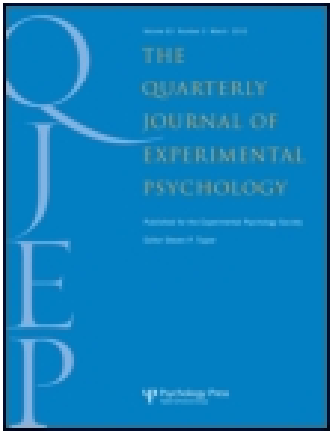

The Quarterly Journal of Experimental Psychology
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/pqje20

The phonological loop: Some answers and some questions

Alan D. Baddeley [a] & Janet D. Larsen [b]
[a] University of York , York, UK
[b] John Carroll University, University Heights , OH, USA
Published online: 13 Apr 2007.

THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY
2007, 60 (4), 512–518

Psychology Press
Taylor & Francis Group

Rejoinder

The phonological loop: Some answers and some questions

Alan D. Baddeley
University of York, York, UK

Janet D. Larsen
John Carroll University, University Heights, OH, USA

Jones, Hughes, and Macken (2007) claim that their data and our own are inconsistent with a multicomponent working-memory model. We explain in greater detail how the model can account for the data and can address their more specific criticisms. Both sides accept that data relating to the presence of a phonological similarity effect throughout the list depend on list length. We accept that, at this point, all explanations of their interaction are speculative and require further empirical investigation. We examine J, H, & M's interpretation of their and our results in terms of an auditory modality effect, observing that their interpretation of this effect is not well supported by the literature. We suggest that their account assumes a very narrow basis for a general theory of short-term retention, in contrast to a phonological loop interpretation, which forms part of a well-developed and articulated model of working memory.

Jones, Hughes, and Macken (2007) (henceforth J, H, & M) have proposed a radical revision of the way in which short-term verbal memory is conceptualized, based initially on the observation that a critical interaction between input modality, articulatory suppression, and phonological similarity was principally confined to the recency portion of the serial position curve and backed up by their subsequent studies incorporating prefix and suffix items. In responding to their earlier studies, our main aim was to point out that our own data were inconsistent with the suggestion that the critical interaction was principally based on recency, suggesting caution before strong conclusions were drawn from the existing evidence.

We are pleased to note that in their response to our comments, Jones et al. do not dispute the data indicating that the association with recency does not apply to shorter lists. Nor do they deny that this is one of many examples whereby evidence of verbal/phonological coding is robust for relatively short lists, but tends to disappear with longer lists, particularly when an additional load is imposed, for example by articulatory suppression

Correspondence should be addressed to Alan D. Baddeley, Department of Psychology, University of York, Heslington, York YO10 5DD, UK. E-mail: ab50@york.ac.uk

We are grateful to Susan Gathercole and Graham Hitch for comments on an earlier draught. The support of MRC Grant G9423916 to A.B. is gratefully acknowledged.

http://www.psypress.com/qjep DOI:10.1080/17470210601147663

(e.g., Larsen & Baddeley, 2003), irrelevant speech (Hanley & Broadbent, 1987; Salame & Baddeley, 1986), or reading disability (Johnston, Rugg, & Scott, 1987). They do, however, take issue with us on a number of points. First, they question whether the phonological loop model can handle our own data. They then challenge our speculations regarding, first, the possible reasons for the interaction between list length and the phonological similarity effect and, second, our interpretation of prefix and suffix effects within the working-memory model. We consider these in turn.

Are our data inconsistent with our model?

A number of features of our interpretation concern J, H, & M, some reflecting our use of a procedure involving a delay; this stems from the fact that the data were part of an extended study following up earlier research using this technique (Larsen & Baddeley, 2003). J, H, & M very reasonably point out that, given an interval filled with articulatory suppression, one would not expect the phonological store to be contributing directly to performance. That is certainly the case, if the loop were regarded as functioning in isolation. However, the whole concept of working memory is one of a multicomponent system, whereby a number of storage processes are used to optimize performance. The model itself involves three storage components—namely, the phonological loop, the visuo-spatial sketchpad, and the episodic buffer. It is also assumed to have access to information from outside the system and hence to be able to benefit from both long-term memory and sensory memory systems.

J, H, & M describe the phonological loop as the keystone of working memory, the implication being that without it the whole system would collapse. As the evidence with patients with impairments to the phonological loop indicates, this is far from the case (Vallar & Shallice, 1990); of much greater functional importance are the central executive and episodic buffer components. The loop system is, however, capable of storing small amounts of sequential information accurately and with relatively little executive demand, allowing it to play an important role in the acquisition of both first and later languages (Baddeley, Gathercole, & Papagno, 1998) and to provide a simple but effective mechanism for controlling behaviour (Baddeley, Chincotta, & Adlam, 2001; Luria, 1959a, 1959b; Vygotsky, 1962). The phonological loop is probably the most extensively investigated component in the fields of cognitive psychology, neuropsychology, and cognitive development. It does, however, have a relatively small capacity, just a few digits, unless supported by concurrent rehearsal and by other components of working memory, most notably the episodic buffer (Baddeley, 2000).

Let us consider then the task of retaining a sequence of six letters over a delay. In the absence of concurrent suppression, the auditorily presented items will be stored in the phonological loop and recycled using the articulatory system. In the case of visual presentation, a further stage will be involved whereby the visual stimuli will be recoded so as to allow registration in the phonological store. A number of sources of evidence discussed by Baddeley (2000) indicate that this is accompanied by registration in a further multidimensional store, the episodic buffer, which is capable of taking advantage, not only of visual and phonological codes, but also of any other sources of constraints provided, for example, by syntax and semantics in the case of sentence memory.

Consider next the situation in which articulatory suppression is required throughout this operation. In the case of the visually presented material, this will proceed as before, except for the absence of a contribution from phonological recoding of the visual stimuli and, of course, subsequent articulatory rehearsal. As we know both from neuropsychological evidence (Vallar & Shallice, 1990) and from earlier studies involving suppression and visual presentation (e.g., Larsen & Baddeley, 2003), subjects can still store sequential information in the absence of these processes, although there is a clear impairment in level of performance. In the case of auditory presentation, we would expect the phonological loop to hold incoming material for a matter of seconds, probably longer than is the case with visual material, resulting in a phonologically coded trace within the multimodal episodic buffer. This would have two advantages over visual presentation—namely, the extended exposure time due to the storage capacity of the phonological store, together with a phonological code, which, as we noted earlier, is well suited to maintaining sequential information. In short, the loop will influence performance, but it will do so indirectly through its impact on the episodic buffer, which in turn raises the issue of rehearsal.

The question of rehearsal is important because the episodic buffer is, of course, itself required to maintain information, if not by articulation, then how? As explained elsewhere (Baddeley, in press), we suspect that subvocal articulation is a somewhat atypical form of rehearsal, possible only because the phonological material retained in most studies can be mapped directly onto a familiar spoken response such as a digit. It seems likely that rehearsal in visual, semantic, and other systems may reflect a more general process of attentional activation and reactivation. There is good evidence to suggest that such attentionally based rehearsal is available for verbal material as well as for visual and semantic information. For example, articulatory suppression is not sufficient to cause the forgetting of consonant trigrams in the classic Peterson task in healthy subjects (Baddeley, Lewis, & Vallar, 1984), whereas such suppression is effective in Alzheimer's disease patients with defective executive processes (Morris, 1984, 1986). Forgetting does of course occur when suppression is replaced by a demanding backward counting task, suggesting that a crucial and general feature of rehearsal is attention. We suggest, therefore, that attentional rehearsal, although not subvocal rehearsal, is possible during the filled delay in our experiments.

A related point raised by J, H, & M concerns *semantic encoding* and whether this should always give rise to better recall, since it was found to be advantageous in a study by Hanley and Bakopoulou (2003). Our response here is that the relative advantage of different coding modes will depend on the material, the conditions, and the strategy selected. When meaningful words are used, and pairs are made up that allow the two words to be readily combined into a meaningful unit, then semantic coding will occur and be beneficial, whereas this tends not to be the case when arbitrary pairings are used at relatively rapid presentation rates (Baddeley & Levy, 1971; Jeffries, Lambon Ralph, & Baddeley, 2004). Repeated presentation will also facilitate the use of semantic coding (Baddeley, 1966). There is, furthermore, evidence that under appropriate conditions, it is possible to combine phonological and semantic coding, with subjects apparently depending on phonological coding after short delays and semantic coding after long delays (Baddeley & Ecob, 1970). In short, the phonological loop is simply one relatively small part of a highly flexible interactive system that is operated strategically. Under some circumstances involving the rapid presentation of arbitrary items, it is likely to be a more effective strategy than semantic coding. This leads onto the next issue.

Why does phonological similarity have less impact on longer lists?

As mentioned earlier, there appears to be considerable evidence for this phenomenon, but little in the way of direct investigation. We ourselves are certainly guilty in this respect, relying too heavily on speculation as to the possible role of strategy. Given a limited capacity, the phonological store is likely to become overloaded. Will that lead to a simple loss of the usability of that code, or will the code still be useable but be abandoned in favour of other codes that may be more capable of maintaining performance as list length increases?

The problem in tackling this question is not that the various options are logically incompatible, as suggested by J, H, & M, but rather that of failure to develop adequate methods to determine strategy use that are independent of the effects being studied. This is a difficult, but by no means impossible, task. Ways ahead are likely to include explicit strategic instructions, which Hanley and Bakopoulou (2003) have found to be effective, probably coupled in due course with neuroimaging evidence such as that presented in a recent paper by Petersson, Gisselgard, Gretzer, and Ingvar (in press) who performed a functional connectivity analysis of a positron emission tomography (PET) study of the influence of irrelevant speech on the serial recall of digits. They found that the pattern of connectivity changed under irrelevant speech from the strongly linked left temporo-parietal and frontal activation that typifies verbal short-term memory (STM) under baseline conditions, to a link between the temporo-parietal and the medial temporal area typically involved in episodic long-term memory. They suggest that this may reflect a switch from reliance on the phonological loop to a greater involvement of the episodic buffer.

J, H, & M account for the length-similarity interaction in terms of a perceptual interpretation of the influence of phonological similarity on recall. Like us, they assume both a strategy and an automatic order-encoding system that participants "exploit—to recover the order of auditory to-be-remembered items", going on to suggest that "this strategy is only likely to be effective (or at least to be more effective) for items at list boundaries". As both their data sets and ours suggest, however, this does not seem to be the case for shorter lists. Similarity is assumed to influence, performance "when the sequence exhibits acoustic changes on a common ground". The arbitrary assumption is then made that acoustically similar sequences fail to meet this requirement, whereas dissimilar sequences do, though presumably only with short lists.

This assumption of the importance of acoustic fluctuation also characterizes the changing-state hypothesis, whereby Jones (1993) explains the irrelevant speech effect, although in this context, phonologically similar sequences are assumed to fluctuate sufficiently to fulfil the changing-state criterion just as effectively as do dissimilar sequences (Jones, Beaman, & Macken, 1996). In short, the explanation of the interaction of list length with phonological similarity offered by J, H, & M is just as arbitrary as our strategy interpretation. We need more evidence here, rather than more speculation.

Reliance by J, H, & M on a perceptual interpretation of STM requires them to explain the occurrence of the full range of phonological coding effects, similarity, word length, and irrelevant speech that occur with visual presentation when articulatory suppression is avoided. Our own assumption is that all of these effects involve a single system employing a common postperceptual phonological store. Evidence that it is postperceptual in nature comes from patients whose speech perception is normal, but whose phonological STM is grossly impaired, while other patients show exactly the opposite pattern of good memory and impaired perception, providing a clear dissociation between perception and memory (Baddeley & Wilson, 1993; Vallar, 2006; Vallar & Shallice, 1990).

The assumption of a postperceptual phonological store also provides a good account of both behavioural and neuroimaging studies of sign language, where comparable phenomena to those observed in auditory memory are found in an equivalent anatomical location, together with evidence for an additional visual input-based location (Rönnberg & Rudner, in press; Rönnberg, Rudner, & Ingvar, 2004). J, H, & M deal with this wide range of evidence by proposing an unspecified "gestural" system, which presumably corresponds to their postulated perceptual system. We would like to see more evidence for such a far-reaching assumption of equivalence between auditory perception and motor behaviour.

Prefix and suffix effects

The working-memory model does not currently include a component for either visual or auditory sensory memory systems, treating them as part of the relevant perceptual systems rather than as components of memory. Such sensory memory systems are of course still assumed to be capable of influencing overall performance, but have not been included in the multicomponent model because there does not at present appear to be a generally accepted coherent account of how they operate. Despite many years of investigation and multiple attempts at theorization, there still seems to be little agreement on how to explain suffix effects. The balance of opinion, however, appears to be in favour of our own assumption of a two-component explanation only one of which is modality based, rather than the unitary perceptual view advocated by J, H, & M. A recent review of the literature by Bloom (2006) identified only 1 paper supporting a single-component theory (Bloom & Watkins, 1999) compared to 34 supporting a two-component theory. Although Bloom himself (2006) advocates a unitary view, largely on the grounds of detecting a minor semantic modulation of the suffix effect, he explicitly rejects a changing-state hypothesis of the type that J, H, & M propose.

Jones (1993) aims, very appropriately, to relate his theories to concepts from psychoacoustics, basing his own theorizing on the perception of "auditory objects". Unfortunately, however, this view is rather less securely based than might at first appear. A recent review of the concept of an auditory object by Griffiths and Warren (2004) observes that "it is not clear what properties these should possess, how they might be represented in the brain, or how they might be related to the more familiar objects of vision", concluding that "the concept of an auditory object challenges our understanding of object perception" (Griffiths & Warren, 2004, p. 887). That does not, of course, imply that the concept is invalid, but it does raise doubts as to whether it provides a firm foundation for a theory of memory.

It is certainly the case that prefix and suffix effects are not currently included in our working-memory model—that does not, however, imply that such effects will not influence performance. Consider, for example, the visual subsystem that Neisser (1967) termed iconic memory. This initially appeared to provide a highly coherent model of STM (Sperling, 1960). However, as the model developed, it became clear that iconic memory did not constitute a unitary memory system; at least one component appears to operate at a retinal level, being disruptable by a flash of light, whereas others operated at a later, pattern-based, stage, at a point beyond which the information from the two eyes is combined (Turvey, 1973). Neither of these processes would be likely to provide a suitable model for the visuo-spatial sketchpad, which can handle auditory information regarding visuo-spatial targets almost as effectively as visually presented items (Baddeley, in press; Logie, 1995). That does not of course mean that performance on a task involving the sketchpad could not be influenced by iconic memory, given that visual presentation is used. This could either be positive, as with the observation of Phillips (1974) that recognition memory for patterns is substantially greater when a complex pattern is presented in the exactly the same location, or negative as would occur if a complex pattern is subsequently masked (Sperling, 1960). We are therefore happy to accept that prefix and suffix effects will impair performance on immediate serial verbal recall tasks, but do not regard them as a suitable model for the rest of the system.

CONCLUSION

Despite our apparently unbridgeable differences, we suggest that our views and those of J, H, & M are in some respects relatively similar. We both accept that there is a modality contribution to auditorily presented STM tasks. We both accept the role of strategy and a rehearsal process that plays an important part in retention, although this is relabelled as a "gestural" component by J, H, & M rather than as verbal rehearsal, and we both accept the need for a subsequent multidimensional coding component, termed the object-oriented episodic record by Jones (1993) and the episodic buffer by Baddeley (2000), although Jones (personal communication, August 10, 2006) does not regard these as equivalent.

We differ in that J, H, & M deny the need for short-term phonological storage. This allows them to simplify their system by omitting one component. However, in doing so they deny a model, the phonological loop, that can be computationally specified and is backed up by experimental, neuropsychological, neuroimaging, and developmental evidence. They propose to replace it with a perceptual system coupled with an unspecified gestural rehearsal system that is left to perform the remaining functions currently attributed to working memory. We suggest that their requiem for the phonological loop is somewhat premature.

Original manuscript received 14 November 2006
Accepted revision received 15 November 2006
First published online 12 February 2007

REFERENCES

Baddeley, A. D. (1966). The influence of acoustic and semantic similarity on long-term memory for word sequences. *Quarterly Journal of Experimental Psychology*, *18*, 302–309.

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, *4*, 417–423.

Baddeley, A. D. (in press). *Working memory, thought and action*. Oxford, UK: Oxford University Press.

Baddeley, A. D., Chincotta, D., & Adlam, A. (2001). Working memory and the control of action: Evidence from task switching. *Journal of Experimental Psychology: General*, *130*, 641–657.

Baddeley, A. D., & Ecob, J. R. (1970). Simultaneous acoustic and semantic coding in short-term memory. *Nature*, *277*, 288–289.

Baddeley, A. D., Gathercole, S. E., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, *105*, 158–173.

Baddeley, A. D., & Levy, B. A. (1971). Semantic coding and short-term memory. *Journal of Experimental Psychology*, *89*, 132–136.

Baddeley, A. D., Lewis, V. J., & Vallar, G. (1984). Exploring the articulatory loop. *Quarterly Journal of Experimental Psychology*, *36*, 233–252.

Baddeley, A. D., & Wilson, B. A. (1993). A case of word deafness with preserved span: Implications for the structure and function of short-term memory. *Cortex*, *29*, 741–748.

Bloom, L. C. (2006). Two-component theory of the suffix effect: Contrary evidence. *Memory, and Cognition*, *34*, 648–667.

Bloom, L. C., & Watkins, M. J. (1999). Two-component theory of the suffix effect: Contrary findings. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *25*, 1452–1474.

Griffiths, T. D., & Warren, J. D. (2004). What is an auditory object? *Nature Reviews Neuroscience*, *5*, 887–892.

Hanley, J. R., & Bakopoulou, E. (2003). Irrelevant speech, articulatory suppression and phonological similarity: A test of the phonological loop model and the feature model. *Psychonomic Bulletin & Review*, *10*, 435–444.

Hanley, J. R., & Broadbent, C. (1987). The effects of unattended speech on serial recall following auditory presentation. *Bristish Journal of Psychology*, *78*, 287–297.

Jefferies, E., Lambon Ralph, M. A., & Baddeley, A. D. (2004). Automatic and controlled processing in sentence recall: The role of long-term and working memory. *Journal of Memory and Language*, *51*, 623–643.

Johnston, R. S., Rugg, M. D., & Scott, T. (1987). Phonological similarity effects, memory span and developmental reading disorders: The nature of the relationship. *British Journal of Psychology*, *78*, 205–211.

Jones, D. M. (1993). Objects, streams and threads of auditory attention. In A. D. Baddeley & L. Weiskrantz (Eds.), *Attention: Selection, awareness and control* (pp. 87–104). Oxford, UK: Clarendon Press.

Jones, D. M., Beaman, P., & Macken, W. J. (1996). The object-orientated episodic record model. In S. Gathercole (Ed.), *Models of short-term memory* (pp. 209–238). Hove, UK: Psychology Press.

Jones, D. M., Hughes, R. W., & Macken, W. J. (2007). The phonological store abandoned. *Quarterly Journal of Experimental Psychology*, *60*, 505–511.

Larsen, J., & Baddeley, A. D. (2003). Disruption of verbal STM by irrelevent speech, articulatory suppression and manual tapping: Do they have a common source? *Quarterly Journal of Experimental Psychology*, *56A*, 1249–1268.

Logie, R. H. (1995). *Visuo-spatial working memory*. Hove, UK: Lawrence Erlbaum Associates Ltd.

Luria, A. R. (1959a). The directive function of speech in development and dissolution, Part I. *Word*, *15*, 341–352.

Luria, A. R. (1959b). The directive function of speech in development and dissolution, Part II. *Word*, *15*, 453–464.

Morris, R. G. (1984). Dementia and the functioning of the articulatory loop system. *Cognitive Neuropsychology*, *1*, 143–157.

Morris, R. G. (1986). Short-term forgetting in senile dementia of the Alzheimer's type. *Cognitive Neuropsychology*, *3*, 77–97.

Neisser, U. (1967). *Cognitive psychology*. New York: Appleton-Century Crofts.

Petersson, K. M., Gisselgard, J., Gretzer, M., & Ingvar, M. (in press). Interaction between a verbal working memory network and the medial temporal lobe. *NeuroImage*.

Phillips, W. A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception and Psychophysics*, *16*, 283–290.

Rönnberg, J., & Rudner, M. (in press). The cognitive neuroscience of signed language. In L. Bäckman & L. Nyberg (Eds.), *Memory, aging, and brain*. Hove, UK: Psychology Press.

Rönnberg, J., Rudner, M., & Ingvar, M. (2004). Neural correlates of working memory for sign language. *Cognitive Brain Research*, *20*, 165–182.

Salame, P., & Baddeley, A. D. (1986). Phonological factors in STM: Similarity and the unattended speech effect. *Bulletin of the Psychonomic Society*, *24*, 263–265.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied*, *74*, 1–29.

Turvey, M. T. (1973). On peripheral and central processes in vision: Inferences from an information processing analysis of masking with patterned stimuli. *Psychological Review*, *80*, 1–52.

Vallar, G. (2006). Memory systems: The case of phonological short-term memory. A festschrift for Cognitive Neuropsychology. *Cognitive Neuropsychology*, *23*, 135–155.

Vallar, G., & Shallice, T. (1990). *Neuropsychological impairments of short-term memory*. Cambridge, UK: Cambridge University Press.

Vygotsky, L. S. (1962). *Thought and language* (E. Hanfmann & G. Vakar, Trans.). Cambridge, MA: MIT Press.