

The Case for Unfelt Feelings

Katherine Tullmann

Northern Arizona University

She sat silently meditating, in a fixed attitude, for a few minutes. A few minutes were sufficient for making her acquainted with her own heart. A mind like hers, once opening to suspicion, made rapid progress. She touched — she admitted — she acknowledged the whole truth. Why was it so worse Harriet should be in love with Mr. Knightley, than with Frank Churchill? Why was the evil so dreadfully increased by Harriet's having some hope of a return? It darted through her, with the speed of an arrow, that Mr. Knightley must marry no one but herself!
 — Jane Austen, *Emma*¹

Greater numbers of cognitive scientists accept that emotions can be unconscious. Less accepted — theoretically, intuitively, and empirically — is the possibility of unconscious emotional feelings. Building off David Rosenthal's and Fred Dretske's work on the types of consciousness, this paper makes the case for unfelt emotion feelings; i.e., unconscious feelings. I argue for two main claims: (1) not only emotions proper, but emotional feelings, can be unconscious; (2) we can often best learn of the emotional feelings of others and ourselves by observable behaviors, expressions, and bodily reactions. I explore three levels of data to support these claims: empirical studies on affective priming, phenomenal first person experience, and behavioral third person observational evidence. I also explore several implications of the claims for theories concerning the nature of emotions, emotional consciousness, and the functional role of emotional feelings.

Keywords: emotions, feeling theory, consciousness

In Jane Austen's novel, *Emma*, the protagonist's love for the gentlemanly Mr. Knightley manifests in a number of ways throughout her years of knowing him, all without her conscious awareness of it. In real life, we frequently glean a friend or family member's emotions (anger, embarrassment, annoyance, sorrow, etc.) even if she is unaware of it herself. Or we may ourselves come to realize, like Emma, that we are harboring a certain emotion of which we previously were not conscious.

Many thanks to David Rosenthal, Jesse Prinz, and Joshua Keton for their comments on this paper. Correspondence for this article should be addressed to Katherine Tullmann, Ph.D., Department of Philosophy, Room 106 Building 23, Babbitt Academic Annex, Northern Arizona University, 803 S. Beaver St, Flagstaff, Arizona 86011. Email: Katherine.tullmann@nau.edu

¹Austen, J. (2008). *Emma*. Waiheke Island: The Floating Press, p. 645.

The *Emma* example seems to present cases in which a subject lacks conscious awareness of his emotions. Michael Lacewing (2007) details three possible interpretations of situations that seem to include unconscious emotions. First, it's possible that Emma did *not* undergo an occurrent emotion. Perhaps she had unconscious beliefs and judgments that influenced her behavior, but she wasn't in an *emotional* state. Second, it's possible that part of Emma's emotion was conscious, but not each component of it. Perhaps Emma consciously experienced the bodily sensations associated with joy (for example) but the corresponding beliefs or judgments that comprise joy were unconscious. Finally, it may be possible that Emma experienced joy, while denying that she felt joyous. In other words, the cognitive or behavioral components of joy may have been consciously available to Emma but the feeling of the emotion was unconscious.

How can we adjudicate between these three interpretations of Emma's emotions? We cannot rely solely on introspection to adjudicate between these explanations of the phenomena, since the mental state in question is either unconscious (maybe preconscious or completely inaccessible to conscious awareness) or simply not present. Introspection may grant a sensitive subject some information about her occurrent *conscious* experiences but not necessarily her unconscious ones.

The first interpretation seems correct if we accept that emotions are constituted by conscious affective feelings (hereafter, simply "feelings"). Indeed, there is an assumed view in both the philosophic and scientific literature that feelings are necessarily conscious states or components of states. Indeed, this position is not often argued against even amongst theorists who argue that emotions are comprised by feelings (see Lacewing, 2007).² There may be a great deal of unconscious emotional processing, affective bodily responses, and emotional expressions, but feelings are always conscious. For our purposes, a feeling is the qualitative aspect of an emotion. Just as the visual perception of a delicious apple is accompanied by the qualitative feature of red, and dark chocolate has a certain flavor, so too do many emotions have a qualitative feel: the feel of anger, sorrow, or joy.

The second and third interpretations of the case studies also rely on theoretical assumptions. Specifically, these views are committed to the idea that emotions are the types of mental states that may be unconscious (Prinz, 2005; Rolls, 2005; Rosenthal, 2008). It is widely accepted that mental states possessing intentional content beliefs, thoughts, and judgments can be unconscious. It is less obvious that *qualitative* states can be unconscious. For, if mental states intrinsically possess qualitative character, and the qualitative character of a state is necessarily conscious, then the state itself must also be conscious.

We can delineate the question of unconscious feelings in two ways: (1) feelings are thought to be necessarily conscious; (2) if (1) is true, and emotions are

²See Whiting (2007) for an exception.

necessarily constituted by feelings, then emotions are also necessarily conscious. Following other prominent authors in the literature, I will refer to the view that emotions are constituted by conscious feelings the *feeling theory* (Kriegel, 2011; Whiting, 2009). I call the idea that feelings are necessarily conscious (whether or not feelings constitute emotions) the *strong feeling theory*.

There exist many compelling arguments in favor of the feeling theory. Some authors have argued for the second interpretation of the case studies presented above: the view that emotions are partially constituted by conscious feelings (Goldie, 2009; Helm, 2009). But few theorists argue in favor of the last interpretation, that emotions may be constituted by feelings but that those feelings may be unconscious. I think this is because unconscious feelings are counterintuitive. Feelings are, well, felt. What would it mean for a feeling to be unconscious? Isn't this just a contradiction of terms?

In what follows, I argue that there is a genuine question as to whether feelings are necessarily conscious. The claim that feelings are necessarily conscious follows from a vague understanding of "feeling." But must we accept the assumed understanding of "feeling"? I argue that we do not. I will argue for two central claims concerning feelings. First, there is some evidence that feelings — the "what it's like" or "qualitative" component of emotions — can be unconscious. This is a different claim than saying that *emotions* can be unconscious; I will attempt to make the case for the stronger claim that there can be *feelings* of which we are completely unaware. My second claim concerns the epistemic implications of unconscious emotion feelings: we often learn of the feelings of others and ourselves by observable behaviors, expressions, and bodily reactions. Actions and verbal reports — not conscious qualitative feelings — are often the best indication of our emotional states. My overall goal is to preserve the feeling theory of emotions while rejecting the strong feeling theory. That is, we may grant that emotions are constituted (wholly or in part) by feelings but deny that those feelings must be conscious. This view accepts the third interpretation of the Emma example while rejecting the first two.

I begin by presenting an overview of the feelings theory and how those theories see the relationship between emotions and consciousness. I then examine the meaning of "feeling" and propose a way of understanding the feelings involved in emotions in a way that makes sense of our intuitive understandings of them but also pays homage to how the term is empirically studied. Next, I present evidence in favor of my two claims, that feelings can be unconscious and that behaviors are often the best way of learning of one's own emotional state. I will spend the remainder of the paper discussing the connection between feelings, consciousness, and the function of emotions.

Feelings and the Feeling Theory

William James famously stated that emotions *just are* conscious bodily changes in reaction to stimuli in our environment, and “every one of the bodily changes is FELT, acutely or obscurely, the moment it occurs” (James, 1890/2007, p. 451). James’s contention that emotions are simply conscious feelings remains controversial, even in the wake of contemporary neuroscientific evidence for subpersonal emotion processing. Many philosophers and cognitive scientists argue that mere feelings cannot adequately capture the intentional, evaluative, and motivational aspects of emotional experiences (Currie, 1995; Goldie, 2002; Gordon, 1987; Nussbaum, 2001; Solomon, 1993; Walton, 1990). Nevertheless, many of these same theorists agree with James that feelings are an important component of emotions, even if feelings do not comprise all there is to emotions.

There are several main camps in the debate about the nature of emotions. First, cognitivists hold that a thought, belief, or judgment necessarily constitutes an emotion (Carroll, 1990; Currie, 1995; Nussbaum, 2001; Solomon, 1993; Walton, 1990). On some views, the bodily responses that make up feelings may very well accompany an emotion, but do not constitute it. In other words, there can be emotions in which there simply is no feeling, conscious or unconscious. Other cognitivists hold that both affective and cognitive elements are necessarily involved in our emotions (Aristotle, 1941; Frijda, 2000; Lazarus, 1991). Finally, non-cognitivists argue that emotions simply are bodily responses, perceptions of bodily responses, or feelings (Damasio, 1994; de Sousa, 2004; LeDoux, 1996; Prinz, 2004a, 2004b; Robinson, 2005; Zajonc, 1984, 1990). This is not to say that emotions are nonintentional or nonrepresentational; rather, the non-cognitivist point is that emotions are non-propositional and are constituted by a perceptual or bodily process.

Summarizing these positions, we can see that there are several potential aspects of emotions that may be unconscious. First, there are affective responses — the bodily changes that result during emotions that may or may not be necessary components of them. Alternatively, the cognitive judgment, belief, or thought that constitutes a genuine emotion may be unconscious. Third, the cause or object of an emotion may be unconscious. Finally, the feeling of an emotion may be unconscious. Note that one of these aspects may be necessarily conscious, while it is perfectly plausible that the others are not. For example, the feeling theorist may grant that the belief that constitutes an emotion may be unconscious while the feeling must always be conscious (see, for example, Zajonc, 2000). I take it that the most controversial position is that the feeling need not be conscious. The challenge for this view is to show that the feeling itself is unconscious and not some other aspect of the emotion — the belief, judgment, or thought involved, the causal or neural processing.

Indeed, many theorists would be happy to admit that emotions can be unconscious, thus rejecting the strong feeling theory (Haybron, 2008; Prinz, 2005; Schwitzgebel, 2008). It is less clear that these theorists are willing to grant that feelings can be unconscious. Berridge and Winkelman (2003) persuasively argue that many emotion theorists tacitly adopt the strong feeling theory, often by simply stipulating that feelings (whatever they may be) are necessarily conscious without arguing for this point. For example, Nico Frijda (2000) describes feelings in terms of hedonic experiences of pleasure or pain — which, presumably, must be conscious. Appraisal theorists such as G.L. Clore (1994) and Phoebe Ellsworth (1994) make similar claims. They hold that affective experiences must be experienced (Clore, 1994) or, at least, are generally experienced in one important aspect of our emotional lives, the subjective character of feelings (Ellsworth and Scherer, 2003).

The feeling theory makes some intuitive sense. Our conscious feelings are the most salient aspect of our emotional lives. We seek out situations, objects, and people that cause us to consciously feel positive emotions. We avoid other situations, objects, and people because they have the opposite effect; they foster conscious negative emotions. And since our feelings are at least sometimes introspectively available to us, we are free to report them (“I am so happy now” or “That news made me so angry!”) and act on their basis. So, one reason why the strong feeling theory may also seem persuasive is that it fits naturally with this functional aspect of emotions. Proponents of a strong feeling theory, including the neuroscientist Joseph LeDoux (1996, 2012), maintain that conscious emotions serve a unique function. Conscious emotions, unlike unconscious emotion processing, may add extra efficacy to our actions and deliberation. Once conscious, we can “adjust and fine tune” the state as well as modify related thoughts, intentions, desires, etc. For example, our automatic and unconsciously triggered fight or flight response during a fearful situation may, to a certain extent, be nullified or adjusted to facilitate further actions. I will focus on LeDoux’s view in the following sections, not as the only strong feeling theorist, or even as a self-described one, but as an example of the strong feeling theory at work and how the view pervades our typical understanding of emotions.

I do not wish to argue for or against one version of the feeling theory. Instead, I want to explore how these theories understand feelings. Let’s begin by examining definitions of “feeling” already present in both the philosophical and cognitive scientific literature on emotions. Here are several representative views:

1. Panksepp (2007): “Affective feelings” are “basic, internally felt neurodynamics reflecting intrinsic survival values that are experiences but not necessarily reflected upon” (p. 114).
2. Smith et al. (2018): “Emotional experiences” [EE] are “the conscious experience of many different aspects of that response, such as cognitions, bodily

- sensations, and motivated actions — as well as the conscious recognition of that overall response as belonging to a particular emotion-concept category (“fear,” sadness,” etc.). EE ... pertains to what an individual consciously experiences/recognizes in a given moment or situation” (p. 670).
3. Sizer (2006): a feeling is “a phenomenally conscious, sensation state that one occupies (as opposed to an action one performs)” (p. 110).
 4. Lacewing (2007): “Commonsense suggests that the type of awareness involved in feeling an emotion essentially involves phenomenology, i.e., the epistemological access that enables direct and noninferential avowal of an emotion necessarily involves, even if it cannot be reduced to, “feelings.” ... I leave it open, as a matter of dispute, as to whether “feelings” are or involve qualia, or are entirely reducible to intentional content” (p. 83).
 5. Goldie (2009): “the view that there are emotional feelings of a kind that can be directed immediately towards objects in the world beyond the bounds of the body: these feelings are bound up with cognition and perception, and are not the mere effects of cognition or perception. Above all, though, they must be feelings, and in deference to that, I call them *feelings towards*” (p. 232).
 6. Kriegel (2011): Feelings are “some combination of cognitive, conative, somatic, and/or affective phenomenology” (p. 428).
 7. Whiting (2009): “[E]motions are conscious is to say at a minimum that there is something that it is like for us to undergo emotion, or that emotions have a characteristic feel or phenomenology. For instance, it might be held that fear has an edgy feel, and anger has an irritable or hot-headed feel. That being said, note that the claim that emotion feels a certain way doesn’t on its own tell us anything about the character of emotional phenomenology. To be sure, that claim is consistent with holding that emotional phenomenology takes the form of bodily sensation or feeling” (p. 311).

I consider these views to be quasi-definitional, or as to at least present a working definition of “feeling.” As it turns out, this important aspect of our mental lives is exceedingly difficult to define. Importantly, though, each of these theorists takes some version of the feeling theory seriously. Each believes that feelings play some constitutional role in emotions. Nevertheless, the views differ in terms of emphasis and clarity when it comes to describing an emotional feeling. Panksepp (2007), as a cognitive neuroscientist, considers “affective feelings” in terms of neurodynamics (how various parts of the neurological system interact with each other) and the evolutionary import of feelings (how feelings aid in a subject’s survival). Panksepp also states that feelings are “internally felt.” While we might grant that feelings are functionally significant for a subject — i.e., they contribute to, influence, or directly cause behavior — this view doesn’t elucidate what it means for a feeling to be felt. We see similar vagueness in Sizer’s and Lacewing’s views: feelings

are phenomenological or qualitative, whatever that might mean. Whiting's comments gesture toward phenomenal consciousness in a similar vein to Ned Block's concept of P-consciousness. I will return to this point momentarily.

Ryan Smith, Peter Goldie, and Uriah Kriegel each propose expanded versions of feelings in which those feelings may include motivational/action-oriented, cognitive, and external components. For instance, the feeling of joy may involve cognitive phenomenology (for Kriegel and Smith) and that cognition is directed towards an object in one's environment (for Goldie). These views suggest that feelings are intentional, a position that Lacewing also considers. Feelings are about something, a view which has many supporters in the emotion literature but is not universally accepted (Whiting, 2009).

Each view characterizes feelings as conscious states in themselves or as components of a mental state. There are several points in favor of this characterization. First, as Whiting suggests, we talk about feelings (and emotions in general) as if they are necessarily conscious. When a friend of mine says "I felt hurt when you forgot my birthday," I take this to mean that my friend consciously experienced the bodily and other phenomenological sensations she associates with sorrow and hurt. The way we speak expresses what Whiting calls a folk theory of emotion: that emotions are feelings, and feelings are necessarily conscious. This coincides with both the feeling theory of emotions and the strong feeling theory of conscious feelings. Moreover, introspection often reveals to us that emotions are constituted by conscious feelings. Such introspective experiences are also taken as evidence in favor of both feeling theories (Hatzimoysis, 2007; Kriegel, 2011; Whiting, 2009).

However, as I suggested in the previous section, we must be careful here. Introspection is not always a reliable source of knowledge regarding our own mental states. Subjects may be poor introspectors. They do not regularly introspect or may not be able to parse various aspects of their occurrent states, let alone be able to *report* one's own mental states. This is the charge against introspective psychology in general: introspection is not a fully reliable source of information about mental states and processing. Importantly, introspection cannot tell us much about unconscious states, which, insofar as they are inaccessible to reflection, are not introspectable.

If we start from the assumption that feelings are necessarily conscious, then we may easily conclude that there cannot be unconscious feelings. It would take a radical reinterpretation of "feeling" to think of them as something that can be unconscious. My skepticism concerning the claim that feelings are necessarily conscious stems from a worry about the use of commonsense/folk emotion theory and introspection as evidence in favor of theories of emotion (see Schwitzgebel, 2008). I think we have good reason to mistrust first-person accounts of the mind and theories that use these accounts as a primary source of evidence. However, the point can be made that, if we can't use first-person or introspective evidence to support a theory about feelings, what else is there?

This question stems from a certain conception of mental states and consciousness that is prevalent in the philosophical literature. The emotion consciousness debate mirrors similar questions concerning qualitative states in general. Some philosophers argue that qualitative states (or, qualia), such as pains and other bodily sensations, are conscious by definition (see Nagel, 1974 for the conscious qualia argument *par excellence*). Thus, any mental state that possesses qualitative character must also be conscious. Ned Block (1995, 2009) has made a case for the consciousness of a mental state's qualitative character, distinguishing between *Access* and *Phenomenal* consciousness. A state is Access-conscious if "in virtue of one's having the state, a representation of its content is 1) inferentially promiscuous, that is, poised for use as a premise in reasoning, 2) poised for rational control of action, and 3) poised for rational control of speech" (Block, 1995, p. 231). Phenomenal consciousness, on the other hand, is *experience*. It is the state's qualitative character, its "what it's like-ness." P-consciousness includes the qualitative character of sensations, feelings, perceptions, thoughts, desires, and emotions (1995, p. 230).

Block argues that P-consciousness cannot be defined in any non-circular way. The best one can do is "point to the phenomenon" (1995, p. 230). This means that P-consciousness cannot be understood in terms of unconscious processing or information that is available to third person observation, such as facial expressions and behaviors. Defining P-consciousness in this way assumes that phenomenal experience is necessarily conscious. There are no unconscious experiences. Thus, there are no unconscious qualitative aspects of emotions (or any other state). It is possible that the intentional and informational aspects of an emotion can be unconscious, and so one cannot report or act on it. But the phenomenal character is conscious; we would feel the emotion but not remember that feeling or be able to talk about it.

Extrapolating from Block's view, it makes sense from a first-person perspective that feelings are always conscious. This is how we experience feelings. It would be natural to assume that those feelings would simply be missing if they are not experienced. Further, since the feeling is what we know, it may also be safe to assume that feelings constitute emotions. So not only are feelings necessarily conscious on this view, but so are emotions themselves.

I want to push back against Block's claim that we cannot define qualitative states such as feelings. The worry is that feelings described solely in terms of phenomenology place those feelings apart from playing a psycho-functional role in behavior, a view that several of the theorists described above seem to want to preserve (Panksepp, 2007; Sizer, 2006; Smith et al, 2018). Other theorists have attempted to define qualitative states in non-circular ways, *pace* Block (see LeDoux and Brown, 2017; Rolls, 2005; Rosenthal, 2008).

One way we might do this for feelings is to emphasize their functional roles in mental processing and in behavior: how feelings impact what we think and how

they impact what we do. Sizer (2006) makes a similar point. Like LeDoux, Sizer wishes to preserve a role for feelings as information carriers: “It is plausibly the function of a feeling to play an informational or representational role, something it does in virtue of its phenomenal content standing in certain relations to other states and states of affairs” (2006, p. 110); “Our ability to feel our emotions plays an important role, however. It allows us to make conscious connections between a certain stimulus or environmental feature, our thoughts, and an emotion” (2006, p. 119). Feelings can also inform and attune subjects to the external environment. Feelings alert a subject about external states of affairs and about one’s internal mental landscape. Feelings may also have a more direct role in motivating behavior; I will return to this possibility in a following section.

So, if we think of feelings as playing a functional role in a subject’s behavior or mental landscape, we can also begin to think of a way to define feelings beyond their phenomenology. Another point is that all mental states have some corresponding brain state. I am assuming some form of physicalism here; a substance or state dualist would likely have some qualms with this point. For instance, we may not know the exact brain state that underlies an occurrent belief, but we accept that some such brain state occurs (I’m being purposely vague here on the relationship between the mental and the brain state; it’s possible that the mental state supervenes on the brain state or that the mental state simply is the brain state). Similarly, when an emotional feeling occurs, there is some corresponding brain state underlying it. This is important because the neural basis of mental states supports a theory of functional role: brain states are caused by other brain states, and further cause more brain states. We may wish to deny that qualia play a functional role; to grant this is to accept that qualia are epiphenomenal. If emotions are feelings, feelings are qualia, and qualia are epiphenomenal, then it follows that emotions are epiphenomenal. This is, however, a conclusion I think most theorists wish to avoid.

Let’s return to the question of feelings and consciousness. Feelings are typically thought of as “what it’s like” to have certain bodily states, including emotions. In these terms, feelings are the experience of *X*, where *X* is some affective state. On the face of it, this definition isn’t terribly informative. We must now try to understand “experiences” and “affective states,” in some non-circular way. Affective states are typically thought of as those which are constituted by or cause some sort of felt bodily change. Pain, for instance, triggers a neurological response that often results in the conscious feeling of pain. Emotions are also affective states: the perceptual recognition, thought, or judgment of a mental or environmental trigger may cause the subject to feel the pangs of sorrow, or the arousal of anger.

The central question here concerns whether *experiences* are necessarily conscious. They are often treated as such. For instance, in discussing perceptual experience, Susanna Siegel claims that perceptual experiences are, by definition, conscious (Siegel 2010; also, Block above). Still, it’s worth asking whether

experiences may be *unconscious*. We must inquire about the nature of an experience to answer that question. This proves challenging; philosophers of mind and cognitive scientists often use the term “experience” without defining it — again, in much the same way that Block treats qualia, as a *sui generis* concept that, by definition, cannot be further defined. Nevertheless, I think we can make some progress here. A perceptual experience occurs when a subject undergoes an occurrent (as opposed to dispositional) percept. That is, the subject has a perception of something in her environment: a perception of a coffee mug, or cat, or pine tree. Our subject may have an experience of a belief — say, that it is cloudy outside — when that belief is manifest. Note that this way of thinking about experiences is silent on the question of consciousness: an occurrent belief may be conscious or unconscious, while still motivating behavior. The same applies to perception, emotions, pains, desires, and judgments — in principle, any mental state.

Now let’s consider defining feelings in terms of functional role and brain states: to have a feeling is to have an occurrent mental state, X, where X is constituted by or causes a bodily change that plays a psycho-behavioral functional role and is appropriately related to (supervenes on, is constituted by or some other physicalist relation) a brain state. This is the notion of “feeling” that I will work with in this paper. I think it has the benefit of capturing the phenomenological account of feeling — feelings in this sense may be consciously experienced — but is neutral on whether feelings are necessarily conscious. It also highlights the functional role and neural basis of feelings, and so goes beyond commonsense and introspective accounts of feelings (see LeDoux and Brown, 2017; Rolls, 2005).

Evidence for Unconscious Feelings

In this subsection, I will provide three layers of evidence that address the possibility of unconscious feelings and unconscious emotions. Together, they present evidence in support of the claim that emotions may occur unconsciously, even if emotions are constituted by feelings. This preserves the feeling theory, but radically amends it. The evidence presented here is intended to put chinks in the armor of the assumption that feelings are necessarily conscious. If we grant this evidence even some initial plausibility, then we may be forced to question our assumptions about emotional consciousness and the consciousness of qualitative character in general.

Evidence from Subliminal Processing

I will begin with evidence from work in cognitive psychology and neuroscience on the power of unconscious affective processing to influence our conscious

attitudes and behaviors.³ In the mid-1900s, members of the New Look movement argued that perceptions are “constructions that integrate sensory information about physical stimuli with internal factors, such as needs, goals, attitudes, and emotions” (LeDoux, 1996, p. 55). Their experiments showed that participants could have ANS (autonomic nervous system) responses (visceral reactions like increased heart rate, perspiration, pupil dilation, etc.) to emotionally-charged stimuli even while they were not consciously aware of the stimuli. In particular, “taboo” and “dirty” words are recognized more quickly than those that do not have emotional connotations (1996, p. 56). Richard Lazarus’s work on subliminal perception showed similar results (Lazarus and McCleary, 1951). Patterns of letters were briefly flashed on a screen, too quickly for a verbal identification of the letters. The patterns that were previously accompanied by electric shock generated increased ANS response in subsequent trials, even though the participants reported no conscious awareness of the stimulus.

Zajonc’s work in subliminal emotional priming has likewise been extremely influential in showing that emotional processing can take place unconsciously (Murphy and Zajonc, 1993). An emotionally-laden priming stimulus, such as a frowning or smiling face, is very briefly (5 ms) shown to the participant, followed by a masking stimulus. The mask eliminates the possibility that the participant can recall the prime. Another emotionally-neutral stimulus, such as a Chinese ideograph, is then shown for a long enough period for it to be consciously perceived. The participant is then asked to rate how much he likes the current stimulus. The participant’s preference rating directly corresponded to the valence of the masked priming stimulus; participants liked the stimulus if they were primed with the smiling face, disliked it if they were primed with the frowning face.

This work indicates that subliminal emotional processing can result in conscious mental states. And although unconscious priming has recently come under some scrutiny (see Newell and Shanks, 2014), there seems to be a consensus that unconscious processes and states can affect conscious decision-making and other behaviors. The above data further indicate the possibility of unconscious feelings, as evidenced by the prevalent ANS responses. The participants must have been aware of the unconscious, emotionally-laden stimuli in some way, or else they would not have responded as they did (unconsciously aware, not consciously so). The idea is that feelings at least co-occur with bodily responses; if the bodily response is present, then we have some indication (although, I grant, not a guarantee) that the feeling is present as well. If a conscious feeling is a bodily response that plays a functional role, then changes in bodily states may still qualify as feelings, but unconscious ones.

³See LeDoux (1996) for a full review of this material. See also John Bargh’s (1992) work on implicit attitudes.

One might not find this conclusion very convincing. Surely there is more to emotions than the kind of low-level stimulus processing tested by the above studies. In response to claims like this, Berridge and Winkielman (2003) developed a study that suggests the presence of unconscious feelings, as opposed to mere emotional processing. Participants were asked to rate their subjective emotional reactions immediately after their exposure to subliminal emotionally-laden stimuli — again, smiling or frowning faces. This was masked by a consciously perceived neutral face. They gave the participants a pitcher of fruit-flavored water, which they could pour for themselves after the subliminal exposure. The authors found that exposure to the subliminal emotional expressions influenced the subjects' subsequent pouring and drinking of the flavored water. Those exposed to happy faces poured and drank about 50% more than those exposed to neutral faces. Those exposed to angry faces poured significantly less than those exposed to neutral faces.⁴

This study seems to indicate that we can have unconscious emotional reactions to unconsciously perceived stimuli. There was a behaviorally demonstrable effect from the unconscious perception and the subjects were unable to introspect the cause of their pouring and drinking behaviors or indicate at all that they had been influenced by the unconscious stimulus.

One could argue, though, that these results are not evidence of unconscious feelings. Rather, the results support some other kind of unconscious emotional processing, such as information mechanisms that would have later caused an actual conscious affective reaction, if they had not been blocked by a mask. Unconscious *information* processing led to the pouring/drinking results, not the presence of an affective feeling.

Berridge and Winkielman argue that this interpretation of the data is inconsistent with the abundant research on subliminal facial expressions that elicit genuine affective changes, including physiological and behavioral effects (activation of the amygdala and other parts of the brain associated with affect), spontaneous mimicry, and skin conductance responses — effects that are generally indicative of affective *feelings*. Still, it could be that the subliminal frown or smile elicited a “free-floating cognitive belief that something good or bad is happening,” which then influenced the interpretation of the subsequent stimuli, such as the drink (2003, p. 191). This would lead the participants (and, in particular, the thirsty ones) to pour and drink that flavored water. The authors find this interpretation of their data unconvincing. Some of the participants were asked to introspect their mental states immediately after the neutral mask. If the information view is correct, then these participants should report that they felt happy or angry, good or bad, because they assigned the free-floating belief to *themselves*. But this

⁴See Berridge and Winkielman (2003) for more details. The authors also controlled for factors such as attention, focus, thirstiness, etc., that may have shaped the results.

is not what they found; the participants did not self-attribute an emotional state. Nevertheless, exposure to the affective state influenced their behavior. Because of the connection between feelings and behavior, the authors suggest that the participants had unconscious affective feelings.

Whiting (2007) presents another potential issue with the authors' interpretation of this study:

Pace Winkielman et al. why not think that the study results show only that there can be changes in emotion that people can fail to consciously reflect on or form certain beliefs about? Indeed, there are a number of reasons why study participants might have been unable to reflect on changes in felt emotion. For instance, the changes might have been too subtle and/or short-lived to be registered or reflected on (which is consistent with supposing that such changes might nevertheless have had significant effects on behavior). And although the authors of the study acknowledge this kind of response to their argument (Winkielman et al. 2005: 132), they say little to remove or mitigate the worry. (pp. 317–318)

Summarizing this point, it is possible that in supposed cases of unconscious emotions that a subject's feelings really are conscious but occur too quickly for the individual to report or reflect on. Thus, the subjects report that the feeling didn't consciously occur. In a more recent paper, Winkielman, Berridge, and Sher (2011) acknowledge this limitation in their research, and merely suggest that further research is required to adequately distinguish between a subject's failure to reflect on and report feelings and the actual inexperience of a feeling: new methods in non-verbal reporting or more introspection training may be required to determine whether Whiting's interpretation is correct, or whether the emotion is unconscious. This challenge highlights the difficulty of measuring, reporting, and understanding unconscious mental states in general. Unconscious states are unreportable. I think the best evidence in favor of Whiting's interpretation (that feelings are necessarily conscious) is introspective. This highlights why the value of a functional approach to deduce the presence of an unreported state is so important: the presence of a behavior typically associated with a feeling may be evidence that the corresponding state occurs. The alternative position would be to deny that the subjects in this study undergo an occurrent emotion state. This results in the worry about emotional epiphenomenalism.

Behavioral Evidence

We often attribute mental states to others based on their behaviors, verbal reports, facial expressions, and bodily reactions (perspiration, tension, etc.). For example, I guess that my friend is thirsty if she goes to the kitchen to get a glass from the cupboard. I would also guess that she believes that the third cupboard from the left contains glasses if that's the one toward which she reaches. I predict that my friend will then go to the sink and turn on the tap. Sometimes attributing

emotional states is as simple as this. You watch as your sister gets increasingly irritated and angry with her colleague's insensitive comments about another coworker. She crosses her arms, frowns, and scoffs at her colleague's remarks. You notice as your brother becomes frightened and anxious as he watches a Hitchcock thriller, as he grabs the sides of his chair or clutches his face in nervous anticipation, eyes wide. Your best friend's anguish over her favorite sports team's recent loss is palpable and extreme. Her whole body seems to sag and she hangs her head in dejection. Indeed, we have no trouble attributing both emotion *types* and *content* based on one's facial expressions, bodily responses, and behaviors. The subject herself will be also able to tell us about her emotional state in many cases since the feeling is conscious to her. But what about situations in which one is not consciously aware of one's feelings? Could one's expressions and behaviors reveal *unconscious* emotions and feelings?

Consider a second example from Jane Austen's novel, *Emma*. Emma witnesses the cruel Mr. Elton snub her friend Harriet by refusing to dance with her. Emma then watches Mr. Knightley "rescue" Harriet by dancing with Harriet himself. Emma is overjoyed, and surely consciously so. Let's imagine that this joy manifests itself later on, so that when Mr. Knightley and Emma are dancing and she recalls Mr. Knightley's kind act, she smiles to herself, dances just a little lighter, and laughs more readily than usual.

Emma seems to be acting out of joy. If Emma reflected on her laughter, smile, and other bodily changes, she may wonder what it was that caused these changes in her behavior. And, being the intelligent young woman that she is, Emma may conclude on the basis of these behaviors that she feels abundantly happy with Mr. Knightley. Yet Emma is so focused on the dance and her partner's conversation that this joy may not be apparent to her at this later time.

I think examples like this are commonplace. Other people may have better knowledge of our emotional states than we do, since facial and verbal expressions and body language are generally reliable indications of what one is feeling.⁵ There are several points to make here. First, Emma's joy strikes me as a case of unconscious emotion as opposed to emotional processing. She was happy without being aware of it, not processing some information in a happy-type way. Of course, proponents of the strong version of the feeling theory will argue that this couldn't be an actual emotion; Emma isn't joyful until she becomes conscious of it. This is one possible interpretation of the situation. But we can also propose that Emma indeed did have an emotional response that included reliable behavioral responses.

Second, following from the discussions concerning affective priming, it seems like this could be a case of unconscious affective *feeling*. If one's tone of voice, facial expressions, and ANS bodily responses that co-occur with feelings are present,

⁵ One can, of course, mask one's true feelings or pretend to have feelings that one lacks. See Goldie (2002) for an interesting discussion of emotional pretense.

then we have some evidence that the feeling is present, but not consciously felt. One could alternatively argue that these are unconscious processes that indicate and lead to feelings, but are not themselves genuine feelings (which are, by definition, consciously experienced). But, again, there doesn't seem to be any non-question begging reason to support this assumption.

Finally, as a general point, it seems likely that we can learn of other's unconscious behaviors from their behaviors and expressions, as well as from our own. Emma realizes that she is happy after noticing her unusual behaviors and reflecting on them. This realization will likely put Emma's other thoughts and behaviors into a new perspective; now she knows that she danced happily and that she smiled *out of joy*. This sort of inference-drawing may not (or need not) typically occur, but it can. Thus, observing the outward behaviors and expressions of ourselves and others can be an important source of information about one's emotional state, conscious or unconscious.

First-Person Experience

I would like to present one last example as evidence for unconscious emotions, this time from the perspective of first person experience. Jesse Prinz (2004a) offers the following example:

Imagine that you are given to a particular phobia, say fear of flying. As your plane takes off on an intercontinental flight, you find yourself entering a state of acute anxiety. Your heart is racing, you clutch the armrests, your body stiffens against the seat as if it were being pulled back by gravity. Sensing your distress, the friend you are traveling with begins to tell you a humorous anecdote about her visit to a baldheaded barber. For a moment you lose yourself in her inane yarn. In that brief interval, you are not aware of your fear. In fact, you experience amusement. But then you recall where you are, and terror immediately returns. You notice that the whole while you have been clutching the armrests with equal vigor. Even during the moment of greatest absorption in the barber anecdote, you never stopped displaying the bodily signs of fear. (p. 202)

Prinz suggests that there are two possible interpretations of this kind of case. Your fear is conscious while you are thinking about flying and focusing on the take-off. But when you are distracted by your friend's tale, your fear disappears. You are no longer afraid, but amused. Then the fear — including the qualitative character — starts right back up again when the story ends. Alternatively, it could be that you were afraid throughout your friend's story, but the qualitative aspects of fear were unconscious. The fear became conscious again once the story ends and you are reminded of your surroundings.

The second interpretation is the one most conducive to my claims, and perhaps also the most theoretically plausible. This is because only in this story does

the behavioral data make sense. You never fully relaxed during your friend's tale; your body was tense, your hands clutching the armrest. Those bodily responses to your surroundings were still present. This suggests that your fear did not dissipate, but rather that you were not consciously aware of it. It also suggests the further claim that your feelings are unconscious, at least assuming the functionalist interpretation. I think that our own first-person experiences corroborate this interpretation. Emotions and feelings may not be conscious to us when we watch a scary film, find ourselves in a tense situation, or enjoy a live concert. However, we may still take actions that evince bodily behaviors that are emotionally relevant in these situations.

Together, the three types of examples that I have discussed in this section — emotion priming, third person access, and first person experience — provide some folk psychological and empirical support for the possibility of unconscious emotions and feelings. This motivates both my first and second claims for this paper: emotional feelings may be unconscious and observable behaviors and bodily reactions are the best evidence we have of unconscious emotions and feelings. Introspection and first-person awareness also, to an extent, support the unconscious emotion view.

As a final point, we can consider the implications of the metaphysics of mind on this question. If we take some variety of physicalism for granted, then it makes sense to say that feelings co-occur with physical responses and alterations in the brain and body. The cause of a qualitative state — whether emotional or some other mental state type — will be detectable. The presence of the feeling's cause is some indication that the feeling is present as well. Furthermore, again assuming physicalism, certain behavioral responses and brain states are indicative of a mental state. So, if the typical causes, brain states, and behavioral responses to emotional stimuli are present, then we may reasonably infer (if not guarantee) that the feeling is present as well, whether a subject would claim that the feeling is consciously experienced or not.

Implications

I have highlighted several different ways in which we can interpret the evidence I have presented here. First, we can say that there was no occurrent emotion present. For instance, the emotion turned on and off during the flight, Emma didn't really experience joy, and the subliminal emotion processing did not give rise to a full-fledged emotion. This view preserves the feeling theory. None of these subjects had conscious feelings and so they did not have emotions. Another interpretation is that these emotions were unconscious. There are two ways to approach this interpretation. First, we could say that, while the emotion itself was unconscious, the subject nevertheless had conscious feelings. Your flight-averse friend did not have the relevant beliefs or thoughts about the danger of flying

(say) at the right time, so did not feel fear. The feelings associated with fear may have been present, but the cognitive states that comprise fear were unconscious.

I propose a third option: the emotions were unconscious in each case and this includes the feeling component of the emotion (whether the feeling exhausts the emotion or is merely one part of the emotion). This preserves a modified version of the feeling theory: emotions may be constituted by feelings but those feelings need not be conscious. The motivation for this view stems from the role that feelings are thought to play in behavior. In each case, we saw how emotions predict certain behaviors (clinging to an armrest, facial expressions, etc.) even though the subject was not consciously aware of the cause of those behaviors.

The main challenge to this view is to show that feelings are not epiphenomenal. Following Robert Gordon (1987), Sizer states that it may not be “the emotion that gets us running from the bear, crying out in fear or laughing at the joke. The emotion is merely the feeling or experience of those changes” (2006, p. 117). In other words, it is not the feeling itself that causes behavioral responses. Rather, it is some other mental state doing the functional work: perhaps the perception of the bear or the belief that such creatures may cause harm. Following Sizer, I think that this view denies emotions the causal and explanatory power with which we usually associate to them. We saw above that Sizer attributes the power of informational attuning to emotions: emotions are causally efficacious in virtue of their ability to make a subject aware of her external or internal happenings.

The problem is that this functional role seems to require feelings to be conscious. I must be consciously aware of my feeling, its cause, and its object in order to pick up on its informational capacity. We can't do this with unconscious feelings. This may be one more strike against the possibility of unconscious feelings: feelings must necessarily be conscious in order to display the functional role we have described them as having.

Feeling's Functional Role

Many theorists advocate unconscious emotions. However, few support the claim that feelings can be unconscious. Still fewer theorists are willing to grant that conscious feelings lack special functionality. I wish to explore this position and argue that unconscious emotions can have functionality in virtue of their unconscious qualitative character. This separates my view from those of, say, Prinz and Zajonc.

What is the function of emotions, in general? It's often argued that emotions promote an organism's well-being; emotions generally help organisms to appropriately respond to situations that bear on their well-being, that are harmful or beneficial in some way. The question for us to consider is whether the emotional state — including its qualitative character — must be conscious in order to fulfill that function. Many theorists reject the view that qualitative properties of a mental state play a functional role. This is the point of Block's distinction between

Access and Phenomenal Consciousness, for instance. On this view, it is not the qualitative character of emotions that influences behavior, but rather some other component. That is not to say (as Sizer argues) that conscious awareness of our feelings does not ever lead to behavior after introspection. But the qualitative character of the emotion *itself* is not playing the functional role. I have taken the alternate position in this paper: qualitative character of emotions may have some influence on bodily responses, and so may have functionality in this respect. If so, our question concerns whether these responses and the feeling itself need be conscious in order to influence behavior.

Let's begin with a picture of how emotion processing leads to action. Here, I draw on LeDoux's extensive work on emotion processing (LeDoux, 1996, 2012; LeDoux and Phelps, 2008). According to LeDoux, the emotion process begins when a stimulus in one's environment (or an internal stimulus such as a thought or memory) causes a mental representation of something that bears on one's well-being. We form a mental representation of the stimulus as corresponding to an emotional property based on the memory of similar objects, or from learned or innate concepts and triggers. For example, humans may have an innate predisposition to fear heights or looming objects and a learned fear of monsters under the bed, certain people, and social interactions (Damasio, 1994). When we perceive, imagine, or think about a looming object, barking dog, or slithering snake, we form a mental representation that triggers an emotional response, drawing on stored concepts and memories of these objects.

The representation of an emotion's object takes place subconsciously and subpersonally. We respond to perceived objects as pleasant or unpleasant, things to avoid or approach. Information about the stimulus is transferred to areas of the brain that are involved in encoding emotional reactions, such as the sensory thalamus, the affective division of the striatum, the orbitofrontal cortex and, finally, the amygdala, each of which is involved in initiating bodily reactions and behaviors such as an increased heart rate or freezing response (LeDoux, 1996; Mello and Villares, 1997; Rolls, 2005; Schroeder and Matheson, 2006).

Running parallel to the subcortical processing, the sensory signals also take a slower track to cortical regions of the brain for further processing and input from cognitive faculties. LeDoux (1996, 2012; LeDoux and Phelps, 2008) calls the initial pathway the "low-road" of emotions, and the slower pathway the "high-road." The sensory information involved in the initial appraisal reaches the amygdala or hippocampus after first being processed by the sensory thalamus. The same information is processed by the sensory thalamus, and is also sent to sensory cortex for cognitive processing and availability for consciousness. There, the subject's knowledge about the stimulus, her beliefs about her current environment, and desires concerning her future goals may all influence how the stimulus information is processed. She may also deliberate, draw inferences, and make decisions concerning the potential value of the object.

The question remains as to whether the conscious feeling of the emotion has any added utility to the above processes. LeDoux (1996) suggests that it does. He argues that emotion processing, especially via the subcortical pathway, leads to emotional reactions, while the cortical pathway, which generally results in conscious feelings, leads to actions. One of LeDoux's examples highlights how this works:

As [a] hiker walks through the woods, he abruptly encounters a snake coiled up behind a log on the path. The visual stimulus is first processed in the brain by the thalamus. Part of the thalamus passes crude, almost archetypal, information directly to the amygdala. This quick and dirty transmission allows the brain to start to respond to the possible danger signified by a thin, curved object, which could be a snake, or could be a stick or some other benign object. Meanwhile, the thalamus also sends visual information to the visual cortex.... The visual cortex then goes about the business of creating detailed and accurate representation of the stimulus. The outcome of the cortical processing is then fed to the amygdala as well. Although the cortical pathway provides the amygdala with a more accurate representation than the direct pathway to the amygdala from the thalamus, it takes longer for the information to reach the amygdala by way of the cortex. In situations of danger, it is very useful to be able to respond quickly. It is better to have treated a stick as a snake than not to have responded to a possible snake. (1996, p. 166)

Imagine that the hiker actually saw a coil of vine. Once the details of the visual stimulus are processed by visual cortex, and this information is sent back to the amygdala, the emotional response that the hiker underwent (freezing), will subside. On the other hand, if the stimulus actually *is* a snake, the hiker will go from his initial freezing response to another more appropriate action, such as cautiously stepping around the snake or going back the way he came. In the first case, the conscious appraisal of the stimulus leads to a reduction of the automatic reaction. In the second case, the conscious appraisal of the stimulus leads to different precautionary actions. Once we have a conscious feeling as the result of a potentially dangerous stimulus, we can modify our behavior to match what is required by the situation.

Emotions presumably come to be conscious in the same way as any other mental state. Different theories will have different explanations of how this works. LeDoux proposes a variation on the global workspace theory of consciousness similar to those developed by Bernard Baars (2007) and Stanislas Dehaene and Lionel Naccache (2001; see also LeDoux, 1996; LeDoux and Phelps, 2008). The brain has a large capacity for unconscious events but only a limited capacity for consciousness. Thus, different neural mechanisms must compete for consciousness. Once certain mental processes are conscious, they can be used for "accessing, disseminating, and exchanging information, and for exercising global coordination and control" (Baars, 2007, p. 240). The global access of a state leads to increased rational capacities (Block, 2009).

LeDoux's version of the global workspace theory emphasizes the importance of working memory (LeDoux, 2012; LeDoux and Phelps, 2008). LeDoux and

Phelps (2008) state: “our hypothesis is that the mechanism of consciousness is the same for emotional and non-emotional states and that what distinguishes these states is the brain system that consciousness is aware of at the time” (p. 171). Along with long term memory, working memory has the ability to integrate information from across sensory modalities into unified representations. These representations are useful for controlling mental activities and behaviors once they become conscious (p. 171). With conscious emotions, input from various neural systems makes working memory respond to the fact that the fear system (for example) has been activated.

This version of the global workspace theory seems to make sense of the special utility of conscious emotions. Engaging in appropriate action to emotional stimuli can be very useful, especially since our initial automatic reaction can get things wrong — we may confuse a coil of rope for a snake, or misinterpret another’s grimace for a smile. If the slower access to working memory of emotional input can supersede these mistakes, then conscious emotional feelings do seem to have an important function. But why does availability for executive control require that a state be conscious? Can’t higher level cortical processing be unconscious and lead to increased functionality? LeDoux assumes that full-blown emotions are conscious because feelings are conscious. So the feelings are playing the active functional role. He does not consider that the unconscious, automatic processing, or even unconscious *cognitive* processing, is really what’s doing the functional work.

Smith et al. (2018) provide another version of the global workspace theory in terms of affective experience. They describe the psychological and neurological processes by which affective responses generate upon perception of an emotionally salient object. Some responses are automatic, such as one’s facial expression, body posture, muscle tension, heart rate, respiration, etc. Other responses to the object may be cognitive: changes in one’s attention, motivation, decision making, etc. Many of these changes are automatic (quick, effortless) and not under voluntary control (p. 671). Like LeDoux, Smith et al. cite evolutionary pressures for granting organisms with these automatic capacities. New stimuli acquire the same affective responses through repeat patterns of sensory input and conditioning. So, affective responses serve as appraisals of objects in one’s environment, as being “novel or familiar, relevant or irrelevant to one’s current concerns, congruent or incongruent with one’s current goals, in or out of one’s control, and consistent or inconsistent with one’s norms and values” (2018, p. 671). Importantly for our purposes, these appraisals need not be consciously experienced in order to have behavioral effects. Indeed, the authors suggest that affective responses are generally unconscious before they arise to consciousness, as a global workspace theory would suggest (see also Smith and Lane, 2016).

The basic assumption here is that once the emotional feeling is conscious, we can modify it, moving from automatic reaction to regulated action, as an executive

model of consciousness implies. I would like to consider another interpretation of this data. Recall that we are considering the possibility that emotions need not be consciously felt in order to play a functional role. We saw some evidence for this above, with the studies concerning unconscious priming. We can now offer a theoretical model to support this claim.

Rosenthal (2008) argues that mental states are often unconscious before a subject becomes consciously aware of them. This makes sense if we accept the view that mental states are not intrinsically conscious. It is possible, then, that the properties of the unconscious state can influence behavior and lead to action even before the subject is aware of the state itself.⁶ One might think that this may be adequate for simple actions, but it cannot account for more complex or rational behaviors (comparable to LeDoux's action/reaction distinction). But once we accept that mental states can be either conscious or unconscious, one must show that the unconscious states can't play a strong role in bringing about certain actions. Assuming that emotional feelings must be conscious suggests that full-blown emotions are the end result of the preceding emotional processing. But, as we have seen, it is possible that emotions need not be conscious and so shouldn't be equated with conscious feelings. If unconscious emotional feelings are a possibility, then it's relevant to ask whether the function of emotions may result from them. A complex action, as opposed to a simple one, would only require more mental states to bring it about and each of those states may itself occur unconsciously.

If Rosenthal's interpretation is correct, then any mental state that first occurs without being conscious can later become conscious. The prior unconscious occurrence undermines the idea that only consciousness leads to regulated or modified behaviors. If the conscious state was first unconscious, or preceded and causally determined by a different unconscious state, then the real utility may come from one of them. This applies to emotions as well: unconscious feelings and emotions may have a great deal of utility.

It is not enough to simply claim that consciousness adds functional efficacy to a mental state. One has to make the further claim that the behaviors that supposedly result from the conscious control are not also *preceded* and *causally determined* by unconscious states. Even LeDoux claims that the truly important response occurs unconsciously or pre-consciously: as quoted above, it is much better to respond to the coil of rope as if it was a dangerous snake, just in case. Although conscious information often makes our emotional responses more appropriate, in the long-term, evolutionary story, it is beneficial to have these initial unconscious

⁶Rosenthal responds to work on the timing of decision-making, particularly in studies by Benjamin Libet and his colleagues. See Libet (1985), as well as Haggard (2005) and Haggard and Libet (2001), and Graves et al. (2010) for updated versions of the volition studies. See Gomes (2002) for alternate interpretations of these studies.

responses which are present in humans and other non-human animals that do not seem to have the same cortical capabilities as we do.

Once we grant that unconscious feelings are possible, we may also consider the further possibility that unconscious feelings can play roughly the same role in behavior as conscious ones. This is because emotional feelings can play functional roles in influencing our behavior; previously considered evidence suggests that unconscious feelings can play these roles as well. If so, then feelings do not seem to gain any special utility in virtue of being conscious. It is not that conscious feelings play no functional role. Rather, we have no reason to think that they must be conscious to influence behavior and, in fact, in many (maybe most) cases, unconscious feelings may be doing the real functional work.

The point of this discussion is not to show that mental states are epiphenomenal in virtue of being conscious — far from it. Rather, I wish to thwart a potential problem for the view that feelings can be unconscious: if qualitative character does, in fact, play a functional role then it would likely be in virtue of being conscious. By questioning this assumption as I have done, we should reconsider the possibility that feelings are necessarily conscious.

What Jane Knew

If I am right, then we have some reason to think that feelings can be unconscious. These unconscious feelings may influence and motivate our behaviors. The masterful folk psychologist, Jane Austen, recognized this back in the early 1800s. But what Jane knew has not received adequate attention in the contemporary work on emotions and consciousness. Proponents of a strong feeling theory of emotions claim that all emotions are consciously felt. In other words, emotions are necessarily conscious states, even though much of the processing involved in those states initially occurs subconsciously. This view suggests that conscious emotions have a unique function: conscious feelings are available to executive control and aid in our reactions to emotionally charged situations. I've attempted to show that the evidence for this claim is inadequate since it does not sufficiently rule out the possibility that the same effects are not achieved by unconscious states. Just because the structure of emotions and emotional processes can be interpreted as leaning toward a certain theory of consciousness does not mean that those are the *only* interpretations.

I've taken emotional consciousness to be an especially interesting case of state consciousness partly because of the strong subjective qualities that emotions have and the important role they play in our lives. Yet we can still explain the subjective character of mental states independently of arguing that state consciousness has a special function. I hope that this paper provides evidence for that claim.

References

- Aristotle (1941). *Rhetoric*. In R. McKeon (Ed.), *The basic works of Aristotle*. New York: Random House.
- Austen, J. (2008). *Emma*. Waiheke Island: The Floating Press.
- Baars, B. J. (2007). The global workspace theory of consciousness. In M. Velmans and S. Schneider (Eds.), *The Blackwell companion to consciousness* (pp. 236–246). Malden, Massachusetts: Blackwell Publishing.
- Bargh, J.A. (1992). Being unaware of the stimulus vs. unaware of its interpretation: Why subliminality per se does matter to social psychology. In R. Bornstein and T. Pittman (Eds.), *Perception without awareness* (pp. 236–255). New York: Guilford.
- Berridge, K. C., and Winkielman, P. (2003). What is an unconscious emotion? (The case for unconscious liking). *Cognition and Emotion*, 17, 181–211.
- Block, N. (1995). On a confusion about a function of consciousness. *Behavioral and Brain Sciences*, 18, 227–287.
- Block, N (2009). Comparing the major theories of consciousness. In M. Gazzaniga (Ed.), *The cognitive neurosciences IV* (pp. 1111–1122). Cambridge, Massachusetts: MIT Press.
- Carroll, N. (1990). *The philosophy of horror, or paradoxes of the heart*. New York: Routledge.
- Clore, G. (1994). Why emotions are never unconscious. In P. Ekman and R. Davidson (Eds.), *The nature of emotion: Fundamental questions* (pp. 285–290). New York: Oxford University Press.
- Currie, G. (1995). The moral psychology of fiction. *Australasian Journal of Philosophy*, 73, 250–259.
- Damasio, A. (1994). *Descartes' error*. New York: Harper Collins Publishers.
- Dehaene, S., and Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: Basic evidence and a workspace framework. *Cognition*, 79, 1–37.
- de Sousa, R. (2004). Emotions: What I know, what I'd like to think I know, and what I'd like to think. In R. C. Solomon (Ed.), *Thinking about feeling* (pp. 61–75). New York: Oxford University Press.
- Ellsworth, P. (1994). William James and emotion: Is a century of fame worth a century of misunderstanding? *Psychological Review*, 101, 222–229.
- Ellsworth, P., and Scherer, K.R. (2003). Affective processes in emotion. In R.J. Davidson, H.H. Goldsmith, and K.R. Scherer (Eds.), *Handbook of affective sciences* (pp. 572–595). New York: Oxford University Press.
- Frijda, N.H. (2000). The psychologists' point of view. In M. Lewis and J.M. Haviland Jones (Eds.), *Handbook of emotions* (pp. 59–74). New York: The Guilford Press.
- Goldie, P. (2002). *The emotions*. New York: Oxford University Press.
- Goldie, P. (2009). Getting feelings into emotional experience in the right way. *Emotion Review*, 1, 232–239.
- Gomes, G. (2002). Problems in the timing of conscious experience. *Consciousness and Cognition*, 11, 191–197.
- Gordon, R. (1987). *The structure of emotions*. Cambridge: Cambridge University Press.
- Graves, T. L., Maniscalco, B., and Lau, H. (2010). Volition and the function of consciousness. In W. Sinnott-Armstrong and L. Nadel (Eds.), *Conscious will and responsibility: A tribute to Benjamin Libet* (pp. 109–123). New York: Oxford University Press.
- Haggard, P. (2005). Conscious intentions and motor cognition. *TRENDS in Cognitive Science*, 9, 290–295.
- Haggard, P., and Libet, B. (2001). Conscious intention and brain activity. *Journal of Consciousness Studies*, 8, 47–63.
- Hatzimoysis, A. (2007). The case against unconscious emotions. *Analysis*, 67, 292–299.
- Haybron, D.M. (2008). Do we know how happy we are? On some limits of affective introspection and recall. *Nous*, 41, 394–328.
- Helm, B.W. (2009) Emotions as evaluative feelings. *Emotion Review*, 1, 248–255.
- James, W. (2007). *Principles of psychology*. New York: Cosimo. (originally published 1890)
- Kriegel, U. (2011). Towards a new feeling theory of emotion. *European Journal of Philosophy*, 22, 420–442.
- Lacewing, M. (2007). Do unconscious emotions involve unconscious feelings? *Philosophical Psychology*, 20, 81–104.
- Lazarus, R. (1991). *Emotion and adaptation*. New York: Oxford University Press.

- Lazarus, R., and McCleary, R. (1951). Autonomic discrimination without awareness: A study of subception. *Psychological Review*, 58, 113–22.
- LeDoux, J. E. (1996). *The emotional brain*. New York: Simon and Schuster.
- LeDoux, J.E. (2012). Rethinking the emotional brain. *Neuron*, 73, 653–676.
- LeDoux, J.E., and Brown, R. (2017). A higher-order theory of emotional consciousness. *PNAS*, 114, 2016–2025.
- LeDoux, J.E., and Phelps, E.A. (2008). Emotional networks in the brain. In M. Lewis, J. M. Haviland-Jones, and L. Feldman Barrett (Eds.), *Handbook of emotions* (third edition, pp. 159–179). New York: The Guilford Press.
- Libet, B. (1985). Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behavioral and Brain Sciences*, 8, 529–566.
- Mello, L., and Villares, J. (1997). Neuroanatomy of the basal ganglia. *Psychiatric Clinics of North America*, 20, 691–704.
- Murphy S., and Zajonc, R.B. (1993). Affect, cognition, and awareness: Affective priming with sub-optimal and optimal stimuli. *Journal of Personality and Social Psychology*, 64, 723–39.
- Nagel, T. (1974). What is it like to be a bat? *The Philosophical Review*, 83, 435–450.
- Newell, B.R., and Shanks, D.R. (2014). Unconscious influences on decision-making: A critical review. *Behavioral and Brain Sciences*, 37, 1–19.
- Nussbaum, M. (2001). *Upheavals of thought: The intelligence of emotions*. Cambridge: Cambridge University Press.
- Panksepp, J. (2007). Affective consciousness. In M. Velmans and S. Schneider (Eds.), *The Blackwell companion to consciousness* (pp. 114–129). Malden, Massachusetts: Blackwell Publishing.
- Prinz, J.J. (2004a). *Gut reactions*. Oxford: Oxford University Press.
- Prinz, J.J. (2004b). Embodied emotions. In R. C. Solomon (Ed.), *Thinking about feeling* (pp. 44–59). New York: Oxford University Press.
- Prinz, J.J. (2005). Are emotions feelings? *Journal of Consciousness Studies*, 12, 9–25.
- Robinson, J. (2005). *Deeper than feeling*. New York: Oxford University Press.
- Rolls, E. (2005). *Emotion explained*. Oxford: Oxford University Press.
- Rosenthal, D. M. (2002). The timing of conscious states. *Consciousness and Cognition*, 11, 215–220.
- Rosenthal, D.M. (2008). Consciousness and its function. *Neuropsychologia*, 46, 829–840.
- Schroeder, T., and Matheson, C. (2006). Imagination and emotion. In S. Nichols (Ed.), *The architecture of the imagination: New essays on pretence, possibility, and fiction* (pp. 19–40). Oxford: Clarendon Press.
- Schwitzgebel, E. (2008). The unreliability of naïve introspection. *Philosophical Review*, 117, 245–273.
- Siegel, S. (2010). *The contents of visual experience*. New York: Oxford University Press.
- Sizer, L. (2006). What feelings can't do. *Mind and Language*, 20, 108–135.
- Smith, R., and Lane, R.D. (2016). Unconscious emotion: A cognitive neuroscientific perspective. *Neuroscience and Biobehavioral Reviews*, 69, 216–238.
- Smith, R., Killgore, W.D.S., and Lane, R.D. (2018). The structure of emotional experience and its relation to trait emotional awareness: A theoretical review. *Emotion*, 18, 670–692.
- Solomon, R. (1993). *The passions: Emotions and the meaning of life*. Indianapolis: Hackett.
- Walton, K. (1990). *Mimesis as make-believe*. Cambridge, Massachusetts: Harvard University Press.
- Winkielman, P., Berridge, K., and Sher, S. (2011). Emotion, consciousness, and social behavior. In J. Decety and J.T. Cacioppo (Eds.), *Handbook of social neuroscience* (pp. 195–211). New York: Oxford University Press.
- Whiting, D. (2007). Consciousness and emotion. In R.J. Gennaro (Ed.), *The Routledge handbook of consciousness* (pp. 310–321). New York: Routledge.
- Whiting, D. (2009). The feeling theory of emotion and the object-directed emotions. *European Journal of Philosophy*, 19, 281–303.
- Zajonc, R.B. (1984). On the primacy of affect. *American Psychologist*, 39, 117–123.
- Zajonc, R.B. (2000). Feeling and thinking: Closing the debate over the independence of affect. In J.P. Forgas (Ed.), *Feeling and thinking: The role of affect in social cognition* (pp. 31–58). New York: Cambridge University Press.