

Processing capacity limits are not explained by storage limits

Graeme S. Halford
University of Queensland

Steven Phillips
Electrotechnical Laboratory

William H. Wilson
University of New South Wales

Abstract. Cowan's review shows that a short-term memory limit of four items is consistent with a wide range of phenomena in the field. However he does not explain that limit, whereas an existing theory does offer an explanation for capacity limitations. Furthermore processing capacity limits cannot be reduced to storage limits as Cowan claims.

In his excellent review, Cowan concludes that short term memory storage is limited to four items, and he notes this corresponds to the limit in processing capacity defined by Halford, Wilson & Phillips (**BBS**, 1998). Furthermore his conclusion that the limit is in the number of integrated objects, independent of the complexity of each, agrees well with the observation of HW&P (1998) that humans are limited to relating four entities, irrespective of their complexity. However these correspondences do not imply that processing limits can be subsumed under storage limits, as Cowan claims.

The fact that the size of the limit is 4 in both cases is not a strong argument for identification because, given that the limit is small, the same number could occur in both contexts by coincidence. Alternatively, storage and processing systems could be distinct but with equal capacities to facilitate transfer from one to the other. There are a number of reasons why processing cannot be subsumed under storage. To take a straightforward example, there clearly is a difference between simply holding the numbers 7 and 4 in short term store, and adding them to yield the sum, 11. In general storage, in the sense of internal representation, is a prerequisite for processing, but cognitive processing cannot be reduced to storage. Furthermore higher cognitive processes require representations that have properties beyond those required for storage, including omnidirectional access and analogical mapping (HW&P, 1998).

Cowan's position is that a concurrent short term memory load can be held in the activated portion of long term memory while other information is being processed in the focus of attention. Lack of interference between processing and short term storage is explained because the focus of attention can be devoted to either storage or processing, but need not be devoted to both at once. However this still implies that storage and processing are distinct, and also implies there would be no tradeoff between the two. It is not fundamentally different from the position of HW&P (1998).

Cowan offers no explanation for the limit he observes in storage capacity, whereas HW&P (1998) offer a natural explanation for processing capacity limits. In this model, conceptual complexity is defined by the arity, or number of arguments that can be bound into a single relation. Human adults are typically limited to processing one quaternary relation in parallel. Each component of the relation is represented by a vector, and the binding is represented by the tensor product of the vectors. Thus the binary relational instance $\text{larger}(\text{elephant}, \text{mouse})$ is represented by $v_{\text{larger}} \otimes v_{\text{elephant}} \otimes v_{\text{mouse}}$. The rank of the tensor product is one more than the arity of the relation. The more complex relations are represented by tensor products of higher rank, the greater complexity of which explains why more complex relations are associated with higher processing load. However the size of the component vectors has much less effect on processing load, so the fact that the

limit is not related to the size of the entities is also explained. Thus, in terms of our relational model, there is a limit on tensor rank entailed by the rapid growth of the number of tensor units as rank increases. Given that a short term memory store of capacity 4 is connected to a tensor-like system for processing, the limit of 4 on store size is a consequence of the fact that for most cognitive tasks, processing of the objects in the store is a necessity.

The links between storage and processing phenomena are worth exploring. In Section 2, Cowan argues that the unity of conscious awareness implies the contents of attended channels should be integrated or combined. Similarly, category clusters (discussed in 2.7 and 3.4.2) imply a link between instances of the category. Cowan further contends, in 3.1.3, that the short term storage limit is observed only with items recalled in correct serial positions. Given that the slots of a relation are identified, serial position can be coded as a relation ordered-items(item1,item2,item3,item4). The observation of no limit with free recall would then suggest that it is ability to represent the relation, rather than the items, that is subject to the limit. This would appear to be consistent with the relational complexity theory of HW&P (1998). Furthermore, it clearly points to explaining storage limits in terms of complexity of relations that can be represented. This would also explain the finding of Nairne (1991, referred to by Cowan in 3.4.3) that errors occur up to three positions from the correct position. The reason would be that the items are represented as a quaternary relation, which contains only four slots. The further finding, in 3.4.5 that participants could predict the 7th item from items 3, 4, 6 may also indicate that the task is represented as a quaternary relation.

These phenomena indicate links between entities that are important, but the nature of these links is not really clear, and the issue is clouded by the lack of a well specified theory in Cowan's paper. Some properties of relational knowledge defined by HW&P (1998) seem to be involved in the phenomena discussed above, but it is not clear that they all are. We could define the relational instances fruit(apple, banana, orange, pear) and fruit(lychee, pineapple, passionfruit, guava), etc. Organizing memory storage as quaternary relations in this way would account for recall of items in clusters of four. However it would also predict a lot of other properties of relational knowledge that Cowan has not demonstrated. For example, relational knowledge has the property of omni-directional access (HW&P, 1998) which means that, given any $n-1$ components of a relational instance, the remaining component can be retrieved. Thus, given the quaternary relation proportion(4,2,?,3) we can determine that the missing component must be "6" because it is necessary to complete the proportion $4/2 = 6/3$. However it is far from clear that category clusters share this property. If given a list [apple, banana, ?, pear] there is no particular reason why we should recall "orange". Thus category clusters do not entail the kind of constraints that are entailed in relations. Another property of relational knowledge is that analogical mappings can be formed between corresponding relational instances (Holyoak & Thagard, 1995). Again, it is not clear that analogies can be formed between category clusters.

Storage is not a simple, unitary matter, but can take many forms. Furthermore, the form in which information is stored affects the form in which it is processed. Some of the possibilities, together with possible implementation in neural nets, are:

Item storage – implemented as a vector v_i of activation values over a set of neural units.

Associative links between items, implemented as connection weights between units in different vectors.

Superposition of items – implemented as summation of item vectors. This is tantamount to a prototype.

Superimposed items bound to a category label, such as fruit(apple) + fruit(banana) + fruit(orange) + fruit(pear). This is equivalent to a unary relation and can be represented by a Rank 2 tensor

$$v_{\text{fruit}} \otimes v_{\text{apple}} + v_{\text{fruit}} \otimes v_{\text{orange}} + v_{\text{fruit}} \otimes v_{\text{orange}} + v_{\text{fruit}} \otimes v_{\text{pear}}$$

Item-position bindings: ordered-fruit(first,apple) + ordered-fruit (second,orange) + , . . , + ordered-fruit(fourth,pear)}. This is a binary relational instance and can be implemented by the tensor product

$$V_{\text{ordered-fruit}} \otimes V_{\text{first}} \otimes V_{\text{apple}} + V_{\text{ordered-fruit}} \otimes V_{\text{second}} \otimes V_{\text{orange}} + \dots + V_{\text{ordered-fruit}} \otimes V_{\text{fourth}} \otimes V_{\text{pear}}$$

Binding items into n-ary relations where n has a maximum value 4. This can be implemented by a tensor up to Rank 5:

$$V_{\text{fruit}} \otimes V_{\text{apple}} \otimes V_{\text{orange}} \otimes \dots \otimes V_{\text{pear}}$$

These representations have different characteristics. They permit different retrieval operations, and impose different processing loads. Importantly, at least some of their properties can be captured by neural net models. The Rank n tensor would explain why processing load increases with the number of entities related, and consequently suggests why the capacity limit tends to be low. However the earlier representations are not sensitive to processing load in this way. It should be clear from these examples that storage and process are intimately related, and that a theory of capacity must include both aspects of computation. However, while their interaction may be complex, it is not arbitrary. Our theory specifies a unique set of properties for processes involving relations of different arities.

Conclusion

Cowan has done the field a great service by showing that a broad range of observations is consistent with the limit of four entities that had been proposed previously by HW&P (BBS, 1998). However his claim to reduce processing capacity to storage capacity is not substantiated. Furthermore he offers no explanation for the limit, and glosses over the fact that at least one existing theory offers a potential explanation as to why the limit should be small.

References

Halford, G.S., Wilson, W.H. & Phillips, W. (1998) Processing capacity defined by relational complexity: Implications for comparative, developmental and cognitive psychology. *Behavioral Brain Sciences*, 21(6), 803-831.

Holyoak, K.J. & Thagard, P. (1995) *Mental leaps: Analogy in creative thought*. Cambridge, MA: Bradford.