/0091-3057(90)90011-6)



- Tsikandilakis, M., Bali, P., Yu, Z., Madan, C., Derfuss, J., Chapman, P., & Groeger, J. (2021). Individual conscious and unconscious perception of emotion[U+202F]: Theory, methodology and applications. *Consciousness and Cognition*, 94, 103172. <https://doi.org/10.1016/j.concog.2021.103172>
- Tsuchiya, N., Wilke, M., Frässle, S., & Lamme, V. A. F. (2015). No-Report Paradigms[U+202F]: Extracting the True Neural Correlates of Consciousness. *Trends in Cognitive Sciences*, 19(12), 757-770. <https://doi.org/10.1016/j.tics.2015.10.002>
- Vaidya, A. R., Pujara, M. S., Petrides, M., Murray, E. A., & Fellows, L. K. (2019). Lesion Studies in Contemporary Neuroscience. *Trends in Cognitive Sciences*, 23(8), 653-671. <https://doi.org/10.1016/j.tics.2019.05.009>
- Vale, R., Evans, D. A., & Branco, T. (2017). Rapid Spatial Learning Controls Instinctive Defensive Behavior in Mice. *Current Biology*, 27(9), 1342-1349. <https://doi.org/10.1016/j.cub.2017.03.031>
- van der Ploeg, M. M., Brosschot, J. F., Versluis, A., & Verkuil, B. (2017). Peripheral physiological responses to subliminally presented negative affective stimuli[U+202F]: A systematic review. *Biological Psychology*, 129, 131-153. <https://doi.org/10.1016/j.biopsycho.2017.08.051>
- Vetter, P., Badde, S., Phelps, E. A., & Carrasco, M. (2019). Emotional faces guide the eyes in the absence of awareness. *ELife*, 8, e43467. <https://doi.org/10.7554/eLife.43467>
- Vilarem, E., Armony, J. L., & Grèzes, J. (2020). Action opportunities modulate attention allocation under social threat. *Emotion (Washington, D.C.)*, 20(5), 890-903. <https://doi.org/10.1037/emo0000598>
- Wieloch, T., & Nikolich, K. (2006). Mechanisms of neural plasticity following brain injury. *Current Opinion in Neurobiology*, 16(3), 258-264. <https://doi.org/10.1016/j.conb.2006.05.011>
- Williams, L. M., Liddell, B. J., Kemp, A. H., Bryant, R. A., Meares, R. A., Peduto, A. S., & Gordon, E. (2006). Amygdala-prefrontal dissociation of subliminal and supraliminal fear. *Human Brain Mapping*, 27(8), 652-661. <https://doi.org/10.1002/hbm.20208>
- Winkielman, P., & Berridge, K. C. (2004). Unconscious Emotion. *Current Directions in Psychological Science*, 13(3), 120-123. <https://doi.org/10.1111/j.0963-7214.2004.00288.x>
- Winkielman, P., Berridge, K. C., & Wilbarger, J. L. (2005). Unconscious affective reactions to masked happy versus angry faces influence consumption behavior and judgments of value. *Personality and Social Psychology Bulletin*, 31(1), 121-135. <https://doi.org/10.1177/0146167204271309>
- Winkielman, P., & Gogolushko, Y. (2018). Influence of Suboptimally and Optimally Presented Affective Pictures and Words on Consumption-Related Behavior. *Frontiers in Psychology*, 8. <https://www.frontiersin.org/articles/10.3389/fpsyg.2017.02261>
- Yang, E., Zald, D. H., & Blake, R. (2007). Fearful expressions gain preferential access to awareness during continuous flash suppression. *Emotion*, 7(4), 882-886. <https://doi.org/10.1037/1528-3542.7.4.882>
- Yarkoni, T., Poldrack, R. A., Van Essen, D. C., & Wager, T. D. (2010). Cognitive neuroscience 2.0[U+202F]: Building a cumulative science of human brain function. *Trends in Cognitive Sciences*, 14(11), 489-496. <https://doi.org/10.1016/j.tics.2010.08.004>
- Zaki, J., Davis, J. I., & Ochsner, K. N. (2012). Overlapping activity in anterior insula during interoception and emotional experience. *NeuroImage*, 62(1), 493-499. <https://doi.org/10.1016/j.neuroimage.2012.05.012>