Wonder as a metacognitive emotion*

Abstract

Although wonder has been the subject of much discussion within the philosophy and cognitive science of emotions concerning its perceptual and spiritual aspects, its cognitive aspects are not as clear. The main effort has been to clarify the effects this emotion has on cognition, notably a broadening of its structures to accommodate a perceptual content marked by beauty, vastness, and complexity of detail. However, emotions can have the same effect on cognition without thereby being cognitive emotions themselves. In an attempt to demarcate the emotion of wonder, we will advance a thesis that aims at specifying its cognitive dimension, namely: wonder is an emotion that is a constitutive part of a metacognitive process by which the agent becomes sensitive to the limits of her conceptual apparatus. What results from this process are precisely the typical sensations which accompany this emotion, such as belonging and reverence. The paper is structured as follows. First, we will bring to the fore the problems inherent in understanding the cognitive emotion of wonder. Second, we will evince a pertinent case of metacognition and show in what way that capacity can be constituted by emotions. Third, we will elucidate the metacognitive character of wonder. Finally, we will offer a case to make clear the explanatory potential of our analysis, namely, the role of wonder in the formation and maintenance of religious belief.

Resumo

Apesar de o maravilhamento ter sido objeto de muita discussão na filosofia e ciência cognitiva das emoções com respeito aos seus aspectos perceptuais e espirituais, seus aspectos cognitivos não estão tão claros. O principal esforço tem sido esclarecer os efeitos que essa emoção exerce na cognição, notadamente um alargamento de suas estruturas para acomodar um conteúdo perceptivo marcado pela beleza, vastidão e complexidade de detalhe. No entanto, emoções podem ter esse mesmo efeito sobre a cognição mesmo que não sejam, elas mesmas, cognitivas. Na tentativa de delimitar a

* I am deeply grateful to my colleague José Eduardo Porcher, who read, revised and commented on early versions of this work, and for his other efforts on my behalf.

\(^a\) Faculdade Jesuíta de Filosofia e Teologia (FAJE), Belo Horizonte, MG, Brasil. Doutor em Filosofia, e-mail: deluca.11@gmail.com
The problem

In the last few years, the emotion of wonder has received a lot of attention from philosophers and cognitive scientists. A considerable portion of the literature converges on the idea that wonder is a complex emotion because it involves different dimensions, namely, cognitive, perceptual, and spiritual. The cognitive dimension refers to the influence of the emotion of wonder in the processing of sensory information. The perceptual dimension refers to the agent’s access to some content marked by phenomenological richness and which captures her attention. Finally, the spiritual dimension refers to an attitude of reverence and inferiority and, at the same time, belonging concerning what is experienced.

These dimensions are not neatly separated from each other. Wonder typically involves a violation of cognitive expectations through the experience of a perceptual content marked by beauty, vastness, and complexity of detail. It mobilizes attention and interest in the experienced content, which causes the agent to engage in exploratory activities concerning the relevant objects. Such activities are characterized by a non-instrumental, usually contemplative, relation with the object in question. Openness and receptivity, common traits of the experience of wonder, often culminate in a feeling of belonging concerning the experienced phenomena. Finally, perceptual experiences that induce the emotion of wonder have an impact on cognition, expanding the agent’s conceptual repertoire (RUDD et al., 2012). In what follows we will be concerned mostly with the cognitive dimension of wonder, which
has received less attention than the perceptual and spiritual dimensions (FINGERHUT; PRINZ, 2018).

Let us consider the conditions an emotion has to fulfill to be conceived as a cognitive emotion. Traditionally, cognitive accounts take emotions to be a sort of appraisal or evaluative judgment. Strong cognitive accounts conceive of emotions as propositional attitudes (GORDON, 1987). Fear, for example, could thus be defined as a belief that there is some imminent danger coupled with the desire to avoid it. Weak cognitive accounts, in their turn, conceive of emotions not as propositional attitudes but as inextricably linked to them (SOLOMON, 2006). Thus, if an agent believes that something is wrong with her life, she may be in an emotional state of anxiety. Each in their way, cognitive accounts situate emotions in the larger web of propositional attitudes.

Here, then, is the problem: experiences which induce wonder, as we mentioned, are characterized by violating the agent’s cognitive expectations. Therefore, it does not seem as though wonder belongs to the domain of propositional attitudes. In particular, given that wonder goes beyond the agent’s conceptual repertoire, it is doubtful that it is inferentially related to the agent’s other propositional attitudes. On the other hand, it does not seem to be the case that wonder is a basic emotion capable of evincing involuntary reactions (i.e., immune to cognitive activity). Thus, if emotional states of wonder exceed the agent’s cognitive repertoire, how can we affirm that it is a cognitive emotion? Are not these experiences beyond our descriptive capacities?

The recent literature on wonder shows two ways of dealing with these questions. The first way is asserting that wonder is a cognitive emotion because it generates intentional and flexible commitments (NUSSBAUM, 2003). From an explanatory point of view, the appeal to cognition here serves to exclude cases of violation of expectations which are not liable to be characterized as wonder. After all, these cases may generate different kinds of reactions. Some of these are flight or self-protection responses, in which case the perceptual content is colored by the emotion of fear and possesses negative valence.

Many behaviors induced by emotions, especially basic emotions, are fixed or stereotypical. In these cases, there is no employment of cognitive capacities
whatsoever. However, the kinds of behavior caused by wonder are actions in response to interest and curiosity concerning experienced phenomena. One way to understand this flexibility is precisely by pointing to the impact of wonder in cognition. Note that this does not change when one takes these behaviors to be appraisals or evaluative judgments since one might say that these are cognitive appraisals concerning the content that impresses us by its beauty, vastness, the complexity of detail, and so on.

The second way concerns the alterations in cognitive processing due to wonder. According to Keltner and Haidt (2003), the violation of expectation generated by the perceptual content in a state of wonder is not assimilated, but rather accommodated. Assimilation occurs in cases of perception of trivial objects where information is simply added to existing conceptual schemes. The opposite happens in accommodation, where perceptual attention falls upon information which deviates from the agent’s conceptual schemes. This deviation leads to the updating or creation of new schemes so that such information can be stored in a minimally coherent way in the agent’s web of beliefs. Thus, experiences of wonder tend to amplify information processing. This modification, or amplification, of conceptual schemes, also demonstrates the cognitive dimension of this emotion.

There is something in common between these two proposals. Both attempt to highlight the cognitive aspect of wonder by pointing to the effects this emotion has on human cognition. However, pointing to the effects of emotion on cognition does not amount to showing that the emotion itself is a cognitive emotion nor, as we will see in the case of wonder, a metacognitive emotion. In sum, it does not show it to be a constitutive part of the cognitive process.

Consider affective empathy, the capacity to respond with an appropriate emotion to another’s mental states. It is known that affective empathy has an impact on social cognition (MAIBOM, 2017). We can, for example, develop affective empathy for people with whom we do not have any close relationships. This process, however, does not entail that affective empathy is cognitive empathy (i.e., the capacity to understand another’s perspective or mental state). Although it is a fact that affective empathy affects pertinent cognitive resources such as imagination and theory of mind (GOLDMAN, 2006, p. 113-140), those resources are not themselves constituted by
emotions: it is perfectly possible to imagine another's point of view without any affectivity involved.

Besides, other emotions also affect cognition without thereby being cognitive emotions. Emotional experiences with positive valence, such as joy, for example, can also broaden cognition (JOHNSON et al., 2010). For this reason, we believe neither of the two approaches outlined above does justice to the specificity of wonder's cognitive character. It is necessary to show, precisely, in which way wonder is a cognitive emotion.

**Emotions and metacognition**

We propose to investigate the cognitive character of wonder, showing that it is an emotion that not only exerts a causal pressure on cognition but is intrinsically metacognitive. However, this proposal is not based on traditional cognitive theories. We will argue that a distinctive feature of wonder is the monitoring of the agent's cognitive apparatus. In this way, we can understand the violation of expectations caused by a deviation of the perceptual content about the conceptual apparatus of the agent.

This self-monitoring is, of course, implicit. In wonder, what is available to the agent's consciousness is a perceptual content endowed with the aspects already mentioned, not her own cognitive resources. In the metacognitive process, the agent does not explicitly reflect on her cognition. To clarify the idea of implicit monitoring of our cognitive abilities, we will bring to light recent developments on the idea of metacognition.

Agents endowed with metacognition are those capable of representing or monitoring the cognitive system itself (PROUST, 2010). The metacognitive process involves different levels of complexity. The terms “represent”, and “monitor” clarify this difference. At a higher level, metacognitive ability explicitly represents the cognitive system. At the lower level, metacognitive ability implicitly monitors and controls the cognitive system.

Consider first the higher level. Here metacognition requires the possession of second-order propositional attitudes. Metacognitive ability is conceived as the self-attributitional capacity of propositional attitudes. For example, consider an agent who
assigns the following statement: ‘I know I intend to perform $p$.’ Note that the possession of explicit meta-representations, i.e. representations about representations, is a necessary condition for the formation of these propositional attitudes.

Part of the proponents of this explanation argues that such meta-representations are acquired through a common theory of mind (ToM), whereby we access other people's minds (BARON-COHEN, 1997). Whether based on a common theory of mind or simulation, self-attribution of propositional attitudes would be tributary to the attribution of propositional attitudes to others. In both cases, metacognitive abilities involve self-reflection and thus a high cognitive cost.

Recently, however, research in cognitive psychology has pointed to a more basic level of metacognition (PROUST, 2010; HAMPTON, 2009; KORNELL, 2009). At this level, there is implicit monitoring and control over one's cognitive performance. These metacognitive functions are partly responsible for flexibility in task performance but do not depend on the possession of meta-representations. In general, these studies postulate three central arguments:

First, meta-representations about cognition itself would affect the formation of prior intentions or long-term action strategies. For this reason, meta-representations would be inadequate to think of the cases in which agents need to know, here and now, whether or not they can perform a task, or how they should reorganize due to unexpected changes occurring in the course of action.

Second, it seems possible that an agent may have meta-representations about her own cognitive system without her having any control over her cognition. In this case, meta-representations are linked to self-knowledge or self-narrative, but it is not certain that such aspects may play a role in controlling cognitive processing.

Third, there are many cases of control and monitoring of the cognitive process itself from feedbacks occurring throughout a task in nonhuman animals, for example in some primates and dolphins, for which meta-representations have no adherence. Finally, what is at stake is a contrast between self-attributional and self-evaluating theories of metacognition (PROUST, 2010). The self-assessment level commonly occurs implicitly.
Given the differences between self-attribution and self-assessment, we can show that the emotion of wonder constitutes the basic level of self-assessment metacognition in cases of perceptual experience with vast content. To do this, we must first place our conception of emotions within a general framework, which is the theoretical framework of the somatic theory of emotions (JAMES, 1884; DAMÁSIO, 1999; PRINZ, 2004)\(^1\). Secondly, we will try to show a case in which emotions perform a metacognitive function. Finally, thirdly, we will clarify the metacognitive character of the emotion of wonder.

The conception of emotions to which we subscribe largely follows the somatic theory. According to this theory, emotions involve perceptions of states or changes in one's own body. The perceptive experience that results from this mapping, in turn, is precisely what we call emotions. A set of reactions endowed with a particular phenomenology is constitutive of this perceptual experience. Thus, experiencing an emotion of fear involves the perception of one's own body (i.e., proprioception). There are several somatic reactions pertinent to this experience, such as tremors, sweating, lack of salivation, among others. Such reactions constitute the emotion of fear and not joy, for example.

In the wake of Prinz (2004), we note that from the fact that emotions are linked to self-perception, it does not follow that they are devoid of directionality. Emotions are not merely perceptions of body states themselves, but body states in relation to the world. The emotion of anger includes not only a set of internal manifestations but also, for example, the perception of offensive behavior by others. In some situations, the emotion of fear can lead to flight behavior — consider distressing bodily manifestations in someone encountering a snake on a field trip. But in other situations, this willingness to escape simply does not occur. For example, in the case of someone whose emotions are aroused in a thriller. The emotion of sadness can be understood as a response to loss, or a set of associated properties such as grief or deprivation. Finally, as recent research shows, emotions are related to their contexts and objects (WILSON-MENDELEHALL et al., 2013).

\(^1\) Note that the implicit metacognitive level, which involves self-assessment, is compatible with somatic theory since it does not define wonder as a propositional attitude that can figure in appraisals or evaluative judgments.
It should be noted that the idea of emotions as self-perception of body states certainly does not imply the self-monitoring of cognitive processing. As we said, nonhuman animals like dolphins are capable of self-monitoring about their abilities to act, but it is not certain that they can monitor, even implicitly, their cognitive abilities. We need to show specifically how emotions can perform a metacognitive function in the strict sense. Given the lower level of metacognition and the conception of emotions described above, we can now tackle this task.

An interesting analysis of emotions with a metacognitive function was proposed by Koriat (2000). It shows us a way in which emotions or feelings can be part of metacognitive processing. Koriat has in mind the known phenomenon of having a name on the tip of the tongue, in which there is a "feeling of knowing" that one is close to retrieving information in one's memory. Koriat highlights three aspects of this phenomenon.

First, this feeling is associated with a cognitive state, concerning which the agent does not, however, access the content in question. There is a feeling of knowing the content, although it is not salient, given the difficulty of the agent to resume it. For this reason, the agent has no representations associated with this content, at least throughout the process. Second, and most importantly, this feeling involves implicit monitoring of the cognitive apparatus itself, without which, of course, the phenomenon we are investigating would not exist. Finally, the intensity of this feeling may cause the agent to continue straining or otherwise simply give up the task. Note that the feeling of having a name on the tip of the tongue not only monitors the cognitive domain, which in this case is a memory. Also, it exerts a causal influence on the agent's relationship with this domain, attenuating or increasing the effort to retrieve certain information.

Can we extend Koriat's analysis to cases of wonder? Consider, at first, a difference between the two cases. Certainly, the metacognitive emotion of having a name on the tip of the tongue is not, at least necessarily, induced by the perception of an external phenomenon. Moreover, in this case, there is a specific cognitive faculty, memory, which is the object of agent monitoring. On the other hand, it seems that there is no specific cognitive faculty monitored in instances of wonder. The literature on wonder points to a violation of the conceptual system of agents in general
Despite these differences, we believe that the common ground between the case analyzed by Koriat and the case of wonder is that the same metacognitive process occurs.

**The metacognitive character of wonder**

While it violates cognitive expectations, wonder is not a cognitive emotion in the sense proposed by traditional theories. That is, it is neither a propositional attitude nor is it inferentially related to propositional attitudes within the agent's cognitive system. On the other hand, contemporary explanations that take the emotion of wonder to be a mechanism to induce accommodation and thereby widening cognitive structures also do not do justice to the metacognitive aspect of wonder. However, these considerations must not lead us to take a wonder to be strictly a bodily manifestation. While violating cognitive expectations, wonder states must somehow be operative at the cognitive level (but not as appraisals, as we have seen). So how should we account for the metacognitive character of wonder?

Consider the wide view one might gaze upon as one contemplates the sunset from a mountaintop. The agent's visual field is characterized by sensory amplitude, endowed with different physical phenomena, a profusion of aspects, colors and so on. What is remarkable about this experience is that the sensory capacities in question are broader than the conceptual capacities. This leads to the idea that our conceptual repertoire does not entirely exhaust all the content given us in perception. As we have seen, it is precisely this insufficiency of the conceptual repertoire in view of the richness of perceptual content, in addition to its beauty, vastness or complexity of detail, that characterizes experiences of wonder.

Now we can say that this perceptual experience induces a metacognitive process that is constituted precisely by the emotion of wonder. That is to say, the explanation for the unfolding of wonder, such as the sensations of reverence and smallness in the face of vast content, must point precisely to a process in which the agent implicitly monitors her cognition. In this process, the agent is sensitive to its limits of information processing. Being sensitive to one's limits does not imply making
second-order judgments based on meta-representations. From there arises a set of sensations that can be understood as the effects of self-monitoring.

In their article, Shiota and colleagues sought to create an environment that would foster wonder. To do so, they subjected several people to situations of perceptual vastness, in which they would be prone to manifest belonging and reverence as a function of the stimuli received by perception. The result is that these people had more recurrent thoughts of the presence of something larger than themselves, of disregard for ordinary tasks and connection with the world (SHIOTA, KELTNER, MOSSMAN, 2007, p. 952-953). This points to the thesis that the sensations arising from wonder states come from a violation of expectation due to perceptually rich stimuli. We can now say that this violation of expectation necessarily requires self-monitoring, the result of which is the typical sensation that comes from wonder.

**A case: wonder and religious belief**

In this last section, we aim to show the explanatory potential of our approach to wonder by suggesting that our approach may be fruitful in dealing with the adherent aspects of religious belief.

Religious belief has specific characteristics that are not typical in empirical beliefs. The first is its transformative character: the acquisition of religious belief can profoundly alter people's habits. The second is its memorable character: religious belief commonly has a high degree of salience in memory that often manifests itself as mental images. The third is its evocative character: at various times throughout life, people evoke religious belief as an attempt to overcome difficulties or challenges. Taken together, these characteristics account for the adherence of this kind of belief, that is, for the fact that it remains unwavering even at times when the believer has counterarguments and contrary evidence before him (SPERBER, 1996, p. 90). What could explain the adherent character of religious belief?

In an attempt to answer this question, some researchers in the cognitive science of religion have advanced the argument that religious belief is a byproduct of cognitive mechanisms typical of the human species (ATRAN, NORENZAYAN, 2004; BARRETT, 2004; BOYER, 2003). One such mechanism is called the
hyperactive agency detection device (HADD), which operates from the earliest stages of human ontogenesis (BARON-COHEN, 1997). The strongest evidence of the presence of this mechanism is our extreme sensitivity to the intentional behavior of our co-specifics. It is an ability to detect by perception, directly or non-inferentially, intentions that manifest themselves in the behavior of others. Such sensitivity does not occur, on the other hand, in body movements devoid of directionality (WOODWARD, 2009).

This ability evolves during ontogenesis until agents acquire a common theory of mind, which we have alluded to, by which they assign mental states to predict or explain the behavior of others. Even with a theory of mind, how we attribute mental states to others as propositional attitudes also typically requires observation of intentional behavior by which these states are inferred. Cognitive science of religion makes use of this evidence to argue that religious beliefs are a byproduct of human cognition, resulting from a specific violation of our intuitive ontology by the presence of an intentional element that in turn makes the content of belief salient to the attention of agents (BOYER, 2003).

As an example, consider the extraordinary concept of a being endowed with mental states, such as intentions, desires, and beliefs, but devoid of a body. Consider furthermore that it is an omniscient being. We have said above that we are inclined to detect intentions and other mental states by observing a body whose actions are directed at something in the world. This means that in our intuitive ontology, mental states depend on a body in a relationship, epistemic or practical, with something in the physical world.

Now the concept of a being endowed with mental states but without a body violates our intuitive ontology, all the more so because of the predicate of omniscience. In fact, in the process of detecting the intentionality of others, the observed agent is related to a certain portion of the world, given its point of view. This view delimits the content accessed by it. For this reason, the predicate of omniscience is counterintuitive to ordinary interpreters. In a word, this

---

2 In the philosophy of mind, such intentions are known as motor intentions. This type of intention is not the point of arrival for practical reasoning but is present in the agent’s engagement with the environment, his habituation with objects that allows him to adjust his behavior in the world. In this sense, it is an intention that manifests itself directly in behavior. See Pacherie (2006).
counterintuitive content becomes salient precisely because it involves mental aspects. Note, moreover, that we have here an explanation of the adherence of religious belief based strictly on human cognition, i.e., an explanation that does not take emotions into account.

However, researchers such as Fuller (2006), Atran (2002) and Pyysiäinen (2001) have shown, in different ways, that emotions should also make up this explanans. This is precisely due to the conception that religious belief, unlike empirical belief, has the particularities described above. According to Pyysiäinen (2001, p. 84), it is important to note that counterintuitive content is involved in an aura of mystery, which is present in the belief in a being that, independently of a physical body, has mental states, such as intentions. We can observe this in contexts of religious belief transmission: the mysterious character of counterintuitive content arouses emotions of fear, joy, relief, and so on.

How can we account for the emotional character of this counterintuitive content? To avoid a bifurcation in the general framework of religious belief, particularly between emotions and beliefs, one way is to show the presence of emotions in the cognitive domain. Thus, we would like to suggest that our metacognitive approach to wonder may contribute to making this picture coherent, at least in circumstances where counterintuitive content is endowed with beauty, vastness or complexity of detail.

Consider that the mysterious character of the counterintuitive content under discussion, consisting of omniscience and mental states without a body, can be understood as an instance of vast content. Although it is not a content available for perception, let us remember that sensory experiences are not the only causal source of the metacognitive process constituted by wonder. This process can also be induced by testimony3. As we have said, the breach of expectation of our intuitive ontology, induced by a content of vastness, involves sensitivity to one's own cognitive limits, the results of which are already known sensations, which we group here under the label 'awe'. Therefore, when induced by a causal impact with vast content, the emotion of wonder is a metacognitive process.

---


Rev. Filos., Aurora, Curitiba, v. 31, n. 54, p. 818-832, set./dez. 2019
Now, this allows us to observe a circumstance in which religious belief has an adherent character precisely because it results from a metacognitive process colored by wonder. In addition to a violation of our intuitive ontology through counterintuitive content, what may account for the adherent character of religious belief is that this content arouses our emotion of wonder. Due to the sensations that unfold from this metacognitive process, religious belief, as a byproduct of this process, manifests itself as adherent.

Finally, there are indeed different ways of acquiring religious belief. It should also be noted that other emotions may also be present in the causes of religious belief. In any case, our aim was only to indicate the importance of our approach to understanding wonder. Since it is a metacognitive emotion, it can help to make coherent the explanatory framework in which emotional and cognitive elements are related.

References


JAMES, W. What is an emotion? *Mind*, v. 9, n. 34, p. 188-205, 1884.


