Neuroscience and Biobehavioral Reviews 55 (2015) 88–97

Contents lists available at ScienceDirect

Neuroscience and Biobehavioral Reviews

journal homepage: www.elsevier.com/locate/neubiorev

Review

There are things that we know that we know, and there are things that we do not know we do not know: Confidence in decision-making

Piercesare Grimaldi[a,b,c,d,*], Hakwan Lau[b,d], Michele A. Basso[a,c,d]

[a] Departments of Psychiatry and Behavioral Sciences and Neurobiology, University of California, Los Angeles, CA, USA
[b] Department of Psychology, University of California, Los Angeles, CA, USA
[c] The Semel Institute for Neuroscience, University of California, Los Angeles, CA, USA
[d] The Brain Research Institute, University of California, Los Angeles, CA, USA

ARTICLE INFO

Article history:
Received 18 August 2014
Received in revised form 19 April 2015
Accepted 20 April 2015
Available online 28 April 2015

Keywords:
Metacognition
Confidence
Consciousness
Awareness
Monitoring

ABSTRACT

Metacognition, the ability to think about our own thoughts, is a fundamental component of our mental life and is involved in memory, learning, planning and decision-making. Here we focus on one aspect of metacognition, namely confidence in perceptual decisions. We review the literature in psychophysics, neuropsychology and neuroscience. Although still a very new field, several recent studies suggest there are specific brain circuits devoted to monitoring and reporting confidence, whereas others suggest that confidence information is encoded within decision-making circuits. We provide suggestions, based on interdisciplinary research, to disentangle these disparate results.

Contents

1. Introduction

Thinking about our own thoughts and knowledge – encapsulated by the infamous quote from the former US Secretary of Defense, Donald Rumsfeld and paraphrased in the title of this review – is referred to as metacognition. How we know what we know has captured the interest of philosophers since ancient times. Aristotle (1987), in his *De Alma*, speculated that the act of judging our own thoughts is necessary for remembering.

The ability to reflect upon our own thoughts has been considered a logical conundrum for centuries. The idea of recursive monitoring evokes the image of a looker inside the looker, which implies infinite regression. This idea held back advancement of this field for a long time and induced a rational thinker like Descartes, to propose a disembodied self, a soul to solve the problem (Descartes, 1999). Centuries later, French philosopher Auguste Comte thought that the notion that an individual can turn his mental faculties inward is logically impossible (see James, 1983 for a discussion). It is paradoxical, Comte argued, that the mind might divide into two minds to allow self-observation. If there is a looker inside the looker, he reasoned, there needs to be a third looker and so on.

* Corresponding author at: Departments of Psychiatry and Biobehavioral Sciences and Psychology, The Semel Institute for Neuroscience and Human Behavior, UCLA, 635 Charles Young Drive Box 957332, Los Angeles, CA 90095, USA. Tel.: +1 3109833418.
E-mail address: pcgrimaldi@ucla.edu (P. Grimaldi).

http://dx.doi.org/10.1016/j.neubiorev.2015.04.006
0149-7634/

In modern neuroscience we commonly assume that the brain is modular (Fodor, 1983), so it is reasonable to postulate a module specialized in monitoring other modules. Paraphrasing Humphrey (2003): "No one would say that a person cannot use his own eyes to observe his own feet". Although considering the brain as modular possibly saves us from the need to postulate a mind inside the mind, we still do not know how metacognition is encoded in the brain. Is it implemented by the same circuits that encode decisions or do specialized modules (a looker) monitor the activity of decision-making areas?

Today, the introspective nature of metacognition is considered a core part of what makes us human and a necessity to form the basis of conscious awareness (Terrace and Metcalfe, 2005). The role of metacognition in learning and memory drives research in the field of educational psychology (Bransford et al., 2005). Neuroscientists are also beginning to see metacognition as a topic amenable to inquiry.

Here it is important to make a distinction between metacognition and consciousness. Metacognition in perception is linked to mechanisms of conscious perception. The exact relationship between the two may be complicated, and in general, issues surrounding consciousness are controversial (Lau and Rosenthal, 2011). Here we define metacognition as one specific aspect of consciousness, namely the ability for one to introspectively appreciate or monitor the quality of an ongoing perceptual process. Metacognition is not synonymous with consciousness because the latter is associated with a wider plethora of concepts including wakefulness, arousal, self-awareness, control of action, etc. However, metacognition is related to consciousness as is seen in neurological cases of the abolishment of perceptual awareness, such as occurs in blindsight. In these patients, metacognition seems to be impaired: blindsight subjects fail to report confidence in their responses even though their responses, i.e. guesses, reflect good perceptual capacity (Ko and Lau, 2012). In the modern memory literature, metacognition starts to be treated as a scientific subject as early as Hart (1965); see Tulving and Madigan (1970), for a review. In this context, metacognition is often described as either prospective or retrospective (Fleming and Dolan, 2012). Prospective metacognition refers to making judgments or predictions about what information will be available in memory in the future. Retrospective metacognition, in contrast, involves making judgments about a past experience, specifically about whether a memory item has been successfully encoded. How does metacognition as discussed here relate to introspection exactly? Introspection, which has been used for centuries by philosophers to explore our internal word, relies on the insight of the subject. Metacognition, as it is considered in modern literature, depends more on an operational definition, determined by reported levels of confidence in perceptual or memory responses given by the subject. Of course, in assessing metacognition, the hope is that the subjects' reported level of confidence arises from introspection of the ongoing perceptual or memory process. However, one needs to be careful to rule out that reports of confidence are not driven by other factors such as social pressure too (Asch, 1951).

Although metacognition broadly refers to the psychological processes used to plan, monitor and assess an individual's own thoughts, knowledge and performance, here we shall focus on retrospective metacognition, and more specifically, on one aspect of it, namely confidence about the correctness of a decision.

In this review, we explore recent work in the field. We begin by introducing a definition of confidence and how it is measured. We then provide a brief history of ideas that form the basis for current approaches designed to understand confidence in decision-making. Next, we describe recent experimental data that are beginning to unravel the neural basis of confidence in decision-making. Finally, in light of existing conflicting data, we provide some suggestions for how to untangle the question of how the brain encodes confidence. Because much of the work on the neuronal basis of confidence capitalizes on developments in our understanding of oculomotor circuitry and visual perceptual decision-making, we will focus primarily on confidence in the context of vision.

2. What is confidence and how can it be measured?

Confidence is a belief about the validity of our own thoughts, knowledge or performance and relies on a subjective feeling (Luttrell et al., 2013). However, in the past decades, several methods have been developed to measure confidence objectively. Until recently, confidence was studied mainly in humans. Recent work, which we review below, suggests that non-human animals may also experience some levels of confidence in decision-making. In this paragraph, we will first review how confidence is measured in humans and then discuss the recent advances in how confidence is measured in animals.

In studies of human perceptual decision-making, confidence is often measured with retrospective judgment. Subjects give a confidence rating right after a report about a perceptual experience and therefore must base their confidence judgment on the memory of their initial response. For example, a subject might first perform some perceptual task such as reporting their perception of an ambiguous object (do you see a vase or a face?). Then the subject would immediately declare how confident s/he felt about that decision.

Similar to measures of confidence using open-ended ratings, several scales have been developed to measure confidence more quantitatively. The most commonly used is confidence rating. In this scale, the subject is asked to report confidence on a continuous scale ranging from 0% or complete uncertainty to 100% or complete certainty. Alternatively, it can be assessed with discrete fixed levels, or a simple binary choice (confident/not confident, Cheesman and Merikle, 1986; Dienes and Perner, 1999). However the use of ratings has been criticized because some subjects may find it not intuitive or they may be poorly motivated to accurately report their confidence (Persaud et al., 2007). To overcome these limitations, post-decision wagering has been introduced, in which subjects bet money or tokens on their own decisions (Persaud et al., 2007; Ruffman et al., 2001). In this context, subjects should ideally bet low when they are not confident and bet high when they are confident, in order to maximize gain. This task is more engaging and more intuitive for most participants. However, it has been noted that wagering can be influenced by individual propensity to risk (Fleming and Dolan, 2010) and that subjects tend to use only the ends of the scale, probably in order to maximize gains (Sandberg et al., 2010), thus suffering from low sensitivity for intermediate ranges. In an attempt to develop a scale that has both the sensitivity of confidence ratings and the intuitiveness of post decision wagering, the feeling of warmth scale has been developed (Metcalfe, 1986; Wierzchon et al., 2012). In this scale subjects report their confidence as a temperature, ranging from cold (not confident) to hot (very confident), with intermediate options (e.g. chilly or warm). The perceptual awareness scale (Ramsøy and Overgaard, 2004) and the Sergent–Dehaene scale (Sergent and Dehaene, 2004) are also commonly used and were developed to judge the degree of visibility in visual tasks, ranging from no visibility at all to clear perception, with discrete intermediate levels (perceptual awareness scale), or a continuous spectrum (Sergent–Dehaene scale). When applied to confidence, however, these two scales end up being very similar to confidence rating. An extensive discussion of the properties and sensitivities of the different scales is beyond the scope of this review. For a rigorous comparison see, Sandberg et al. (2010) and Wierzchon et al. (2012).