On the interplay between working memory consolidation and attentional selection in controlling conscious access: parallel processing at a cost—a comment on ‘The interplay of attention and consciousness in visual search, attentional blink and working memory consolidation’

Brad Wyble¹, Howard Bowman² and Mark Nieuwenstein³

¹Department of Psychology, The Pennsylvania State University, University Park, State College, PA 16801, USA
²School of Psychology, University of Birmingham, Birmingham B15 2TT, UK
³Department of Experimental Psychology, University of Groningen, Groningen, The Netherlands

Theories of human cognition generally agree in assuming the existence of a fundamental limit in information processing, but these theories are challenged by findings that appear to circumvent these limitations. We argue that theories are needed that offer a principled account of why the limitations become manifest in some cases but not others.

Raffone et al. [1] present a comprehensive theoretical framework that links behavioural effects, such as the attentional blink (AB), visual search and inattentional blindness, to an overarching theory of the neurobiological substrates of consciousness and attention. Central to their model is the global workspace (GW), which is predicated on the notion that working memory, attention and consciousness arise from a network of neurons that can only focus on one representation at any time. This inability to simultaneously process multiple stimuli has been suggested to explain the AB as well as visual search behaviour.

However, there is mounting evidence that the brain is capable of encoding multiple items into short-term memory (STM) at once, such as lag-1 sparing, in which two targets are apparently encoded together [2]. Owing to its strictly winner-take-all behaviour, previous GW-based computational accounts of the AB [3,4] had particular difficulty explaining this effect.

Raffone et al. [1] respond to this evidence by attempting to reconcile the GW theory with such evidence. The theory of attentional control (TAC) boldly pushes the unitary nature of cognition even further by proposing that STM encoding and attentional selection also time-share central processing capacity, so that only one of them can occur at a time. TAC permits lag-1 sparing through an intermediate buffer that can hold multiple items concurrently, as they await an opportunity to be encoded into STM. Thereby, the AB is not caused by direct competition between target one (T1) and T2, but rather by competition between attentional selection and STM encoding.

In this respect, TAC is similar to computational models that we have previously proposed [5–7]. In the episodic Simultaneous Type Serial Token model (eSTST), ongoing consolidation has an inhibitory effect on attentional selection. Furthermore, like TAC, we proposed that attention can remain engaged over several consecutive targets to produce spreading of sparing (see [8] and also figure 1).

However, the theories differ dramatically in that the eSTST model permits parallel encoding of multiple targets. Furthermore, attentional selection is possible during encoding; it is just reduced in strength. We argue that permissiveness of parallel processing is essential to accommodate the full range of available data concerning the attentional blink. In particular, there are two specific empirical benchmarks that seem unattainable by TAC.
The first benchmark is that providing a cue during the deepest part of the AB has a facilitatory effect on the following T2 [9] and no effect on T1 accuracy. According to TAC, encoding and selection can never co-occur, making it hard to explain how T2 can be successfully cued without aborting encoding of T1.

The second benchmark is the time course to initiate encoding during lag-1 sparing. The TAC has an intermediate buffer that allows T1 and T2 to be maintained in a purgatorial state, half-way to consciousness, where they persist until encoding is triggered by occurrence of a non-target item. Enabling serial encoding of multiple targets through such a buffer is certainly an ingenious proposal and has precedent [10,11], but the TAC account is contradicted by electrencephalogram (EEG) data.

TAC explains data at the level of neural mechanisms, which allows it to interface with EEG data reflecting the time course of cognitive processes. For example, the P3 component of the EEG is thought to reflect encoding into STM [12–14].

TAC assumes that T1 memory encoding is postponed when it is immediately followed by a T2. If true, then a lag-1 trial should produce a P3 onset that is delayed by 1 lag (e.g. about 100 ms), and even longer when even more targets are placed in sequence. However, existing EEG data indicate that the onset of the P3 for two targets is not delayed at lag-1 when compared with lag-8 (figure 2).

Furthermore, this buffer imposes a restriction that conscious perception of a critically important stimulus would be delayed if it were followed by another important stimulus. This buffering of input until a period of ‘downtime’ seems counterintuitive for a system with a requirement to process information as rapidly as possible.

In contrast to TAC, other models [5,11] posit that T1 encoding begins as soon as sufficient evidence has accumulated, without waiting for a non-target item, and such models can simulate EEG data from lag-1 and lag-8 trials [15].

We conclude with what seems, to us, to be the essential conundrum of the AB and lag-1 sparing: what is the adaptive utility of suppressing attention during encoding? Raffone et al. [1] propose that such suppression prevents interference from ‘feature and proto-object representations’. This seems an insufficient answer, because lag-1 sparing indicates that the system is capable of encoding multiple items in parallel, with relatively little apparent interference (at least on measures of identification accuracy). Our modelling work suggests that the suppression of attention achieves a higher-order goal, viz., enhancing the episodic information present in the stream of visual input. In other words, the visual system compiles information into temporal chunks, which are passed to STM as discrete episodes roughly equivalent to the duration of visual fixations (approx. 200–300 ms; see also [16]). Multiple items can be encoded in an episode, but at the cost of a loss of temporal information and a decreased ability to encode repetitions. This loss of episodic information is the price that is paid when stimuli are processed in parallel. We suggest further that such an episodic attentional system is a good solution to the problem of encoding information from a series of sequential eye fixations [17].

References


