

Abstract: “Unconscious Working Memory?”

The construct of working memory plays a major role in psychology and neuroscience. Working memory is conceived of as a short-term store responsible for the maintenance and manipulation of information no longer present in the immediate environment. It is typically assumed that there is a tight link between working memory and consciousness, such that only conscious information can be encoded in working memory (Baars 2003). In the past few years, however, this assumption has come under attack: several recent studies suggest that working memory can operate on fully unconscious information (Soto et al. 2011; Soto & Silvanto, 2014; Bergström & Eriksson, 2014, 2015). This has given rise to a lively debate, with some researchers endorsing unconscious working memory, and others staunchly denying it. The general strategy of deniers is to offer debunking explanations that can account for the relevant empirical data without appealing to unconscious working memory (Stein et al. 2016; Prinz 2012). My aim in this paper is to show that the proposed debunking explanations are unconvincing, and thus that we ought to take the notion of unconscious working memory seriously.

For the purposes of the present paper, I focus primarily on the landmark study by Soto et al. (2011), which provides one of the first alleged demonstrations of unconscious working memory. In Soto et al.’s study, participants were briefly presented with a masked (and often invisible) line grating of a particular orientation (the “memory cue”), followed by a delay of 2-5 seconds. After the delay, participants were then shown a second test grating and asked whether it had been rotated left or right relative to the initial memory cue (they were also asked to report their awareness of the memory cue on a scale of 1-4). Crucially, Soto and colleagues found that participants performed at above-chance levels on the task, even on trials where they reported being wholly unaware of the memory cue. This suggests that participants were at least sometimes able to retain unconscious information about the memory cue’s orientation over the delay. According to the authors, the most plausible interpretation of these results is that participants held the grating in unconscious working memory.

Three alternative debunking explanations have been proposed in the literature:

- 1) Participants could have falsely reported *not* seeing the memory cue (i.e. given a “1” rating) even when they did in fact see it. Thus, the results

obtained in the “unconscious condition” might simply reflect contamination from conscious working memory (Stein et al. 2016).

- 2) Participants could have immediately guessed the orientation of the memory cue, using blindsight-like unconscious perception, and then retained the conscious guess in working memory over the delay (Stein et al. 2016).
- 3) Participants could have been relying on a form of sensory memory, such as fragile visual short-term memory (Prinz 2012).

As I argue, however, none of these alternative explanations are convincing. Explanation (1) is rendered implausible by internal features of Soto et al.’s study. Specifically, given the detailed nature of the perceptual awareness scale used in the study (1 = did not see anything; 2 = maybe saw something; 3 = saw the stimulus but not its orientation; 4 = saw the stimulus and its orientation) it is highly unlikely that participants would have selected “1” if they were at all unsure as to whether they saw the memory cue. Explanations (2) and (3) require more detailed comment, but ultimately they too are shown to fail. I conclude that we ought to take the results of Soto et al.’s study at face value: participants are likely relying on unconscious working memory after all.

Word Count: 596

References

- Baars, B. J. (1997). Some essential differences between consciousness and attention, perception, and working memory. *Consciousness and cognition*, 6(2), 363-371.
- Bergström, F., & Eriksson, J. (2014). Maintenance of non-consciously presented information engages the prefrontal cortex. *Frontiers in human neuroscience*, 8, 938.
- Bergström, F., & Eriksson, J. (2015). The conjunction of non-consciously perceived object identity and spatial position can be retained during a visual short-term memory task. *Frontiers in psychology*, 6.
- Prinz, J. (2012). *The conscious brain*. Oxford University Press.
- Soto, D., Mäntylä, T., & Silvanto, J. (2011). Working memory without consciousness. *Current Biology*, 21(22), R912-R913.
- Soto, D., & Silvanto, J. (2014). Reappraising the relationship between working memory and conscious awareness. *Trends in cognitive sciences*, 18(10), 520-525.
- Stein, T., Kaiser, D., & Hesselmann, G. (2016). Can working memory be non-conscious? *Neuroscience of Consciousness*, 1, 1-3.