



Prof. Dr. Evie Vergauwe
Associate professor (SNSF Eccellenza Fellow)
Evie.Vergauwe@unige.ch

Working memory, Cognition and Development Lab
<https://www.unige.ch/fapse/womcogdev/>

Genève, Switzerland
14/07/2021

**Evaluation of PhD thesis of Zuzanna Skóra
“Conscious access to short-term memory. From theory to measurement”
Jagiellonian University in Kraków**

Overview

The doctoral thesis under evaluation consists of 5 parts, presented in 5 corresponding chapters. The first chapter is a general introduction to (1) the topics of attention and consciousness, (2) the question how these two concepts are related, and (3) how working memory (WM) theory could serve as a framework towards answering this long-standing question. The chapter ends with an overview of the research questions and of the work presented in the subsequent chapters. Chapter 2 presents a series of 4 studies (plus a pilot study) that aims at answering the first research question, i.e., whether information in working memory that is held outside of the focus of attention is consciously accessible. Elements from typical working memory studies (e.g., change detection tasks and continuous delayed estimation tasks) are combined with elements from typical consciousness studies (e.g., the Perceptual Awareness Scale, or PAS) in these experiments. Next, the focus of the thesis shifts towards answering the second research question, i.e., to what extent the PAS is a valid measure of the perceptual quality of an experience (and, as such, to what extent the PAS is a valid measure of conscious access). Chapter 3 presents a brief commentary on this topic, and Chapter 4 presents a series of 3 studies that aims at assessing the validity of the PAS empirically, by manipulating potential sources of bias as well as assumed sources of perceptual quality. The final chapter, Chapter 5, presents an overview and brief discussion of the results presented in the previous chapters.

Chapter 1

Chapter 1 provides a generally good overview of the relevant literature, is well written, and demonstrates nicely how working memory theory could be useful as a framework to examine the question of whether attention is necessary for consciousness. In particular, it is proposed that a particular state of working memory, whereby information is held outside the focus of attention, is of interest here, because it allows testing whether information held in that state (i.e. outside of attention) is consciously accessible or not. This constitutes a clever, novel, and original way of addressing the relationship between attention and consciousness. There are, however, some specific aspects of the implementation of this approach that deserve more discussion.

For example, a clear distinction is drawn between working memory and short-term memory (STM), but the proposed distinction is somewhat confusing, because the part of the Cowan model that is typically referred to as activated long-term memory is referred to as short-term memory in

this chapter, and because there are some confusing statements about the relationship between WM and STM in this chapter (“WM on the other hand will be an overarching term ..., involving STM as one of its states”, on p. 15, vs. working memory being associated to the focus of attention in Cowan’s model on p. 19; the first indicates that STM is included within WM, whereas the latter indicates that WM is embedded within STM).

Moreover, it is argued that there is a well-developed methodology in place to capture the different states in working memory, but the fact that different models disagree on the number of items held in the focus of attention and the notion that the focus of attention can zoom in and out, suggests that the argument of a well-developed methodology might not be entirely true. For example, when using a one-item retro-cue, would Cowan assume that there is only 1 item in the focus of attention or that all items are in the focus of attention with one item receiving more resources/using more slots? This is somewhat addressed on p. 23, but I am not entirely convinced that it is possible to claim that uncued items equals unattended items, at least not in the strong way as is necessary here to test the accessibility of unattended items in working memory.

Chapter 2

Chapter 2 presents a pilot study as well as four original, well-designed experiments. The rationale for the different experimental manipulations in the main experiments is well explained, except that it was not entirely clear to me how the results of the pilot study have informed the design of the four main experiments. That is, the pilot study aimed at examining the trade-off between quantity and quality in the attended part of working memory, but why is this examination needed, and which aspect of the results have led the candidate to make specific decisions regarding the main experiments? Most, if not all, main experiments were preregistered, a clear distinction is made between planned and exploratory analyses, and the candidate shows critical thinking when it comes to interpreting the results; several alternative accounts are considered and new, more direct manipulations are proposed when necessary. The series of experiments also shows great experimental insight.

I do have some questions related to Experiment 4, where the candidate moves away from the change detection task and proposes to use the continuous delayed estimation task: (1) It is not entirely clear why exactly this should provide a more fine-grained performance measurement, (2) the candidate proposes to test whether the observations of Experiments 1-3 regarding the relationship between focused attention and conscious accessibility will generalize to the more fine-grained task of Experiment 4, but since there is not really a consistent pattern that emerges from Experiments 1-3, it is not clear what exactly one expects to generalize, and (3) if spatial cuing is known to not be very precise because of some spilling, then why not consider whether other types of cuing exist that are more precise?

Finally, Experiments 1-3 seem to suggest that, overall, unattended items are not consciously accessible, whereas Experiment 4 indicates that all items in WM, including the unattended items, are consciously accessible. Why should one conclude that the accessibility of unattended items depends on the task? Could one not argue that more weight should be given to Experiment 4, since the main reason for using continuous delayed estimation in Experiment 4 was to have a more fine-grained measure?

Chapters 3 and 4

Chapter 3 presents a commentary in which it is argued that the PAS can be a valid measure of consciousness, depending on which theoretical stance the user supports. It is then recommended that future users of the PAS should be explicit about their theoretical assumptions and should explain how the PAS measure is interpreted within the chosen theoretical framework. Chapter 4 continues on this issue and proposes to investigate the theoretical claims behind the PAS in a

series of experiments. In particular, the candidate proposes to examine how responses in the PAS are influenced by the presence of an objective task and by perceived task difficulty, and to what extent changes in PAS ratings reflect changes in perceptual quality. It is explained that this latter question can be answered without referencing to a particular consciousness theory, as opposed to answering to the broader (and perhaps more important) question of whether PAS ratings reflect changes in consciousness. While I understand the argument, not addressing the broader question of whether PAS is a valid measure of consciousness comes a bit as a surprise, especially after the commentary in Chapter 3 which argued for PAS as a valid measure as long as one is explicit about the theoretical framework.

Three well-motivated and well-designed experiments are presented. One aspect that was not entirely clear to me, was why a delay of 1000 ms between presentation and recall is needed if the aim is to assess perceptual quality of the colors. Like in Chapter 2, a clear distinction is made between planned and exploratory analyses, and the candidate shows critical thinking when it comes to interpreting the results. However, the joint discussion of the three experiments reported in Chapter 4 remains rather implicit regarding the implications of the findings for the validity of the PAS.

Chapter 5

Chapter 5 provides a brief overview of the work presented in the previous chapters, together with the corresponding conclusions. The chapter is well-written and easy to follow. However, it does not go much beyond a summary of the results and interpretations reported in the previous chapters. While this summary is useful, it would have been interesting to go further by also including, for example, (1) a critical consideration of the ensemble of the findings (e.g., what are the implications of the observations in Chapter 4 when it comes to the interpretation of the observations in Chapter 2?), (2) a discussion of what the current findings imply for theories of working memory, attention, and consciousness (e.g., more easily explained by one particular model? Are modifications needed to the Cowan model to account for some findings?), (3) a discussion of the usefulness of working memory theory as a framework for examining the relationship between attention and consciousness (e.g., is it as useful as anticipated?), and (4) an explicit discussion of the answers to the research questions, their strength and limitations, and what they contribute to the literature.

Conclusion

My general evaluation of the doctoral thesis is positive, and I recommend acceptance. The research questions are timely, interesting and highly relevant. The work brings together two subfields of cognitive psychology in an elegant way. The candidate demonstrates good knowledge from both subfields and takes advantage of these to propose a novel way of investigating an important issue. The experiments are well-designed and the analyses appear sound. Overall, the proposed interpretations of the results are well-presented, critical, and reasonable, although a more elaborated, joint critical evaluation of all findings reported in the thesis and their theoretical implications would have been welcome in the general discussion. Finally, the doctoral thesis is well-structured and well-written, the figures and tables are useful and of very high quality, and the relevant literature is cited and explained. I enjoyed reading the work and am convinced that it will serve to advance our understanding of the relationship between attention and consciousness.

Pr. Evie Vergauwe

