The need for a separate short-term memory system won't go away just by calling it activated long-term memory: A reply to Cowan (2019)

Dennis Norris

MRC Cognition and Brain Sciences Unit,

15 Chaucer Rd.,

Cambridge CB2 2EF, U.K.

dennis.norris@mrc-cbu.cam.ac.uk

Abstract

In Norris (2017), I explained why the notion of activated LTM (long-term memory) combined with a focus of attention (FoA) was unable to perform the computations required to support STM (short-term memory), and argued that those extra computations must require a separate STM system. Cowan (2019) makes the alternative proposal that this full set of computations is better conceptualized as a unitary system of activated long-term memory (aLTM). To this he adds a pointer system, the ability to perform variable binding, and an unspecified model of STM which acts as a front end to LTM. This appears to be simply an exercise in relabeling. Furthermore, without a computational specification of how the components work, the model lacks the ability to simulate even the most basic STM phenomena. If the model were specified in more detail it seems almost inevitable that it would contain something instantly recognizable as an STM system.

When I quoted Cowan and Chen (2008) as saying that "although the mechanisms of short-term memory are separate from those of long-term memory, they are closely related" (their p. 104), I thought we had little to disagree about. I also consider Cowan's idea of a focus of attention (FoA) to be a useful way of thinking about control processes in working memory. So why is it that Cowan (2019) now feels the need to respond to the arguments I presented in Norris (2017)? The framework he presents builds on Cowan (1999) and Cowan (1988), and he now argues that there is no separation between short-term and long-term memory systems after all. He does this by endowing both the FoA and aLTM (activated LTM) with additional properties and an unspecified model of STM. In order to dispense with a separate STM he then deems all of this to be a unitary LTM system. Cowan has not shown that there is no need for a separate STM system, he has just redefined it out of existence.

Norris (2017) highlighted the inadequacy of models that embody the claim that short-term memory can be supported entirely by activated long-term memory (aLTM). In that article I explained that the computational requirements of storing information over the short-term require additional mechanisms to activation. In particular, memory must be able to support the construction of novel representations that have no pre-existing representation in LTM. At the very least, this requires the ability to store multiple tokens of a given type (the problem of two), and to perform variable binding. Any LTM system must be supplemented by extra mechanisms that are required to store information over the short term. Consider the need to store multiple tokens of a given type. I gave the example of the easily remembered sentence "buffalo buffalo buffalo buffalo buffalo" and argued that it was implausible to assume that LTM stored five copies of the phonological word form /bʌfələʊ/, just in case one was ever asked to repeat this sentence. The binding problem can be exemplified by the sentence "The young boy saw the boy who was singing". As I noted

“Here the problem is not simply representing the order of the words, but appreciating that there are two different boys, one of whom is singing and one of whom is young. It is necessary to both represent multiple tokens and the bindings between each of those tokens and other components of the sentence”. Similar problems arise in the case of visual STM, where memory for an array of random dots of the same size requires binding multiple tokens of dots to their locations.

Cowan (2019) agrees that the simple notion of STM as activated LTM cannot solve these problems. But even in the *Abstract,* he concedes the need for an STM system: “models of STM storage can serve as the front end of an LTM learning system rather than being separate”. The point is reinforced later when he says that: “The viability of an approach involving aLTM with new learning does not depend on coming up with a separate serial order memory model specifically within the embedded-processes framework, inasmuch an adequate model of serial order memory in STM formulated by another investigator also could also serve as the long-term learning mechanism.” (p. 30). Cowan recognizes the need for a model of STM but instead of considering it a separate STM system, he wishes to take an existing model of STM and call it a front end to aLTM instead. In Table 1 he states “separate-STM-copy theories might be reclassified as the front end of long-term learning”. This is simply an exercise in relabeling.

Although Cowan’s acceptance of the need to incorporate a model of STM has undermined his own case for calling his theory aLTM, it is still worth taking a closer look at the theory to examine how the various components work together, and how they relate to standard two-store models. First it should be noted that for more than 25 years there has been an expectation that new theories of STM should be presented as computational models (for review see Hurlstone, Hitch, & Baddeley, 2014). This has at least two advantages. First, we can run simulations to assure ourselves that the model really can simulate the target data-set. Second, it doesn't matter too much if we can't pigeon-hole the model into categories such as 'two-store', 'one store', 'uses activation', or the like. We can see how it works, and then, should we wish, assign it an informative label. Cowan reverses this process. He starts with a label, and then asserts that there could exist a model that might be labeled in this way. If the final step were to construct a computational model that fitted the label and simulated the data, I would have few complaints. I might be tempted to question whether the label seemed appropriate, but at least I would have been convinced that the model works, which is the really important thing. Could there be model that works according to the principles Cowan espouses? Given that the 'model' is only expressed verbally, we cannot be sure.

Cowan's unitary model has five core features, activated LTM, a focus of attention (FoA), a pointer system, rapid learning, and a model of STM. With these extra features aLTM is now assumed to be able to perform variable binding, to instantiate multiple tokens, and to create temporary representations that can support performance in STM tasks. All of the extras are there to fulfill the function of a short-term memory system that performs distinct computational functions from LTM. They allow Cowan to smuggle STM into LTM.

Activation

The core feature of aLTM is activation. Norris (2017) suggested that "it seems reasonable to ask what computational function is performed by activation that enables it to encode, maintain, and retrieve information from STM" (p. 998.) Cowan replies that "Activation, then, is simply the degree of availability for retrieval" (p. 38). But this doesn't answer the question about the computational function served by activation. What it says is that given some behavior (retrieval) we can infer that LTM is in some underlying state called activation, but all we know about that state of activation is that it is something that caused the behavior that we used to infer that activation in the first place. We're no wiser about the computational role of activation. The absence of any clear computational definition of activation is apparent in Cowan's concluding sentence: "The exact meaning of activation and of the two alternatives may change as the pursuit to test them continues; changing definitions is a legitimate part of the progression of a science". (p. 65). In other words, we can always use the term activation, because we can always change what it means.

The FoA and pointers to aLTM

Cowan writes: "The information held with the FoA could be described as a structured set of pointers, .. it would also serve as a portal to LTM learning. For example, to learn the list of digits *739482*, the individual might memorize *739*, then *48*, and then the association between these segments as *739-48*, subsequently incorporating the last digit to encode *739-48-2*. That reiterative process (see Rhodes & Cowan, 2018) would presumably be available for immediate recall" (Cowan, 2019, p.20). I take this to mean that pointers do the job of representing sequences. This gives the FoA all of the computational power needed by an STM system, but Cowan still declines to call it STM. Given that aLTM does not have a representation of *979482* to begin with, the focus of attention must be focusing on something other than a subset of aLTM. The only other thing available is the representation stored in an STM system.

Norris (2017) discussed the issue of how best to label a system that relies on pointers: "If there is a system where the STS contains pointers to LTM, should we really call this a STS, or is it just a pointer system? My own inclination is to stick with the term STS, as the pointers are doing all of the hard work." (p. 1003). But Cowan goes beyond having a simple set of pointers and proposes that “a pointer system is expected in which a structured set of references to information in aLTM would be established” and where “a set of items is apprehended with the FoA and then off-loaded into new LTM representations” (p. 52). Here, the hard work in not being done by pointers, but by a system labelled FoA which can construct structured representations and offload them to LTM. FoA has been allowed to subsume all of the processes normally considered to be part of a separate STM system. My preference remains to call that an STM system.

The role of rapid learning

Cowan proposes that some of the problems with aLTM can be overcome by invoking rapid learning and assumes “that information can be learned quite quickly, so newly learned structures (such as the serial positions of list items, spatial positions of array items, or binding of items to semantic roles) is processed by the FoA and is concurrently learned, resulting in new aLTM material that can be used on the trial (though learning may be imperfect and later retrieval depends on interference and on retrieval cues).” (p. 9). However, increasing the speed of learning does not help aLTM escape its predicament. Rapid learning relies entirely on representations constructed by the FoA, but these must be different from the representations in aLTM, otherwise there would be no need for rapid learning. The FoA has now been given all of the power and storage capabilities of a separate STM system. If you can rapidly learn *979482* you have already managed to solve the problem of two (there are two 9s). Both Cowan and I agree that this cannot be done with aLTM alone.

A model of STM as the front end of LTM

When Cowan suggests that a model of STM might form the front end of LTM learning it is not clear whether he has a particular model of STM in mind. It is also unclear what STM can do that is beyond the capabilities of the newly endowed FoA. The model he devotes most space to discussing is Burgess and Hitch (2006). Like all connectionist models of STM, their model has multiple components with separate interacting layers of nodes. The layers perform the task of representing the specific sequence of items or events and transferring those temporarily constructed representations into LTM. Much the same happens in the model of Page and Norris (2009). As with all computational models of STM, there is a lot of weighty structure here. Cowan wants us to accept that all of this mechanism can simply be reclassified as aLTM or FoA.

The 11 arguments for a separate copy of information in STM

Cowan's Table 1 presents 11 "*Arguments for a separate copy of information in STM*" which he attributes to me. Table 1 here presents my responses to Cowan's responses, along with his description of the original argument. I have ignored arguments 5-7 as I did not make them.

Table 1. Responses to the Arguments for a Separate Copy of Information in STM
See Cowan (2019) for a listing of the references cited in Cowan's response

Description of Argument	**Cowan's Response Against Separate Copy**	**My Response to Cowan**
1. Storage of new configurations is needed in STM	Everyone recognizes there must be new, rapid learning of information in STM tasks (e.g., Keppel & Underwood, 1962), and the newly-learned information is typically still in an activated state, aLTM, at the time of test (Cowan, 1999).	Few could disagree with the first part of this response, but it fails to address the question posed. The original question concerned the need to store novel representations that had no pre-existing representation in LTM. This cannot be achieved just by assuming that the learning is rapid. I also pointed out that there must by continual long-term learning. On first encounter with some

		new event there must be some long-term learning, otherwise every encounter would be the same as the first, and learning would never get underway.
2. Token representations cannot be represented in aLTM, only types	aLTM includes rapid learning of information, and therefore can include the same episodic information about tokens that one adds to LTM (Cowan, 1999; Nairne & Neath, 2001)	The case against aLTM applies regardless of the speed of aLTM. It needs more than go-faster stripes - it simply does not have the necessary representational capacity to do the job. Adding that extra capacity turns it into an STM system. What we need to know is how rapid learning works and exactly how it is supposed to solve the problem.
3. No extant model of STM performance based on aLTM	Including new learning as part of aLTM changes the need, because separate-STM-copy theories might be reclassified as the front end of long-term learning. Many long-term learning models exist. A few models deal explicitly with aspects of aLTM and new learning (Anderson & Matessa, 1997; Cowan et al., 2012).	The need is as great as ever. There are no computational models of STM performance based simply on activated LTM. The models cited are not models of aLTM, and the models in Cowan et al. (2012) do not simulate any of the benchmark phenomena of STM. To resort to reclassifying models of STM as part of aLTM is to admit defeat.
4. STM recall differs from LTM recall in its properties	There is evidence that long-term learning with repetition heavily relies on item-item associations (Zaromb et al., 2006), not just item-position as implied by Cumming et al. LTM with reduced interference looks more similar to STM (Dewar et al., 2010; Ecker et al., 2015a, 2015b). Unlike the usual procedures, STM can use semantic information (Potter, 1993) and LTM can be made to use phonological cues when such cues are best suited to the encoding context (Morris et al., 1977). Order retention suffers in dyslexia	Cumming, Page, and Norris (2003) was not cited in Norris (2017) and it is not clear how item-item versus item-position associations has any bearing on the issue. I did point out that that phonological confusions in STM only occur at short retention intervals, after which confusions are likely to be semantic

	within both STM and LTM (Martinez Perez et al., 2013; Szmalec et al., 2011).	
8. Variable binding must be encoded into STM	Patients with hippocampal damage and LTM deficiency also show a deficit in variable binding, in sentence comprehension requiring variable binding for pronoun assignment (Kurczek et al., 2013)	The argument was that we must have some way of performing variable binding. aLTM fails to offer an account of how these computations might be performed. Given Cowan's reluctance to accept the standard interpretation of neuropsychological evidence for a separation between STM and LTM, it is surprising to find him placing such weight on the neuropsychological evidence from a single study. In their abstract Kurczek et al. (2013) say "This finding suggests that the hippocampus plays a role in maintaining and integrating information even over a very short discourse history". Even if the conclusion were that the hippocampus, and only the hippocampus, plays a role in binding, any further conclusion about the role of aLTM depends on the additional assumption that the hippocampus is exclusively involved in LTM and could not be construed as implementing any part of a separate STM process.
9. Neuropathological deficits distinguish STM from LTM	Specific deficits in STM performance could come from deficient processes specific to STM maintenance (e.g., rehearsal: Cowan, 1988; or other kinds of deficient coding: Cermak, 1997; Morey, 2018; Morey et al., in press; Ruchkin et al., 2003). Also, LTM procedure used have not closely matched STM procedures used.	It is always possibly to attribute damage to stores to damage to processes. One need only claim that there is one process for reading out information in the short term and one for the long term. The neuropsychological evidence has recently been the subject of a special issue of the journal Cortex (Shallice & Papagno, 2019). In particular, see Logie (2019) for a critique of Morey, Rhodes, and Cowan (2019). Interestingly, the main theme in that issue is not whether STM and LTM are separate – that was largely taken for granted. The papers focus on presenting evidence for further fractionation of STM and working memory into separate buffers.

10. Tasks are impure measures of either STM or LTM	LTM learning may make use of use the focus of attention once for subspan lists but reiteratively for supraspan lists (Rhodes & Cowan, 2018), and the reiterative process could be impaired.	The response doesn't speak to the argument. Given that tasks are impure measures (Atkinson & Shiffrin, 1968), it is hard to design tasks that involve only STM or only LTM. That is, this is a statement about what follows from the assumption of separate stores.
11. Neuroimaging as a correlation fallacy	The scientific method seeks the most parsimonious and adequate theory that can accommodate all of the evidence, including correlations and causation. The neuroscientific evidence for the embedded-processes approach includes correlational neuroimaging-behavior correspondences (e.g., Chein & Fiez, 2010; Cowan, 2011; Cowan et al., 2011; Kalm & Norris, 2017; Lewis-Peacock et al., 2012; Li et al., 2014; Majerus et al., 2016; Öztekin et al., 2008) and causal TMS evidence (Postle et al., 2006; Rose et al., 2016).	Given that there are no pure measures neuroimaging data that implicates brain regions purported to be involved in LTM, in STM tasks, is simply correlational and is to be expected from the two-store view. Such data should therefore not be taken as evidence that regions assumed to be responsible for LTM are performing the STM task. The scientific method does indeed seek the most parsimonious and adequate theory. However, aLTM is not formulated with sufficient precision to know that it can accommodate the evidence. The appropriate metric of parsimony is not simply a count of the number of stores that a theory claims to have. We also have to count the number of ad hoc assumption too. By adding extra assumption and an extra STM model, the aLTM seems far from parsimonious. It has the potential to explain almost anything. Embedded memory systems will be subject to the same computational constraints as any other STM system. Calling them aLTM (see Figure 1) is simply another exercise in re-labelling.

Conclusion

The conclusion of this response remains the same that of Norris (2017): “A simple activation process would be unable to solve the problem of two, or to store novel representations. It follows that any model that places an emphasis on storage by activated

LTM must be supplemented by some additional mechanism that can represent multiple tokens [and] serial order. That mechanism must be able to perform the variable-binding operation required to construct novel representations. That additional mechanism would then amount to what has conventionally been thought of as a short-term store. In fact, the resulting model would look very much like existing computational models of STM. Some may still prefer to describe this by saying that STM is activated LTM. If they make it clear that there must be some additional mechanism, and explain how that mechanism operates, at least we'll know what they mean."

Cowan admits that there must be some additional mechanism but, with only a verbal description to go on, it is far from clear what he means or even whether his proposals would actually work. It seems very likely that if his proposals were incorporated into an explicit computational model they would work only to the extent that they instantiated the mechanism of some existing model of STM. This is apparent in the claim that "separate-STM-copy theories might be reclassified as the front end of long-term learning". In other words, you need a separate STM system. Short-term memory and long-term memory are still different, unless you pretend otherwise.

References

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 2, pp. 89-195).

Burgess, N., & Hitch, G. J. (2006). A revised model of short-term memory and long-term learning of verbal sequences. *Journal of Memory and Language, 55*, 627-652. doi:10.1016/j.jml.2006.08.005

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin, 104*, 163-191. doi:10.1037/0033-2909.104.2.163

Cowan, N. (1999). An embedded-processes model of working memory. In A. Miyake & P. Shah (Eds.), *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control.* (pp. 62-101). Cambridge, U.K.: Cambridge University Press.

Cowan, N. (2019). Short-term Memory Based on Activated Long-term Memory: A Review in Response to Norris (2017). *Psychological Bulletin*.

Cowan, N., & Chen, Z. (2008). How chunks form in long-term memory and affect short-term memory limits. In M. Page & A. Thorn (Eds.), *Interactions between short-term and long-term memory in the verbal domain* (pp. 86-107). Hove, East Sussex, UK: Psychology Press.

Cumming, N., Page, M., & Norris, D. (2003). Testing a positional model of the Hebb effect. *Memory, 11*(1), 43-63.

Hurlstone, M. J., Hitch, G. J., & Baddeley, A. D. (2014). Memory for serial order across domains: An overview of the literature and directions for future research. *Psychological Bulletin, 140*(2), 339.

Logie, R. H. (2019). Converging sources of evidence and theory integration in working memory: A commentary on Morey, Rhodes, and Cowan (2019). *Cortex, 112*, 162-171.

Morey, C. C., Rhodes, S., & Cowan, N. (2019). Sensory-motor integration and brain lesions: Progress toward explaining domain-specific phenomena within domain-general working memory. *Cortex, 112*, 149-161.

Norris, D. (2017). Short-term memory and long-term memory are still different. *Psychological Bulletin, 143*(9), 992-1009.

Page, M., & Norris, D. (2009). A model linking immediate serial recall, the Hebb repetition effect and the learning of phonological word forms. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 364*, 3737-3753. doi:10.1098/rstb.2009.0173

Shallice, T., & Papagno, C. (Eds.). (2019). *Impairments of short-term memory buffers: do they exist? [Special issue]. Cortex* (Vol. 112).