

Boo! The consciousness problem in emotion

Matthew D. Lieberman

To cite this article: Matthew D. Lieberman (2018): Boo! The consciousness problem in emotion, *Cognition and Emotion*, DOI: [10.1080/02699931.2018.1515726](https://doi.org/10.1080/02699931.2018.1515726)

To link to this article: <https://doi.org/10.1080/02699931.2018.1515726>



Published online: 28 Aug 2018.



Submit your article to this journal [↗](#)



Article views: 104



View Crossmark data [↗](#)



Boo! The consciousness problem in emotion

Matthew D. Lieberman

UCLA Psychology Department, University of California, Los Angeles, CA, USA

ABSTRACT

This paper will examine the conscious aspects of emotion (i.e. emotional experience), arguably the defining features of emotion. I will argue that emotion IS emotional experience and, consequently, that emotion researchers rarely study emotion itself. I will suggest a research agenda for examining the conscious aspects of emotion and end with a consideration of appraisal theory and how it can be made more relevant to the study of emotion by treating appraisals as components of a pre-reflective perceptual process rather than as causal antecedents of a cognitive process that can be self-reported on.

ARTICLE HISTORY

Received 24 May 2018
Revised 7 August 2018
Accepted 8 August 2018

KEYWORDS

Experience; consciousness; awareness

In the mid-1990s, Jeffrey Gray, one of the grandfathers of affective neuroscience, gave a talk at Harvard University. Most of the talk focused on his foundational work identifying the neural bases of anxiety. However, he concluded his talk by suggesting that psychologists had been neglecting consciousness for too long and that try as they might ignore it, it will always eventually jump out and say “Boo!” because it is the central feature of our psychological lives. I believe emotion research has a consciousness problem and this article is an attempt to offer up my own “Boo!” on these issues.

Just as *Cognition & Emotion* is turning 30 (congratulations!), this is a milestone for me as well. Thirty years ago, I started college and became a psychology major at Rutgers University. I was very excited to start studying subjects like emotion and consciousness because these are real attention-getters in teenagedom. The range, intensity, and complexity of emotions became strikingly apparent in these years. My emotional experiences often felt like bolts of lightning – brief hallucinogenic episodes that seemed to come out of nowhere, fundamentally altering how much of the world looked and felt to me and leading me to act in ways that I could scarcely recognise as my own once the episode subsided. As unnerving as some of those experiences were, it is difficult to imagine

anything that was as captivating as emotional experiences.

It was with this mindset that I took my first psychology courses. I was more than a little disappointed to discover that anything related to consciousness was largely *verboten*, a taboo topic to be avoided. I could not fathom why psychologists would avoid the singular topic, consciousness that made their field unique from all others – but avoid they did. Skinner and the behaviourists replaced consciousness with behaviour. Cognitive psychologists replaced consciousness with cognitive operations. Neuroscientists then replaced the study of consciousness with the study of neural activity. And these days, everyone replaces everything with computations. To be clear, these are all worthy contributions to psychology, but I could not understand why they each took their place of prominence at the expense of conscious experience receding further into the background of the field.

Emotion is emotional experience

In some areas of research, the consequences of ignoring conscious experience are not so grave. Understanding how humans perform math may not depend so much on accounting for conscious experience. Adding “6 + 3” does not phenomenologically

feel like much. I can imagine a relatively complete account here with a sole reliance on cognitive and computational processes. But relative to all the other topics of psychology, the emotional domain is where sidestepping consciousness does the greatest harm to the thing studied.

Emotion researchers typically define emotion in terms of various channels of emotion. They tend to focus on experiential, behavioural, expressive, physiological, and neural correlates of emotion. The primary finding associated with all of these channels of research is that they often do not hang together in the way one might expect in an *a priori* way (Mauss, Levenson, McCarter, Wilhelm, & Gross, 2005). In other words, all the channels of emotion do not correlate strongly with one another.

But here is this thing. Emotional experience is not a channel of emotion, one part among many that make up the emotion conglomerate. Emotional experience *is* emotion. Everything else – the behavioural, expressive, physiological and neural channels are just correlates of emotion, not emotion itself.

Do androids dream of emotional experience?

It may be heretical to say that non-experiential channels of emotion are not emotion and that only experience is, but that does not make it untrue. Here are two simple proofs.

First, of the five channels of emotion (experience, behaviour, expression, physiology, and neural), each of the four non-experiential channels can occur in the absence of true emotion. We can easily imagine emotional expressions and behaviour occurring in the absence of true emotion; we call this acting. The same physiology activated in different emotions can be set off for other reasons. When it comes to the brain it is a bit trickier. However, I strongly suspect that if we could activate the specific neurons associated with particular emotions but only those neurons, we could get something valenced but not true emotions like a modern-day Marañon-type study (1924). If we have not already activated the neurons associated with a particular situation and a particular set of personal goals (i.e. the world and our current relation to it), I suspect the emotional-neuronal activations alone would not hold and generate an emotion. However, what we know for sure is that if someone has an emotional experience, they are having an emotion and if someone is having an

emotion, they are having an emotional experience. No other channel of emotion on its own guarantees that an emotion has occurred. Only emotional experience does because that is what emotion is.

Our second proof focuses on androids. Whether its Delores from *Westworld* or your favourite android from *Blade Runner*, *The Terminator*, *I Robot*, or *Ex Machina*, we can consider whether they have emotions. Delores is perhaps the most interesting case because *Westworld* shows how fully biologically realised these androids are. I suspect the androids might even have startle and galvanic skin responses. We might wonder whether these androids actually have emotions. But the answer comes down to one thing and only one thing – do they have emotional experience? They may have the most exquisitely contextually sensitive emotional expressions and actions and they can have all the physiology of emotion, but if they do not have the experience of emotion then it is just an incredibly sophisticated simulation – the appearance of emotion, but not actual emotion. All emotion is emotional experience and everything else is a correlate.

What do emotion researchers study?

From the above perspective, any research would need to focus on emotional experience in order to qualify as a study of emotion, we can ask what do emotion researchers actually tend to study and value. To examine this, I looked to the articles in *Cognition & Emotion*. I identified the 100 articles that had been most cited from the history of the journal. This would seem to give a good mix of the kinds of research that are both done most often but also cited well enough to indicate that the field cares about these topics. Each study was classified according to a number of categories that emerged and the same study could count in multiple categories. For instance, a study of emotion regulation in attention deficit hyperactivity disorder would count in both the emotion regulation and clinical categories.

I defined the emotional experience category broadly including studies of emotional imagery, meta-cognitive awareness of one's own emotion, mindful focus on one's emotion, and anything else that seemed to emphasise the conscious aspects of emotional experience. Things of note that were excluded: emotion inductions for the purpose of examining other channels of emotion and attention shift studies (e.g. biased attention towards threatening

stimuli). In these attention studies, attention is not being examined in terms of how it is experienced by participants but rather is a cognitive operation referring to sampling of information from parts of the environment.

Of the 100 most cited articles from *Cognition & Emotion*, only two were unambiguously about emotional experience (Barrett, Gross, Christensen, & Benvenuto, 2001; Gohm & Clore, 2002). A third focused on peaks and ends of emotional experiences as predictors of memory for the emotion (Fredrickson, 2000). If we include this third paper, then this means that only 3 of the 100 top cited papers from this journal (3%) have focused on emotional experience. Among the different categories, emotional experience came in 13th place among all topics.

Given that I am making a claim about emotion researchers and not just emotion researchers who publish in *Cognition & Emotion*, I broadened the analysis to include other journals. I added four other journals that also focus on emotion research including *Emotion*, *Emotion Review*, *Cognitive Affective & Behavioral Neuroscience* (CABN) and *Social Cognitive and Affective Neuroscience* (SCAN) yielding a total of 500 top cited emotion studies.

The results were the same. Of the 500 studies, only 14 focused on conscious aspects of emotion, which is a paltry 2.8% (see Figure 1). I want to be clear that I am not claiming that emotion researchers are not studying fascinating and important topics – topics in which emotion is clearly playing a part. The top five topics focused on uncovering the neural correlates of emotion, understanding how we recognise

emotional expressions in others, how social and clinical factors contribute to emotion (separate topics), and the different methods by which we regulate our emotions.

I am not suggesting that we should spend any less time studying those critical topics. I am instead suggesting that *emotion researchers almost never study emotion itself* as defined in terms of the study of emotional experience. I am suggesting we add something to the study of emotion – the actual study of emotion as an end itself. I am suggesting we create a much-needed *science of emotional experience*.

Why anything?

A few decades ago, the physicist Stephen Hawking shared his humble professional goals, “My goal is simple. It is a complete understanding of the universe, why it is as it is and why it exists at all” (Boslough, 1985). This is the first of the big three questions – the most fundamental, important questions there are. Hawking’s is the first: *Why does anything exist and why is the physical universe the way it is?* The second: *Why does life exist and why are living forms the way they are?* The third gets us closer to home: *Why does consciousness exist and why is consciousness the way it is?* If there is to be a science of emotional experience, it would surely start as a subheading to this third question.

Thus, the fundamental question for a science of emotional experience is: *Why do emotions feel like anything?* Why are not emotions just nonconscious

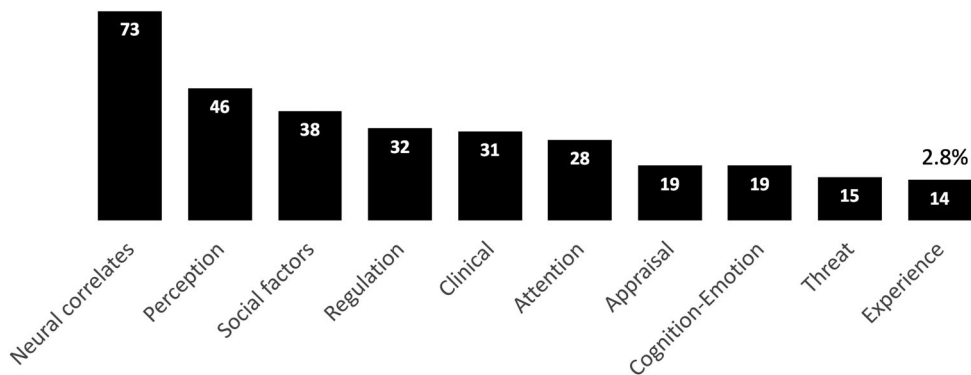


Figure 1. What do emotion researchers study? Based on the 500 most cited articles from each of five emotion-focused journals (100 each from *Cognition & Emotion*, *Emotion*, *Emotion Review*, *Social Cognitive and Affective Neuroscience*, and *Cognitive, Affective, and Behavioral Neuroscience*) what topics were studied most frequently.

Note: The same paper can be counted for more than one topic. Also, categories with fewer papers than “experience” are not included in this figure.

computations that integrate situational information with representations of current goals to produce appropriate actions, all without the muss and fuss of consciousness? We have lots of systems and processes in the human brain and body that operate just fine without any conscious components. Why is emotional experience one of the exceptions to the rule that almost everything in the universe is not conscious?

I have no answer for this question – none whatsoever. There are various psychological theories of consciousness (Baars, 2005; Tononi & Edelman, 1998) that might be a starting point, yet I do not have confidence that these theories will help us answer this question – not even my own (Lieberman, 2009). That does not mean the question is not worth asking. In psychology, there is an unwritten rule that we only ask the questions that there is a method for answering, either now or on the horizon. In other fields, this is not always the case. In mathematics, there are certain famous math problems that were formulated decades or even centuries ago that are still unsolved. Fermat's Last Theorem was proposed in 1637 and only solved in 1994. Perhaps psychology ought to incentivise the recognition and clear articulation of important problems that we do not yet know how to solve.

Why this?

The second major question for a science of emotional experience is subtly different. Given that we know that emotions feel like something, why do they feel like this rather than any of the other ways they might feel? Major theories of emotion do not really speak to this issue at all.

Basic emotions theories (Ekman, 1992) typically describe humans as having six or so distinct basic emotions that each follow their own stereotypical path in different channels of emotion that run to completion once activated. There is something right about this model with respect to emotional experience in that these emotions represent the most distinct poles of emotional experience. With the possible exception of surprise, the other emotions are each experienced as different from one another as seems possible. Yet basic emotion theory tells us nothing about why the different emotions feel maximally distinct from one another.

The constructivist account of emotion (Barrett, 2006) has seemingly supplanted the basic emotions account as the dominant account of how emotions are generated. This model suggests that emotions

result from a combination of core affect (e.g. valence and arousal) combined with linguistic processes that help the core affect take the shape of a particular emotional state. As compelling as this account is, it does not describe how emotions feel as they are being generated. My emotions do not feel like core affect plus linguistic differentiation. Any complete account of emotional experience will need to make sense of why emotions might be generated one way and yet feel as if they are generated in a different way.

Emotions as a mode of seeing

The third question for a science of emotional experience focuses on its regularities: *What are the structural features of emotional experience*. Unlike the previous two questions, some progress has been made on this one. For me, the single best work in this area is *The Emotions* (1986) by Nico Frijda, which devotes a chapter to the structure of emotional experience and builds off of the important but obscure theory from Sartre (1939).

In this work, Frijda identifies significant large-scale features of emotional experience but highlights several smaller but critical features of emotional experience. For instance, Frijda points out an emotion typically feels like something that has happened to us rather than something we have done or initiated. He also highlights the sense of eternity associated with emotions. Even though emotions are relatively transient events, internally, the state of affairs that caused them and the emotional state itself feels, in the moment, like it will last forever. When in a state of grief, we cannot conceive that conditions in the world will ever change that current state of feeling.

More broadly, Frijda noted that “Emotional experience ... is primarily a perception: a mode of appearance of the situation ... Emotional experience is perception of horrible objects, insupportable people, oppressive events” (1986, p. 188). Sartre (1939) too described emotion as a kind of perceptual experience describing it as a “transformation of the world” (p. 58) that creates “a world which is cruel, terrible, gloomy, joyful” (p. 80).

Reappraising appraisal theory

Frijda and Sartre's insight of emotional experience as a way of seeing is remarkably useful as we consider appraisal theories, which are probably the closest

things we have to a mainstream theory of emotional experience (see also Barrett, 2006; Barrett, Mesquita, Ochsner, & Gross, 2007). Appraisal theories (Frijda, 1993; Smith & Ellsworth, 1985) focus on the interpretations people have of their circumstances as generating emotions. Depending on how two people have invested differently in the stock market (e.g. investing versus not investing in tech stocks), the exact same situation (i.e. a major drop in tech stocks) could be interpreted differently by each, leading to very different emotions (i.e. anger vs. relief). Each specific appraisal theory lists a number of appraisal dimensions like pleasantness, anticipated effort, certainty, and responsibility. Different models varied from just a handful of dimensions to 16 or more.

Unfortunately, there is a wide gap between what most appraisal theorists have written and what “Appraisal Theory”, writ large, is understood to claim. Ever since Zajonc’s (1984) critique of appraisal theory, many have assumed that appraisal theory claims that appraisals are conscious and causal antecedents of emotion. This combination of claims is utterly implausible on its face because there is no way that people could serially go through all the possible appraisal dimensions consciously, prior to the generation of an emotional state.

Despite Zajonc’s sticky characterisation, I have been unable to find a single appraisal theorist who ever claimed that appraisals are exclusively conscious (Moors, 2010). The closest case would probably be Lazarus and Folkman (1984). After acknowledging something like automatic appraisal processes, they go to say that “our emphasis is much more on complex, meaning-related cognitive activity. Appraisals go far beyond immediate and indeliberate cognitive–affective responses” (p. 26). Other appraisal theories tend to be more balanced in their emphasis, but given that Lazarus and Folkman have been cited over 50,000 times, it may have had disproportionate influence.

Reflective and pre-reflective emotion processes

Since Zajonc (1984), the appraisal debate has been framed in terms of a false dichotomy between conscious and unconscious emotion processes, with the former being deliberative and verbalisable and the latter being automatic and outside of awareness. A third possibility is that appraisal processes are automatic, but also conscious. I have argued that there

are at least two types of conscious processes, reflective and pre-reflective, and the latter of these share the computational features of automatic processes while collectively being the basis of our immediate conscious experience (Lieberman, 2009; Lieberman, Gaunt, Gilbert, & Trope, 2002; Satpute & Lieberman, 2006). These processes were originally described (Lieberman et al., 2002) as producing “the stream of consciousness that we experience as the real world – not just the objects of the real world, but also the semantic and affective associations” (p. 235).

I have used this distinction to make sense of attribution processes in social cognition (Lieberman, 2009) that bear a resemblance to appraisal processes in emotion: “When we observe one person shoving another, this is experienced as an aggressive act, with the aggressiveness experienced as objectively out there in the world, even though the aggressiveness is constructed psychologically based on a number of characteristics separate from the act itself” (p. 301). Here, the aggressiveness and its characteristics (i.e. being negative, unelicited, intentional), which are literally seen, is inherent to our conscious experience of the act, but it also something that does not require reflective thought to generate.

Although appraisal theorists have said the right things about appraisal processes not all involving reflective conscious thought, they have almost always studied appraisals through self-report which is a decidedly reflective process. While reflective and pre-reflective processes can often perform the same tasks (Lieberman et al., 2002), they do so through mechanisms that are qualitatively different. Studying the reflective variant of a process when the process of interest is pre-reflective is counterproductive.

Frijda (1986) provided a solution to these issues that parallels the reflective/pre-reflective distinction. As indicated above, Frijda identified emotion as a mode of seeing and vision is a paradigmatic pre-reflective process with countless simultaneous automatic processes giving rise to not-yet-reflected upon experience. If emotion is like seeing, then it follows that we ought to think of appraisal processes as analogous to perceptual processes. Processing of valence, arousal, controllability, and agency should be thought about and studied the same way that we study motion, colour, texture and shape processing. In both cases, these processes are literally the lenses through which we see the world. Dimensions like agency or motion can be reflected upon, but typically are not and do not need to be in order to contribute to

the experience of emotional or visual seeing, respectively. And when people do reflect on these things the contents of these reflections are not necessarily isomorphic with the pre-reflective originals, and can even alter them. We have seen in a number of affect labelling studies that when people put negative feelings into words, ongoing affect is dampened in subjective, physiological, behavioural, and neural channels of emotion (Torre & Lieberman, 2018). Additionally, once we move to thinking of emotional experience as a kind of seeing, the causal question drops out. We do not really talk about what causes seeing. We do not say that motion, colour, and shape processing cause seeing – rather they are the component processes of seeing. Similarly, appraisals are component processes of emotional seeing (Ellsworth, 2013), rather than a cause *per se*.

Experience thermometers

What we really want in an appraisal theory is an experience thermometer – something that we can put in a person's head and read out what the person is pre-reflectively experiencing and how various appraisals come together to help construct that experience. Although self-reporting on appraisals after the fact was a natural starting point for research, if we take the "emotion as seeing" approach, we should study appraisals differently going forward. Even though I believe appraisal processes are often conscious in the sense of being pre-reflective processes, the words we use to characterise our pre-reflective experience are never a complete characterisation of that experience. All but the most poetic among us are likely to provide a simple label or a few words to characterise what in fact is a temporary massive overhaul of the way we are feeling. I might label two experiences as "happy", but the phenomenological feel of those two experiences might be radically different. Labelling them both as happy tells us something about how they felt, while providing just the tip of the iceberg with most the experience left unsaid.

Vision researchers avoid this problem by almost never asking participants to reflect on what they have seen. At most, they ask participants to indicate when they have noticed something being present or changing. Vision researchers manipulate dimensions of motion and colour in order to see which aspects of change the visual system is sensitive to.

Appraisals ought to be studied the same way. As a neuroscientist, the approach I am taking is to



Figure 2. Neural reconstruction of perceived stimuli. Schoenmakers et al. (2013) showed individuals various letters and used the neural response in the visual cortex to decode and reconstruct the originally perceived letters.

independently identify multivariate neural patterns that are sensitive to particular appraisal dimensions like agency and effort estimation by showing a series of non-emotional vignettes that manipulate a single appraisal dimension. Once we have these multivariate patterns, we can use them as biomarkers of when these appraisal dimensions are engaged (in the same way that activity in region MT is an indicator that motion processing has occurred). With a set of emotional appraisal indicators, we can then observe the ways in which these processes occur before, during, and after emotional experiences and how their interactions given rise to emotional experiences of particular kinds. To be clear, this would not be replacing experience with neural activity. I would still vastly prefer to have a real experience thermometer. However, the neural activity can be characterised in such a way that it can serve as an indicator of particular appraisal processes that together serve as core components of the emotional experience, without invoking reflective processes (i.e. self-report) that might be poor indicators or even alter the underlying pre-reflective emotional processes.

Currently, a similar enterprise is underway within visual perception. Several studies have now shown that a visual perception can be reconstructed (see Figure 2) from the activity of different parts of visual cortex (Schoenmakers, Barth, Heskes, & van Gerven, 2013). It is still early days in this endeavour, but remarkable progress has been made. If the same thing could be accomplished in the domain of emotional experience, then we might be able to finally make good on Frijda's claim that "It should be possible to present, for every emotion, a corresponding description of how the world looks to the subject" (1986, p. 196).

Boo!

The last thirty years have seen remarkable progress in many things related to emotion and how emotion

relates to many things. That said, it is important to recognise that emotion is emotional experience and thus emotional experience should be at the centre of the emotion research enterprise. This move requires admitting that we do not have the answers to all of the big questions or even an obvious way to begin to answer them. This is ok. As long we continue to let every generation know these are the big questions and that we deeply value any serious attempts to take them on, then perhaps by the 60th anniversary of *Cognition & Emotion*, someone far smarter than me will make some real progress on them.

Acknowledgements

The author would like to acknowledge Naomi Eisenberger for comments on an earlier draft of this manuscript.

Disclosure statement

No potential conflict of interest was reported by the author.

References

- Baars, B. J. (2005). Global workspace theory of consciousness: Toward a cognitive neuroscience of human experience. *Progress in Brain Research*, 150, 45–53.
- Barrett, L. F. (2006). Solving the emotion paradox: Categorization and the experience of emotion. *Personality and Social Psychology Review*, 10, 20–46.
- Barrett, L. F., Gross, J., Christensen, T. C., & Benvenuto, M. (2001). Knowing what you're feeling and knowing what to do about it: Mapping the relation between emotion differentiation and emotion regulation. *Cognition & Emotion*, 15, 713–724.
- Barrett, L. F., Mesquita, B., Ochsner, K. N., & Gross, J. J. (2007). The experience of emotion. *Annual Review of Psychology*, 58, 373–403.
- Boslough, J. (1985). *Stephen Hawking's universe*. New York: Quill.
- Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion*, 6, 169–200.
- Ellsworth, P. C. (2013). Appraisal theory: Old and new questions. *Emotion Review*, 5, 125–131.
- Fredrickson, B. L. (2000). Extracting meaning from past affective experiences: The importance of peaks, ends, and specific emotions. *Cognition & Emotion*, 14, 577–606.
- Frijda, N. H. (1986). *The emotions*. Cambridge: Cambridge University Press.
- Frijda, N. H. (1993). The place of appraisal in emotion. *Cognition & Emotion*, 7, 357–387.
- Gohm, C. L., & Clore, G. L. (2002). Four latent traits of emotional experience and their involvement in well-being, coping, and attributional style. *Cognition & Emotion*, 16, 495–518.
- Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping*. New York: Springer.
- Lieberman, M. D. (2009). What zombies can't do: A social cognitive neuroscience approach to the irreducibility of reflective consciousness. In J. Evans, & K. Frankish (Eds.), *In two minds: Dual process and beyond* (pp. 293–316). New York: Oxford University Press.
- Lieberman, M. D., Gaunt, R., Gilbert, D. T., & Trope, Y. (2002). Reflection and reflexion: A social cognitive neuroscience approach to attributional inference. *Advances in Experimental Social Psychology*, 34, 199–249.
- Marañón, G. (1924). Contribution a l'etude de l'action emotive de l'Adrenaline. *Revue Française D'Endocrinologie*, 2, 301.
- Maus, 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.
- Moors, A. (2010). Automatic constructive appraisal as a candidate cause of emotion. *Emotion Review*, 2(2), 139–156.
- Sartre, J. P. (1939). *The emotions: A sketch of a theory*. New York: Citadel Press.
- Satpute, A. B., & Lieberman, M. D. (2006). Integrating automatic and controlled processing into neurocognitive models of social cognition. *Brain Research*, 1079, 86–97.
- Schoenmakers, S., Barth, M., Heskes, T., & van Gerven, M. (2013). Linear reconstruction of perceived images from human brain activity. *NeuroImage*, 83, 951–961.
- Smith, C. A., & Ellsworth, P. C. (1985). Patterns of cognitive appraisal in emotion. *Journal of Personality and Social Psychology*, 48, 813–838.
- Tononi, G., & Edelman, G. M. (1998). Consciousness and complexity. *Science*, 282, 1846–1851.
- Torre, J. B., & Lieberman, M. D. (2018). Putting feelings into words: Affect labeling as implicit emotion regulation. *Emotion Review*, 10, 116–124.
- Zajonc, R. B. (1984). On the primacy of affect. *American Psychologist*, 39, 117–123.