Running head: PHONOLOGICAL LOOP AS A BUFFER STORE

The phonological loop as a buffer store: An update

Alan D. Baddeley

Graham J. Hitch

Department of Psychology

University of York

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

**Abstract**

We regard our multicomponent model of working memory as reflecting a hierarchy of buffer stores with buffer storage providing an effective way of combining information from two or more streams that may differ in either the speed of input or in the features coded. We illustrate this through the case of the phonological loop component of the model. We discuss its gradual development through a combination of evidence from mainstream cognition and neuropsychology with the need for more detailed modelling of issues such as the representation of serial order. A brief account follows of the application beyond the laboratory and clinic of the concept of a phonological loop and the methods designed to study it. We then discuss some criticisms of the overall multicomponent model, concluding with a discussion of the major contribution made by neuropsychological evidence to its development together with some suggestions as to comparative lack of influence from more recent studies based on neuro-imaging.

Keywords: Working memory, phonological loop, buffer store; STM; STM patients

Memory storage can be divided into two broad categories, archival, in which information is assumed to be potentially useable over long periods of time and more temporary buffer storage. Buffer stores serve an important role in allowing two or more streams of information to be coordinated, despite potential differences in speed, durability or format of the contributing streams. They occur in visual perception, where information is temporarily stored at the retinal level as shown by brightness masking which is essentially monocular, at the pattern processing level that occurs beyond the point at which information from the two eyes is combined and then at the level of visual short-term memory (Sperling, 1967; Turvey, 1973). In the field of memory per se, the proposal to distinguish between short-term memory (STM) and long-term memory (LTM) suggested that the existing concept of unitary memory system should be abandoned (Brown, 1958; Peterson & Peterson, 1959), leading to the proposal of STM as a buffer store (Broadbent, 1968). The concept was elaborated and extended by Atkinson and Shiffrin (1968) into their modal model in which an array of sensory buffer stores then feed into a short-term store (STS), a buffer that played a central role in cognitive processing. It was assumed to provide a crucial interface with LTM and also to serve as a limited capacity attentional control system to which they applied the term working memory. Although influential, the modal model itself lacked further development as the authors then moved away from the field of STM.

The account that follows concerns our own subsequent elaboration of the modal model into a multicomponent model of working memory and its fractionation into a number of inter-related buffer stores. We focus particularly on the impact on theoretical development of the concept of a phonological loop and the way in which the broad concept of verbal STM has gradually been extended and refined, often as a result of neuropsychological evidence. Our account will be broken into a number of stages, beginning with the case for distinguishing between LTM and STM, moving on to the question as to whether STM itself should be fractionated into two or more buffer stores. At this point focus will move to the acoustic/verbal domain with concern for what is stored and how, together with the potential evolutionary value of such storage. After considering the need for more precise computational modelling of the phonological loop, we conclude with a discussion of a fourth proposed component, the episodic buffer and the somewhat different role this plays as a more complex central buffer store.

**Distinction between STM and LTM**

The modal model assumed three types of memory storage, a set of perceptual buffers that operate in parallel, essentially forming part of the processes underpinning perception, before feeding into a second short-term store (STS). This served to hold information pending manipulation, while at the same time allowing gradual transfer to a long-term store (LTS). While the modal model initially seemed to give a good account of much of the existing data, it ran into two problems. One of these concerned LTM, where its assumption of a gradual time-based transfer of information from STS to LTS proved unfounded, a problem amplified by the demonstration by Craik and Lockhart (1972) that LTM was much more dependent on the manner in which material was processed than with processing duration, with relatively shallow encoding of the visual features of a word for example being much less effective than processing its meaning. A second problem concerned the assumption that the STS served as a working memory, not only feeding information into the LTS, but serving a much more general function in the overall control of behaviour. The problem came with the report by Shallice and Warrington (1970) of single cases of patients who appeared to have a very specific deficit in STS, with a digit span of two items or less. Such patients should, according to the modal model have a crucial bottle neck in transfer of information to LTM and if the STS serves as a working memory, be greatly handicapped in their everyday life. They and a range of similar cases described subsequently, appear to have normal LTM and, provided their deficit is relatively pure, to have few problems in day to day activities (Vallar & Shallice, 1990). Such patients show exactly the opposite pattern to amnesic patients with a relatively pure deficit in LTM, who typically have normal digit span together with markedly impaired LTM performance resulting in major problems in coping independently (Baddeley & Warrington, 1970; Milner, 1966; Wilson & Moffat, 1992). Such neuropsychological evidence thus supported the need to separate LTM and STM, but was inconsistent with the dominant Atkinson and Shiffrin (1968) model.

The neuropsychological evidence was paralleled by a series of studies focusing on the same issue but using healthy participants. These seemed at the time to suggest two further important differences between STM and LTM. One of these concerned the observation that a number of standard tasks appear to have separable long and short-term components. Probably the clearest of these was free recall in which a list of unrelated words was presented and participants invited to recall them in any order they wished. This typically gives a serial position curve with somewhat better recall of the initial items (the primacy effect) and substantially better recall of the last few (the recency effect). A delay of 10 seconds filled by counting to prevent rehearsal will remove the recency effect but have virtually no effect on other items, suggesting that recency reflects the output for some temporary STS (Glanzier & Cunitz, 1966). Consistent with this hypothesis was the fact that recency is well preserved in amnesic patients, despite greatly impaired performance on earlier items (Baddeley & Warrington, 1970), while the patients studied by Shallice and Warrington (1970) showed the opposite pattern.

A note of caution however was introduced by the discovery that under certain circumstances, clear evidence of long-term recency could be found both in carefully devised laboratory tasks (Bjork & Whitten, 1974) and in everyday situations such as remembering where you parked your car (Baddeley & Hitch, 1993; da Costa Pinto & Baddeley, 1991). A further problem emerged when Baddeley & Hitch (1974) showed that a concurrent task that would grossly impair performance on most STM tasks had no effect on recency in verbal free recall, again quite at odds with the modal model of recency as reflecting the contexts of a short-term buffer store.

A second source of evidence that seemed at first to be consistent with a simple distinction between STM and LTM came from the observation that whereas STM seems to rely heavily on an acoustic or phonological code and be relatively uninfluenced by factors such as semantic similarity (Baddeley, 1966a) the opposite is the case for LTM which appeared to be semantically-based (Baddeley, 1966b). Similar distinctions were reflected in other paradigms such as free recall in which phonological errors tended to reflect the recency component while semantic errors came principally from earlier, LTM-based items (Craik & Levy, 1970). A similar pattern was found in the immediate recall of prose passages in which surface factors such as syntactic details could be recovered from the most recent sentence in the passage, while earlier sentences appeared to reflect only semantic gist (Sachs, 1967).

Neuropsychological evidence that initially appeared to support this view of amnesia as a failure of semantic encoding came from a study of patients suffering from alcoholic Korsakoff syndrome who appeared to be unable to take advantage of semantic coding in verbal recall (Cermak, Butters & Moreines, 1974; Cermak & Reale, 1978). A study of densely amnesic patients by Baddeley and Warrington (1973) however did not show any signs of failure to process information semantically as reflected for example in their capacity to use semantically-based category clustering in verbal recall. It subsequently proved to be the case that the failure to use semantic coding constructively in the Korsakoff patients reflected additional executive problems resulting from subtle frontal lobe damage rather than representing a basic cause of amnesia (Cermak, 1976)

It thus proved to be the case that the simple assumption that STM was linked to phonological and LTM to semantic coding proved over simplified. Participants will use both types of coding in STM paradigms such as word span, provided the order of the words can be readily linked in a meaningful way, with the semantic coding indeed proving more durable (Baddeley & Ecob, 1970; Baddeley & Levy, 1971; Campoy & Baddeley, 2008). Furthermore, it is clearly the case that long-term phonological learning exists; without it we could never acquire vocabulary or learn a person’s name. By the early 1970s, the evidence that STM comprised a simple phonological buffer store was encountering major problems.

**Working Memory**

In an attempt to simplify the issue, we ourselves decided to test the general hypothesis that verbal STM served as a working memory by investigating the effects of a concurrent STM load on performance of tasks that were assumed to depend upon a limited capacity working memory. We used this technique to simulate a STM deficit as we did not have access to appropriate patients. We argued that if verbal STM is the basis of a limited capacity working memory, then a task such as immediate recall of a sequence of digits, universally assumed to depend on STM, should disrupt cognitive performance. Furthermore, the longer the concurrently stored sequence, the greater should be the disruption. In order to obtain some generality we studied three separate tasks, one involving simple reasoning, a second involving long-term learning while a third was concerned with prose comprehension. Our results were consistent across tasks in showing little effect of a concurrent load of three digits, but a much more substantial effect when the load was increased to six digits. The effect was however, far from dramatic and in one study for example where the load varied from zero to eight concurrent digits, the time taken to solve our reasoning task increased by a maximum of 50%, but with no increase in errors which remained at 5%.

At this point we abandoned the idea of a simple unitary STM system. We proposed instead a three component model comprising an attentional control system of limited capacity, the central executive, aided by two short-term storage systems, one for verbal and acoustic material, the phonological loop and its visual counterpart the visuo-spatial sketchpad as shown in Figure 1. In developing the model, we decided to focus initially on the phonological loop on the grounds that it appeared to be the most tractable, given that we already knew a good deal about it from the earlier verbal STM literature. While the central executive was likely to be the most important, it would clearly involve venturing into the complex field of attention, beyond our area of expertise. Our studies over the next few years focused on two questions, what was the overall structure of the phonological loop and what function or functions did it serve?

Figure 1. The initial working memory model proposed by Baddeley & Hitch (1974)

**The phonological loop**

We began with a simple model that distinguished between a temporary store and a rehearsal process based on vocalisation, which could be either overt or covert. The temporary store could hold information for a matter of seconds before it faded unless the memory trace was refreshed by articulatory rehearsal. This remains the core of our current interpretation of the phonological loop, although with a good deal of detailed modification over the years fulfilling our criteria for a productive theoretical model, namely that it captures existing data and is fruitful in generating new questions leading to its gradual extension and modification. The sections that follow will pursue a number of questions raised by the model, focusing particularly on those for which neuropsychological evidence has proved crucial. Three issues are central, the characteristics of the store, the means of maintaining information within it and finally its ecological function.

One basic question concerning the loop is the nature of its storage code. Conrad (1964) referred to it as an acoustic store, although this could presumably not be literally the case since it can be fed either auditorily or visually, through the covert naming of letters or words. We ourselves initially termed it the articulatory loop, later abandoning the term since we regarded articulation as the method of rehearsal rather than the storage medium itself. This raised the question of an alternative term; we initially adopted phonemic, intending it to be relatively atheoretical and when our psychoacoustics colleagues pointed out that this was not the case, switched to phonological. When this again proved to have unwanted theoretical implications we gave up, accepting our ignorance and leaving the question to those with much better understanding of speech perception and production than ourselves, although as will be explained later we suspect that both articulatory and acoustic codes may be used, typically in parallel, but under certain circumstances potentially separately.

The principal evidence for such a code comes from the initial demonstration by Conrad and Hull (1964) that serial recall of similar sounding letters such as b t v p c is considerably less accurate than a dissimilar set (r w y k f). A later study applied this to words, and as described earlier distinguished this from the effect of semantic similarity which had a much more marked effect on LTM (Baddeley, 1966a;b). Evidence that maintenance of the memory trace was through verbal rehearsal came from the demonstration that articulatory suppression, the requirement to repeatedly utter some irrelevant item such as the word the, markedly impaired STM performance, and when presentation was visual, removed the effect of phonological similarity (Murray, 1968). Further evidence for the importance of subvocal rehearsal came from the word length effect, a demonstration that serial retention of sequences of five unrelated words declined from around 90% with monosyllables to around 50% with words of five syllables. Longer word sequences took longer to articulate with a function suggesting that memory span was approximately equivalent to the number of words that could be articulated within two seconds (Baddeley, Thomson & Buchanan, 1975).

The potential neuropsychological implications of this result were pointed out by Naveh-Benjamin and Ayres (1986) who studied digit recall across a series of languages which differed in the mean syllabic length of the digits. They found that rehearsal rate increased and digit span decreased with mean number of syllables from monosyllabic English digits through Spanish and Hebrew to Arabic with an average length of more than two syllables. This was consistent with the Baddeley et al. (1975) interpretation of the use of subvocal articulation to refresh the fading memory trace. Further supportive evidence was provided by the fact that articulatory suppression removed the word length effect whether presentation was visual or auditory. The Baddeley et al. (1975) interpretation is however still controversial. The word length itself is robust but the fading trace hypothesis, remains only one of several potential explanations ranging from the effect of lexical of neighbourhood size (Jalbert, Neath, Bireta, & Suprenant, 2011), to word complexity (Brown & Hulme, 1995) and interference (Lewandowsky, Oberauer & Brown, 2009). This question is certainly of relevance to the phonological loop model, but fortunately is not critical to the assumption of a phonological loop within a broad working memory model.

When first proposed, we envisaged the articulatory loop as equivalent to a simple tape loop, providing the sole means of temporarily maintaining verbal information in serial order. It gradually became clear that this was an over-simplified view. First of all, patients with a clear but specific STM deficit such as those studied by Shallice and Warrington (1970) had reliably higher digit spans when material was presented visually, suggesting a parallel visual code. Secondly, digit span of healthy subjects when material was presented visually under articulatory suppression was reduced by about two digits, but by no means eliminated as would be predicted by a simple articulatory tape loop model.

A more nuanced version of the phonological loop was beginning to emerge. One source of this came from an attempt to study the role of the loop in reading, which showed that people can extract the meaning from passages of prose while suppressing articulation (Baddeley & Lewis 1981). It was also shown that people are capable of making judgements of homophony (e.g. slay-sleigh) of visually presented words while suppressing articulation, with no increase in speed or errors although making rhyme judgements (e.g. slay, clay) was impaired by suppression (Baddeley & Lewis, 1981; Besner, Davies & Daniels, 1981). This suggested that there may be two underlying verbal codes, one articulatory and one auditory. The articulatory code plays an important role in maintaining information through subvocal rehearsal and allows items to be held during the manipulation necessary to strip away a word’s initial phoneme in order to make a rhyme judgement. The non-articulatory auditory code allows simple homophone judgements but is not suitable for maintenance or complex manipulation. We tentatively referred to these as the “inner voice” and the “inner ear” (Baddeley & Lewis, 1981). Further evidence came from the study of special populations where a broadly similar model proved applicable both to congenitally deaf users of sign language and to lip read stimuli (Rönnberg, 2003; Rönnberg, Söderfeldt & Risberg, 2000; Wilson, 2001).

Further information on the process of rehearsal came from the study of dysarthric patient who had lost the capacity to speak as part of his locked in syndrome. (Baddeley & Wilson 1985). He proved able to make phonologically-based homophony judgements together with standard effects of both phonological similarity and word length in immediate serial recall. We replicated this in a series of other patients with less complete peripheral disruption of speech production, arguing for a separation between the capacity to set up and internally manipulate an articulatory code and the ability to turn this into overt vocalisation. Verbal STM appears to depend on this internal speech-based code rather than its peripheral manifestation in overt speech. At this point we appeared to have a simple but plausible model supported by evidence from both healthy and neuropsychologically impaired participants, a model illustrated by Vallar and Papagno (2002) in Figure 2.

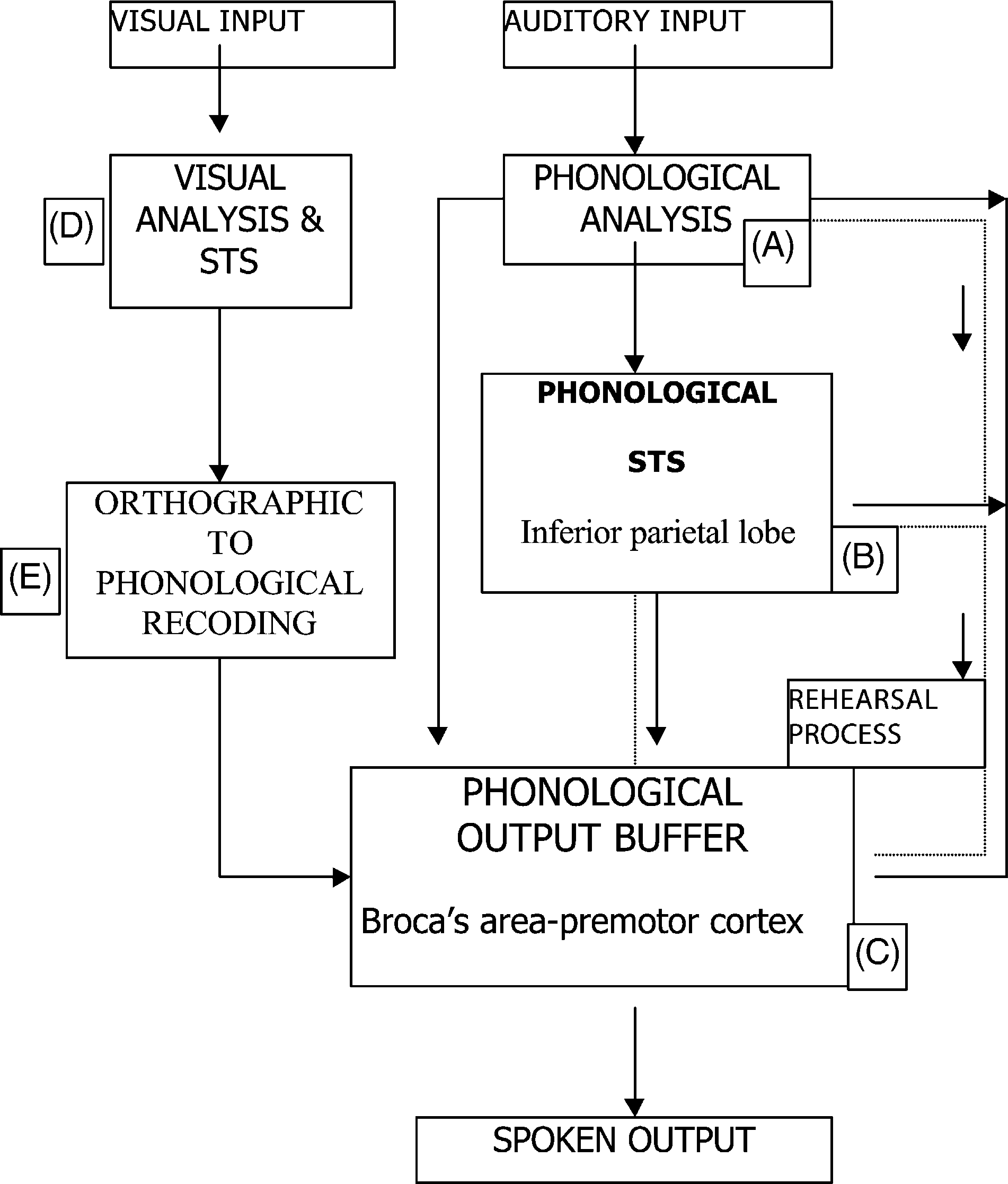


Figure 2. The elaborated model of the phonological loop proposed by Vallar & Papagno (2002).

**Computational modelling of serial ordering**

The phonological loop model did however have a major lack. The system itself, and the phenomena on which it is based, involved serial presentation of verbal material, while the model itself had no realistic mechanism for storing order information, a classic problem for which more detailed modelling was required. More specifically, the concept of a simple tape loop cannot explain the tendency for typical errors in immediate recall to involve getting the items in the wrong order. The main options for a more adequate ordering mechanism are chaining associations linking successive items (Lewandowsky & Murdock, 1989), positional associations linking items and their temporal contexts (Burgess & Hitch, 1992), and non-associative ordinal markers of relative position (Page & Norris, 1998). Chaining is unlikely to be important in immediate recall as it fails to deal with patterns of error when sequences contain repeated items. Thus, if each item serves as the cue for the next, errors should fall after repeated items. However, the typical finding is that it is the second instance of the repeat that is impaired, the Ranschburg Effect (Jahnke, 1969). A related difficulty concerns memory for sequences in which phonemically similar and phonemically dissimilar letters occupy alternate serial positions. According to chaining, errors should fall after similar items whereas errors peak on them, forming a striking sawtooth pattern (Baddeley, 1968; Henson, Norris, Page & Baddeley, 1996). These results motivated the development of computational models using ordinal or positional mechanisms for serial order.

According to the Primacy Model (Page & Norris, 1998) items decay after encoding and this leaves a primacy gradient of trace strength over items when sequence presentation ends. At recall, a simple iterative process of choosing the strongest item and then suppressing it results in items being retrieved in their original order. It is only necessary to assume the comparison process is noisy to generate order errors showing the locality constraint since adjacent items are more similar in strength than those further apart. This ordinal mechanism can be contrasted with the positional model proposed by Burgess & Hitch (1992) in which order is encoded in the form of associations between items and states of a context signal that changes gradually over time. Recall involves rerunning the context signal in steps, retrieving the most strongly associated item at each step, and suppressing it before moving on to the next. In this model errors show the locality constraint because states of the context signal are more similar for adjacent items. Both models agree in assuming two stages in recalling an item, the first processes order information at the lexical level, as outlined above, while the second involves retrieving the item’s phonological features. In both models, errors at the second stage account for phenomena originally attributed to the phonological loop, namely those of phonological similarity, word length and articulatory suppression. However, the models go further by reproducing a range of other phenomena associated with serial ordering, perhaps most notably the sawtooth serial position curve for lists of alternating similar and dissimilar items.

Each serial ordering mechanism has its merits. The ordinal account gives a better explanation of paired transpositions, where two items swap positions with one another in recall, and does better in simulating the amount of primacy in the serial position curve. However, it has more difficulty explaining memory for sequences presented in a regular rhythm (typically as groups of three items). Grouping leads to a dramatic reduction in the total number of order errors but at the same time introduces interposition errors (Henson, 1999; Ryan, 1969), where an item is recalled in the wrong group while maintaining its within-group position. Although interpositions are not inexplicable in an ordinal system, they can be modelled more straightforwardly in a positional system by assuming the context signal has an extra dimension that tracks group position (Burgess & Hitch, 1999). Unfortunately, this approach only works for typical experiments in which all groups are the same size which is known in advance, and cannot be applied to sequences with unpredictably varying rhythms such as occur in everyday speech. This limitation has, however, been overcome in a later variant known as the ‘BUMP’ model in which the context signal is generated automatically by a bank of filters that detect phase and amplitude at different frequencies in the speech stream (Hartley, Hurlstone & Hitch, 2016). This mechanism proved capable of simulating previously unexplained large and systematic effects of different irregular grouping patterns on recall (Ryan, 1969).

This theoretical development illustrates the broader challenge of how to evaluate computational models, in that BUMP is clearly an advance on Burgess & Hitch (1999) but is also more limited in scope, given that it can only be applied to auditory sequences. Beyond such differences in scope, computational models typically embody large numbers of assumptions, some of which are important because they embody core principles such as the nature of order information, while others are potentially insignificant details of implementation. Faced with such complexity, one way of taking stock is to weigh the empirical evidence for each core principle separately. When this is done, the most likely ordering mechanism for verbal sequences is a combination of ordinal and positional selection rather than either alone (Hurlstone Hitch & Baddeley, 2014). What little evidence there is regarding visual and spatial sequences points to a similar conclusion (Hurlstone et al., 2014; Hurlstone & Hitch, 2015). If one accepts this conclusion, it raises further questions such as whether serial order is processed separately in each modality, using a common solution to the problem, or whether there is instead a central, multi-modal serial ordering mechanism. We return to these broader issues later on. For the present we would argue that computational modelling has proved useful not only in developing the concept of the phonological loop to encompass a much wider range of phenomena, but also in generating further questions for future research.

**Evolutionary function**

The next stage concerned the question of what function might be served by the phonological loop? Was it anything more than, as suggested by one of our colleagues, a mere pimple on the nether regions of cognition? The chance to tackle this issue came with the discovery of a patient, PV with a very pure but substantial deficit in STM by our colleagues in Milan, Vallar and Papagno. As with earlier studies of patients with relatively pure STM deficits, PV’s other cognitive capacities appeared to be unimpaired; if the loop performed a genuinely important function beyond verbal STM, then it should be reflected in PV’s performance of certain other tasks of practical importance.

As was typical of other STM patients, PV had a lesion in the temporo-parietal region resulting in grossly reduced digit span and the lack of a clear recency effect in verbal free recall (Vallar & Baddeley, 1984). We began with a hypothesis prominent in the early years of psycholinguistics, namely that the short-term retention of verbal material was necessary to allow the “unpacking” of syntactic and semantic cues that were assumed to be necessary for comprehension. We tested this by presenting sentences of varied length which she had to judge as true or false (Vallar & Baddeley 1987). By keeping the syntactic structure the same and packing the sentence with largely irrelevant detail, we were able to demonstrate an effect of sentence length, and hence of delay, showing that she could correctly accept a sentence such as Ships are lived on by sailors and reject Sailors are lived on by ships, but unlike controls had difficulty with a sentence such as Sailors, it is commonly believed, and with considerable justification, are frequently lived on by ships. Anything less extreme than this however created few problems and it seemed unlikely that the phonological loop had evolved purely to deal with unnecessarily convoluted prose (Vallar & Baddeley, 1987).

A second possibility however was that the loop may have evolved for new phonological learning as would be required to acquire either one’s native or a second language. We chose to test this using a paired associate paradigm in which the stimuli were in her native Italian and the responses either unrelated Italian words such as castello-tavalo (castle-table) or a Russian equivalent such as fiore-sviete (flowers-sviete). We argued that learning the native language pairs would be based on semantic coding, but that this would not be possible for Russian vocabulary. The results fully supported this hypothesis; with auditory presentation, she learnt the native pairs at a normal rate, while she failed on the Italian-Russian pairs; after 10 presentations when matched controls had mastered the list of auditorily presented Italian-Russian word pairs, she had not learned a single one. With visual presentation she was still impaired, though less dramatically, presumably because of her well preserved visual STM (Baddeley, Papagno & Vallar, 1988).

We followed up this result by demonstrating that disrupting the operation of the phonological loop in healthy participants, differentially impaired foreign language vocabulary acquisition. This occurred whether disruption was achieved by concurrent articulatory suppression (Papagno, Valentine & Baddeley, 1991) or by phonological similarity (Papagno & Vallar, 1992). Further supportive evidence came from a demonstration by Papagno and Vallar (1995) that polyglots who were capable of mastering a range of languages tended to have extended memory spans. They went on to report that a highly atypical young woman with Down syndrome who had mastered several languages also had an excellent memory span (Vallar & Papagno, 1993) in contrast to the typical depression of verbal span in this syndrome (Jarrold, Baddeley & Hewes, 1999). A wide range of later studies have established the link between the phonological loop and second language acquisition, with a recent meta-analysis involving over 3000 learners across 79 samples, finding a clear association between performance and measures of both central executive and phonological loop capacity (Linck, Osthus, Koeth & Bunting, 2014).

A further series of studies investigated the link between the phonological loop and the acquisition of their native vocabulary by both typically developing and specific language impaired (SLI) children. We found the SLI group to be particularly impaired in the capacity to hear and repeat back unfamiliar nonwords. One study investigated immediate recall of spoken nonwords of varying length in an eight year old SLI group with normal nonverbal intelligence but the language development of six year olds, comparing them with two control groups matched on either age or language ability. We found a small effect of age together with a much larger SLI effect that increased with nonword length (Gathercole & Baddeley, 1990). The deficit could not be attributed to hearing or articulatory problems and was assumed to reside in the storage component of the phonological loop. Other studies based on the general population again indicated an association between nonword repetition and performance on a standard vocabulary test that involved hearing a word and then pointing to the relevant picture. This association could of course reflect either a role of the phonological loop in helping the acquisition of vocabulary or conversely that having a good vocabulary aids the recall of nonwords. Cross-lagged correlations suggested however that the phonological loop was the principal driver during the early years, although as children became older the relationship became reciprocal (Gathercole & Baddeley, 1989).

In addition to its role in acquiring vocabulary, there is evidence to suggest that the phonological loop may play a significant role in learning to read. It has long been known that delayed reading acquisition tends to be associated with reduced digit span (Miles 1993). It was furthermore suggested by Shankweiler, Liberman, Mark, Fowler and Fischer (1979) that poor reading might be the result of failure to use phonological coding, based on a study comparing verbal STM in good and poor readers in which the latter failed to show a phonological similarity effect. However it subsequently proved to be the case that the fixed sequence length used by Shankweiler et al. to test both groups was sufficiently far beyond the span of the poor readers as to lead them to abandon phonological coding (Hall, Wilson, Humphries Tinzman & Bowyer, 1983). Further evidence that phonological coding tends to be abandoned when sequence length increased was also noted by Salamé and Baddeley (1986). Baddeley, Logie and Ellis (1988) tested a group of boys diagnosed with dyslexia. When tested using span rather than fixed length procedure, they showed clear evidence of phonological coding, together with an error pattern that resembled that of younger children. Studies by Gathercole and Baddeley went on to investigate the role in reading development of both phonological STM and phonological awareness as measured by the capacity to make rhyme judgements. They found the maximum influence of verbal STM during the early stages of letter-name acquisition, while phonological awareness became important at a later stage (Baddeley & Gathercole, 1992; Gathercole, Willis & Baddeley, 1991).

**Computational modelling of learning**

The question arises as to whether computational models of the phonological loop can be extended to address non-word repetition and vocabulary acquisition. Given that most of what we know about STM is based on sequences of familiar items such as digits, letters or words, computational models have tended to focus on serial ordering at the lexical level, leaving ordering at the phonological level unspecified. Extending them to address non-word repetition and word learning involves several challenges. These include the nature of the learning mechanism and the interface with LTM in addition to dealing with serial order at multiple levels.

Hartley and Houghton (1996) showed how non-word repetition can be simulated using a positional coding model similar to that proposed by Burgess & Hitch (1999) but in which associations are formed between phonological elements and states of a modified context signal. Gupta and MacWhinney (1997) sketched out how multiple levels might be combined in a single model of this sort capable of repeating a series of syllables. However, proof of concept involves rather more than this, as was illustrated by the claim that a positional model could explain long-term sequence learning through the gradual strengthening of position-item associations (Burgess & Hitch, 1999). This was refuted by Cumming, Page & Norris (2003), who found no position-specific transfer from a learned sequence to a test sequence in which alternate items maintained the same positions. In retrospect the problem for the positional model should have been obvious as it assumed all sequences are associated with a single context signal, a proposal that would generate massive interference. Burgess & Hitch (2006) illustrated a possible solution that involved relaxing this assumption and including a process whereby an incoming sequence is matched, item by item, against previously learned sequences. Finding a match is equivalent to recognising the sequence has been encountered previously so that further learning can take place, whereas finding no match forces the model to rely exclusively on short-term memory for the current presentation. However, although the revised model predicts experimental data on learning sequences at the lexical level (Hitch, Flude & Burgess, 2009), it has not so far been extended to address learning at the phonological level, crucial for learning novel words. Using a somewhat different approach, Gupta & Tisdale (2009) have modelled the learning of phonological sequences, but have not extended this to sequence learning at the lexical level. A major challenge for future modelling in this area is how to achieve workable combinations of different levels of representation and how to balance the ability to recognise previously encountered sequences against adequate responsiveness to novelty.

**Visuo-spatial sketchpad**

At the same time as developing the concept of a phonological loop and exploring its potential application, we also began to make progress on the visuo-spatial sketchpad, although at a somewhat slower rate. Evidence for separate visual and spatial components came from single case neuropsychological studies (De Renzi, 1992; Farah, Hammond, Levine & Calvanio, 1988), behavioural studies on healthy individuals (Farmer, Berman & Fletcher, 1986; Logie, Zucco & Baddeley, 1990) and from neuroimaging based on PET (Jonides et al., 1993; Smith et al., 1995). Progress was however rather slower with some doubt for example as to whether there exists a rehearsal mechanism equivalent to subvocal articulation in the phonological loop. One suggestion was that this may be based on eye movements (Hebb, 1968) and it is certainly the case that visuo-spatial STM can be disrupted by the requirement to saccade from one point to another repeatedly (Postle, Idzikowski, Della Sala, Logie & Baddeley,2006). Logie (1995) has proposed a model that involved a separation between a store which he termed, the visual cache and a rehearsal mechanism, the inner scribe. Our own current view is that the capacity found in the phonological loop, for rehearsal based on overtly recycling the material by subvocal articulation represents an atypical form of rehearsal which, in other systems involves a process known as refreshing, whereby maintenance is based on the continuing focus of attention on the relevant material (e.g. Barrouillet & Camos, 2012; 2015). There has also been relatively little work on the possible role of the sketchpad in the development of visuo-spatial semantic memory, by analogy with the research on the role of the loop in the acquisition of vocabulary. This may partly reflect the imbalance between research on linguistic semantics and on its visuo-spatial equivalent, which presumably involves information such as the colour and shape and use of visuo-spatial aspects of the world. There has however in recent years been a major increase in research on visual working memory (e.g. Bays & Husain, 2008; Luck & Vogel, 1997; Luck & Hollingworth, 2008; Oberauer & Lin, 2017). While much of this has tended so far to focus on retention of simple stimuli over brief intervals, it seems likely that the issue of rehearsal and transfer to long-term episodic and semantic memory will become areas of increasing research in the near future.

**Episodic Buffer**

In the case of the phonological loop, while there was good evidence for its importance in acquiring phonologically novel word forms, there was little evidence for its role in comprehension, either from the study of patient PV (Vallar & Baddeley, 1984) or from the influence of articulatory suppression on comprehension (Baddeley & Lewis, 1981), with later work suggesting that semantic and syntactic processing appears to operate relatively automatically using processes within LTM (Baddeley, Hitch & Allen, 2009 ). This raised the question of the role of working memory more generally in language acquisition and comprehension. A classic paper by Daneman and Carpenter (1980) was directly stimulated by this question. They used an individual differences-based approach, correlating performance on comprehension of complex prose passages with an apparently simple measure of working memory span. They argued that the essence of working memory is the capacity both to hold information and to manipulate it at the same time, testing this by a method in which participants read out a sequence of sentences and then attempted to recall the last word of each. This proved surprisingly difficult with spans typically of around three sentences but despite its narrow range, it correlated highly with prose comprehension measure of a graduate selection test.

This result has been replicated many times (Daneman & Merikle, 1996) and while it could be argued that the original studies were simply correlating one language test with another, later work indicated that similar results could be obtained by replacing the sentence reading by series of simple arithmetic tasks (Turner & Engle, 1989), or by the even simpler information processing tasks, provided they were performed at a rate that was too rapid to allow interleaved rehearsal of the initial material (Barrouillet & Camos, 2015). Furthermore, such correlations were not limited to language comprehension, extending to a wide range of cognitive activities from the capacity to resist auditory distraction (Conway, Cowan & Bunting, 2001) through performance on a programming course (Shute, 1991) to reasoning, where working memory span was shown to correlate highly with more standard measures of intelligence (Kyllonen & Christal, 1990). This research led to extensive further work attempting to identify the critical feature of working memory span that allows it to be so powerful a predictor (see e.g. Engle, Tuholski, Laughlin & Conway, 1999; Miyake et al., 2000). This research has been fruitful in many ways, but in our own view has still not identified a single “essence” of intelligence.

From the viewpoint of the multicomponent model, this line of research has proved gratifying in demonstrating the broader importance of working memory, but also problematic, raising the question of how it could be incorporated within our model. More specifically, where in the model are the sentences being held during processing and verbal output? A load of three sentences is beyond the capacity of either the visuo-spatial or verbal subsystems while the executive was assumed at the time to be a purely attentional system with no inherent storage capacity (Baddeley & Logie, 1999). A related problem was presented by evidence that immediate serial recall of sequences of words or letters could be influenced at the same time by both visual and phonological similarity (Logie, Della Sala, Wynn & Baddeley 2000). How could the two subsystems communicate given that they are assumed to use quite different basic codes?

In response to this problem, a fourth subsystem was proposed, capable of acting as a buffer store, that is able to hold multidimensional stimuli. It was assumed also to accept input from semantic and episodic LTM in addition to information from the other subsystems of working memory, combining the various sources, as appropriate, into a series of multidimensional episodes. Termed the episodic buffer, it was also assumed to be accessible to conscious awareness. Its assumed capacity to bind information from different sources into coherent objects or episodes serves a function similar to that proposed by Baars (1997; 2002) who likens consciousness to a stage on which a range of different actors appear and are co-ordinated. Our initial version of the revised model is shown in Figure 3 (Baddeley, 2000). Our addition of a further component is vulnerable to the common complaint about information processing models that whenever they run into difficulties, they simply add another box, although one box in 25 years does not seem unduly profligate. We did however regard it as important to demonstrate that the concept of an episodic buffer was a fruitful one, leading to new findings rather than a simply offering a way of providing post-hoc justification of awkward results. To do this, we constrained the model by making an initial strong assumption that full access to the buffer was possible only through the central executive. We went on to test this by using concurrent secondary tasks to interfere differentially with the phonological loop, the sketchpad and the central executive. In an attempt to explore the generality of our results, we studied the binding of features into both visual objects such as coloured shapes, and words into meaningful sentences.

Figure 3. Modification of the multicomponent model to include a fourth component, the episodic buffer (Baddeley, 2000).

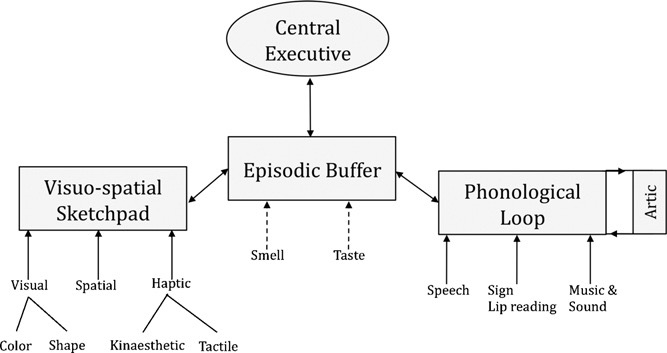
Across a series of studies we consistently obtained the same result, disruption of visual, verbal or executive components of working memory, all impaired overall performance, with the degree of disruption with verbal interference interfering most with language retention, visuo-spatial disruption was greatest for object retention while concurrent executive processing disrupted both. However despite these effects on overall performance, in none of the cases was the capacity to bind features into episodes itself impaired by concurrent working memory tasks. We concluded first of all that the multidimensional buffer system appeared to be able to input information directly from any of a number of sources including not only visual and verbal stimuli, but also directly from long-term semantic and episodic memory. We proposed secondly that, since the executive was not responsible for binding in any of the studies, this must occur “off stage”, with the buffer providing an essentially passive but still important subsystem for binding elsewhere. In the visual case this involves earlier stages of perception while in the case of the verbal binding of words into phrases and sentences, it was assumed to reflect long-term memory-based syntactic and semantic processing. The modified model is shown in Figure 4.

Figure 4. The current version of our multicomponent model of working memory (Baddeley, 2012).

Our model still retains the basic tripartite structure but has now been elaborated in a number of ways. The most obvious is the episodic buffer which makes a clear distinction between its capacity for passive storage, contrasting this with the purely attentional function of the central executive. The buffer also represents those aspects of the visuo-spatial and phonological subsystems that are accessible to conscious awareness. The loop and sketchpad are however underpinned by complex processes that are not directly open to conscious awareness. In the case of the sketchpad the representation of objects potentially reflects information not only of visual features such as colour and shape and of spatial information such as location and spatial relations, but also input from other potentially relevant features such as the tactile and kinaesthetic aspects of perceived objects. Object representations may then link with knowledge in long-term memory. Similarly, we see the phonological loop as providing a confluence point for language-related material, not only sound and phonology but also associated language-based features provided by lip reading and sign language, allowing them to be bound into more complex verbal linguistic episodes. We thus envisage the two systems as being at the confluence point of a number of subsystems with much of the processing likely to operate at a procedural level with integrated episodes subsequently reaching awareness within the episodic buffer. We have tentatively included in the model further potential buffers for taste and smell although at present these are purely speculative.

The current working memory model could thus be seen as a hierarchical structure based on a series of interacting buffer stores. Each store is potentially able to combine information from two or more sources which are likely to be processing information based on different but relatable codes probably operating at different input rates. Combining codes and co-ordinating input rates are both likely to require temporary storage. What we have presented is thus a broad conceptual model, aimed at bringing together existing information from a range of fields and we hope, generating productive new questions. How successful has this been?

**Overview**

In the 40 years since it was originally produced, the multicomponent working memory model has of course not escaped criticism, some no doubt generated by the question of its theoretical nature, is it a theory or a model? And if a model, does it make specific predictions and what are they? We ourselves regard it as providing a theoretical framework offering a coherent account of how memory and attention combine to facilitate complex cognition. As such, the primary criterion has been its usefulness, both in capturing what we already know, both from the laboratory and the neuropsychological clinic, while suggesting further tractable questions. We also aimed to provide an account that is simple enough to be understood beyond the field of experimental cognitive psychology while at the same time producing a series of tools that will indicate whether or not this model appears to fit. By exploring cognition and its breakdown across a range of practical situations we are able to probe the limits of the current model, providing evidence on its strengths and limitations. As such, our approach is closer to that of the philosopher Stephen Toulmin (1953) who regards theories as cognitive maps than to that of Braithwaite (1953) or the early views of Popper (1959) who saw Newtonian physics as the classic model, and the capacity to make precise falsifiable predictions as the acid test of the adequacy of the theory, criteria that would of course be inconsistent with Darwin’s theoretical contribution to science. Our own approach is much closer to that of Lackatos (1968) who proposes that what he terms “dogmatic falsification” should be replaced by the concept of a research programme which may be progressive, in the sense of stimulating further empirically sustained evidence or degenerative in the case of theories whose principal function is defense of the theory’s “hard core”. He rejects the idea of any theory being “true” and accepts that progressive theories may contain anomalies and hence be “false”, to be replace in due course by theories that are less false.

Our policy therefore has been to use our broad framework to generate questions, using the answers to further elaborate the conceptual model. This approach does not of course allow us to decide whether the framework is right or wrong, but has, we believe, been fruitful in generating further knowledge, rather than merely defensive of the status quo, the crucial criterion for models of this broad type (Lakatos, 1968). This degree of flexibility is of course a two-edged sword making the model open to misinterpretation. A good example of this is provided by Nairne (2002) who demonstrates convincingly that the phonological loop component of the multicomponent model does not itself provide an adequate account of working memory. We agree with this conclusion, but not with his treating the phonological loop as “the standard model” rather than simply one interesting but limited component of a much more complex system.

A related problem of the model is its breadth and complexity, extending as it does from perception through attention to episodic and semantic memory. This can lead to attempts to absorb it into one or more of these related systems, then denying its usefulness. For example Allport (1984) claimed to detect a subtle auditory processing deficit in one of the classic Shallice and Warrington patients with a short-term memory deficit, concluding that it was unnecessary to assume a memory problem. This attempt to explain a large and broadly replicated memory deficit in terms of a small problem in hearing has had little general impact, but does illustrate the potential complexity of the route from the sound signal to phonological storage, a route that can be disrupted at a number of points as in the case of the patient suffering from word deafness studied by Baddeley and Wilson (1993a).

A rather more general attempt to account for the phenomena of verbal STM without the need to assume a buffer store is that of Macken, Taylor and Jones (2014) who assume a direct mapping of auditory perception onto action in the form of verbal rehearsal proposing that this obviates the need for an intervening buffer store. However, if this is to be anything beyond a relabeling it will need considerable elaboration before becoming a strong candidate as an explanation of the extensive existing evidence from both healthy individual and patients. Furthermore, it does not currently offer an explanation of the remaining components of our working memory model.

Reflecting on the development of the concept of the phonological loop, the influence of neuropsychological studies, often based on single cases is striking. This contrasts with the much more limited contribution made by more recent studies based on neuroimaging. Why should this be the case? The link between mainstream cognitive psychology and neuropsychology extends back to the work of Milner (1966) and Warrington and Weiskrantz (1968), challenging the dominant concept of a unitary memory system. This led in due course to a more explicit link to developments in cognitive psychology which suggested separate long- and short-term memory systems that map directly onto the neuropsychological evidence (Baddeley & Warrington, 1970; Shallice & Warrington, 1970), followed in due course by patient-based studies concerned with the functional role of such systems (Vallar & Baddeley, 1987; Baddeley, Papagno & Vallar, 1988). This line of research was further extended by studies of dysarthria and dyslexia (Baddeley, Logie & Ellis, 1988; Baddeley & Wilson, 1985; 1993b) that directly influenced cognitive theory.

In contrast, while early studies, largely using PET provided clear evidence for the separation of visual and verbal STM (Jonides et al., 1997; Jonides et al., 1993) and for separating the anatomical locations principally associated with phonological storage and rehearsal (Paulesu, Frith & Frackowiak, 1993), subsequent contributions have been much less clear. The work of Jonides and associates was seminal during these early stages, but subsequently moved away from a fractionation approach to one in which working memory was assumed to be widely distributed across areas concerned with LTM. This duly led to speculations of some complexity (Jonides et al., 2008) that appear to be difficult to test productively using either behavioural or neuroimaging methods. A recent attempt to use meta analysis to fractionate executive control by Nee et al., (2013) tested a number of hypotheses, notably including the claims by two influential approaches to working memory for the importance of inhibitory control (Engle, 1996; Friedman & Miyake, 2004) finding that the only structural distinction to emerge from available neuroimaging data was that between visual and verbal processes. Whilst it is plausible to assume such a separation, it does not represent a major step forward from the PET studies carried out by Jonides and colleagues many years earlier (Jonides et al., 1993; 1997).

This contrast between the apparent fruitfulness of neuropsychological and neuroimaging approaches prompts a number of questions: is the early evidence of fractionation simply misleading? If not, why were single case studies so fruitful in the past but not now? If so, will this continue to be the case? We will discuss these in turn.

As we have already explained, we regard the value of a scientific model as reflecting its capacity to account for existing data, its potential for asking fruitful new questions and finally its applicability beyond the laboratory and clinic. While details continue to be disputed, we would argue that the broad concept of a phonological loop fulfils all three of these (Baddeley, 1997; 2012). We suggest that the contribution made by single case studies reflects a happy coincidence between the development of cognitive theory and the presence of a number of distinguished neuropsychologists who combined day to day patient contact with a sufficient understanding of current developments and cognitive psychology to identify theoretically important cases, coupled with a willingness to work with theoretically-oriented cognitive psychologists. This was probably more characteristic of Europe than North America, where the clinical tradition is somewhat more psychometrically oriented with a tendency to emphasise group rather than individual case studies, although notable exceptions certainly exist. One problem with the single case approach however, is that it depends on this rare and unpredictable association between the patient, the neuropsychologist and the cognitive psychologist, limiting it to a relatively few fortunate investigators. In contrast, neuroimaging approaches do not necessarily require access to patients and can be carried out by any group with sufficient expertise and funding.

Unfortunately however, the inevitable limitations accompanying the development of neuroimaging have made it difficult to apply fruitfully to an area as complex as working memory. One problem is that over much of this period neuroimaging research has focused on anatomical localisation, attempting to correlate activity in specific areas of the brain associated with hypothetical cognitive systems, albeit more recently with attempts to study the links between such areas. Simple localization does not fit well with theories such as our own multicomponent approach which regards working memory as potentially linking many rapidly interacting systems. Consider for example a simple task such as the immediate recall of a sequence of five unrelated words. Experimental behavioural studies suggest that recall relies predominantly on the phonological characteristics of the words. While accepting that such words may indeed activate their semantic associates, these are not typically used in this context (Baddeley, 1966a). Of course, knowing how to pronounce the words will also involve long-term phonological memory in order to retrieve and pronounce the lexical units to be recalled. In short, it is clear that many aspects of LTM, implicit and explicit, declarative and procedural are likely to be active in the operation of working memory. Given the probable complexity of this relationship, it is reasonable for a theorist such as Cowan (2005) when focusing on the attentional aspect of working memory to use the term “activated long-term memory” as a conceptual place-holder for this aspect of his model, just as we have used the term central executive to refer to complexities of attentional control when focusing principally on short-term storage. It is important to note however that, as an explanation, simply regarding working memory as activated long-term memory is no more satisfactory than to “explain” language as activated long-term memory. Computational modelling of the phonological loop (Burgess & Hitch, 2006) provides a concrete illustration of this point in that the underlying mechanisms involve a combination of changes in the activation of pre-existing long-term connections with the formation of novel short-term connections. In our view the question whether working memory is better viewed as activated long-term memory or a separate system is based on a false dichotomy.

A second problem with much of the work over this period stems from the use of methods such as fMRI which depend on averaging activation over a period at least of several seconds, a period during which a range of different working memory processes are likely to occur. In this connection, it is notable that EEG based evoked response methods are proving much more fruitful in investigating the processes underlying working memory (Eimer & Kiss, 2010; Vogel & Machizawa, 2004). We suggest therefore that we remain optimistic that as the range of relevant methods develops, the link between the cognitive psychology and the cognitive neuroscience of working memory will once again prove highly productive.

References

Allport, A. (1984). Auditory verbal short-term memory and conduction aphasia. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and performance X: Control and language processes* (pp. 351-364). Hillsdale, NJ: Erlbaum.

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). New York: Academic Press.

Baars, B. J. (1997). *In the Theater of Consciousness*. New York: University Press.

Baars, B. J. (2002). The conscious access hypothesis: origins and recent evidence. Trends in Cognitive Sciences, 6, 47-52.

Baddeley, A. D. (1966a). Short-term memory for word sequences as a function of acoustic, semantic and formal similarity. *Quarterly Journal of Experimental Psychology, 18*, 362-365.

Baddeley, A. D. (1966b). 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. (1968). How does acoustic similarity influence short-term memory? *Quarterly Journal of Experimental Psychology, 18*, 362-365.

Baddeley, A. D. (1997). Working memory and language processing. *Japanese Society of Aphasiology Journal, 17*(3), 189-200.

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences, 4*(11), 417-423.

Baddeley, A. (2012). Working memory: Theories, models and controversies. *Annual Review of Psychology, 63*, 1-29.

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. (1992). Learning to read: The role of the phonological loop. In J. Alegria, D. Holender, J. Junca de Morais, & M. Radeau (Eds.), *Analytic approaches to human cognition.* (pp. 153-167). Amsterdam: Elsevier Science

Baddeley, A. D., & Hitch, G. (1993). The recency effect: implicit learning with explicit retrieval? *Memory & Cognition, 21*, 146-155.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. A. Bower (Ed.), *Recent Advances in Learning and Motivation* (Vol. 8, pp. 47-89). New York: Academic Press.

Baddeley, A. D., Hitch, G. J., & Allen, R. J. (2009). Working memory and binding in sentence recall. *Journal of Memory & Language, 61*, 438-456.

Baddeley, A. D., & Lewis, V. J. (1981). Inner active processes in reading: The inner voice, the inner ear and the inner eye. In A. M. Lesgold & C. A. Perfetti (Eds.), *Interactive Processes in Reading* (pp. 107-129). Hillsdale, N.J.: Lawrence Erlbaum.

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

Baddeley, A. D., & Logie, R. H. (1999). Working memory: The multiple component model. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 28-61): Cambridge University Press.

Baddeley, A. D., Logie, R. H., & Ellis, N. C. (1988). Characteristics of developmental dyslexia. *Cognition, 29*, 197-228.

Baddeley, A. D., Papagno, C., & Vallar, G. (1988). When long-term learning depends on short-term storage. *Journal of Memory and Language, 27*, 586-595.

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior, 14*, 575-589.

Baddeley, A. D., & Warrington, E. K. (1970). Amnesia and the distinction between long- and short-term memory. *Journal of Verbal Learning and Verbal Behavior, 9*, 176-189.

Baddeley, A. D., & Warrington, E. K. (1973). Memory coding and amnesia. *Neuropsychologia, 11*, 159-165.

Baddeley, A. D., & Wilson, B. (1985). Phonological coding and short-term memory in patients without speech. *Journal of Memory and Language, 24*, 490-502.

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

Baddeley, A. D., & Wilson, B. A. (1993b). A developmental deficit in short-term phonological memory: Implications for language and reading. *Memory, 1*, 65-78.

Barrouillet, P., & Camos, V. (2012). As time goes by: Temporal constraints in working memory. *Current Directions in psychological Science, 21*, 413-419.

Barrouillet, P., & Camos, V. (2015). *Working memory: Loss and reconstruction*. Hove: Psychology Press.Bays, P. M., & Husain, M. (2008). Dynamic shifts of limited working memory resources in human vision. *Science, 21*, 851-854.

Besner, D., Davies, J., & Daniels, S. (1981). Reading for meaning: The effects of concurrent articulation. *Quarterly Journal of Experimental Psychology, 33A*, 415-438.

Bjork, R. A., & Whitten, W. B. (1974). Recency-sensitive retrieval process in long-term free recall. *Cognitive Psychology, 6*, 173-189.

Brathwaite, R. B. (1953). *Scientific explanation.* Cambridge: Cambridge University Press.

Broadbent, D. E. (1958). *Perception and communication*. London: Pergamon Press.

Brown, J. (1958). Some tests of the decay theory of immediate memory. *Quarterly Journal of Experimental Psychology, 10*, 12-21.

Brown, G. D. A., & Hulme, C. (1995). Modelling item length effects in memory span: No rehearsal needed? *Journal of Memory and Language, 34*, 594-621.

Burgess, N., & Hitch, G. J. (1992). Toward a network model of the articulatory loop. *Journal of Memory and Language, 31*, 429-460.

Burgess, N., & Hitch, G. J. (1999). Memory for serial order: A network model of the phonological loop and its timing. *Psychological Review, 106*, 551-581.

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.

Campoy, G., & Baddeley, A. D. (2008). Phonological and semantic strategies in immediate serial recall. *Memory, 16*, 329-340.

Cermak, L. S. (1976). The encoding capacity of a patient with amnesia due to encephalitis. *Neuropsychologia, 14*, 311-326.

Cermak, L. S., Butters, N., & Moreines, J. (1974). Some analyses of the verbal encoding deficit of alcoholic Korsakoff patients. *Brain and Language, 1*, 141-150.

Cermak, L. S., & Reale, L. (1978). Depth of processing and retention of words by alcoholic Korsakoff patients. *Journal of Experimental Psychology: Human Learning & Memory, 4*, 165-174.

Conrad, R. (1964). Acoustic confusion in immediate memory. *British Journal of Psychology, 55*, 75-84.

Conrad, R., & Hull, A. J. (1964). Information, acoustic confusion and memory span. *British Journal of Psychology, 55*, 429-432.

Cowan, N. (2005). *Working memory capacity.* Hove: Psychology Press.

Conway, A. R. A., Cowan, N., & Bunting, M. F. (2001). The cocktail party phenomenon revisited: The importance of working memory capacity. *Psychonomic Bulletin and Review, 8*, 331-335.

Craik, F. I. M., & Levy, B. A. (1970). Semantic and acoustic information in primary memory. *Journal of Experimental Psychology, 86*, 77-82.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing. A framework for memory research. *Journal of Verbal Learning & Verbal Behavior, 11*, 671-684.

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

da Costa Pinto, A., & Baddeley, A. D. (1991). Where did you park your car? Analysis of a naturalistic long-term recency effect. *European Journal of Cognitive Psychology, 3*, 297-313.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behaviour, 19*, 450-466.

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review, 3*, 422-433.

De Renzi, E. (1982). *Disorders of space exploration and cognition*. Chichester: Wiley.

Eimer, M., & Kiss, M. (2010). An electrophysiological measure of access to representations in visual working memory. *Psychophysiology, 47*, 197-200.

Engle, R. W. (1996). Working memory and retrieval: An inhibition-resource approach. In J. T. E. Richardson, R. W. Engle, L. Hasher, R. H. Logie, E. R. Stoltfus, & R. T. Zacks (Eds.), *Working memory and human cognition* (pp. 89-119). New York: Oxford University Press.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General, 128*, 309-331.

Farah, M. J., Hammond, K. M., Levine, D. N., & Calvanio, R. (1988). Visual and spatial mental imagery: Dissociable systems of representation. *Cognitive Psychology, 20*, 439-462.

Farmer, E. W., Berman, J. V. F., & Fletcher, Y. L. (1986). Evidence for a visuo-spatial scratch-pad in working memory. *Quarterly Journal of Experimental Psychology, 38A*, 675-688.

Friedman, N. P., & Miyake, A. (2004). The relations among inhibition and interference control functions: A latent variable analysis. *Journal of Experimental Psychology: General, 133*, 101-135.

Gathercole, S., & Baddeley, A. (1989). Evaluation of the role of phonological STM in the development of vocabulary in children: A longitudinal study. *Journal of Memory and Language, 28*, 200-213.

Gathercole, S. E., & Baddeley, A. D. (1990). Phonological memory deficits in language-disordered children: Is there a causal connection? *Journal of Memory and Language, 29*, 336-360.

Gathercole, S. E., Willis, C., & Baddeley, A. D. (1991). Differentiating phonological memory and awareness of rhyme: Reading and vocabulary development in children. *British Journal of Psychology, 82*, 387-406.

Glanzer, M., & Cunitz, A. R. (1966). Two storage mechanisms in free recall. *Journal of Verbal Learning and Verbal Behavior, 5*, 351-360.

Gupta, P., & MacWhinney, B. (1997). Vocabulary acquisition and verbal short-term memory: Computational and neural bases. *Brain and Language, 59*, 267-333.

Gupta, P., & Tisdale, J. (2009). Does phonological short-term memory causally determine vocabulary learning? Toward a computational resolution of the debate. *Journal of Memory and Language, 61*, 481-502.

Hall, J. W., Wilson, K. P., Humphreys, M. S., Tinzmann, M. B., & Bowyer, P. M. (1983). Phonemic similarity effects in good vs. poor readers. *Memory and Cognition, 11*, 520-527.

Hartley, T., & Houghton, G. (1996). A linguistically constrained model of short-term memory for nonwords. *Journal of Memory and Language, 35*, 1-31.

Hartley, T., Hurlstone, M.J., & Hitch, G.J. (2016). Effects of rhythm on memory for spoken sequences: A model and tests of its stimulus-driven mechanism. *Cognitive Psychology, 87,* 135-178.

Hebb, D. O. (1968). Concerning Imagery. *Psychological Review, 75*, 466-477.

Henson, R. N. A. (1999). Positional information in short-term memory: Relative or absolute? *Memory and Cognition, 27*, 915-927.

Henson, R. N. A., Norris, D. G., Page, M. P. A., & Baddeley, A. D. (1996). Unchained memory: Error patterns rule out chaining models of immediate serial recall. *Quarterly Journal of Experimental Psychology, 49A (1)*, 80-115.

Hitch, G.J., Flude, B. & Burgess, N. (2009). Slave to the rhythm: Experimental tests of a model for verbal short-term memory and long-term sequence learning. *Journal of Memory and Language, 61*, 97-111.

Hurlstone, M., & Hitch, G. J. (2015). How is the serial order of a spatial sequences represented? Insights from transposition latencies. *Journal of Experimental Psychology: Learning, Memory & Cognition, 41*, 295-324.

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

Jahnke, J.C. (1969). The Ranschburg effect. *Psychological Review, 76*, 592-605.

Jailbert, A., Neath, I., & Surprenant, A. M. (2011). Does length or neighborhood size cause the word length effect? *Memory and Cognition, 39*, 1198-1210. doi:10.3758/s13421-011-0094-z

Jarrold, C., Baddeley, A. D., & Hewes, A. K. (1999). Dissociating working memory: Evidence from Down’s and Williams’ syndrome. *Neuropsychologia, 37*, 637-651.

Jonides, J., Lewis, R. L., Nee, D. E., Lustig, C. A., Berman, M. G., & Moore, K. S. (2008). The mind and brain of short-term memory. *Annual Review of Psychology, 59*, 193-224. doi:10.1146/annurev.psych.59.103006.093615

Jonides, J., Schumacher, E. H., Smith, E. E., Lauber, E. J., Awh, E., Minoshima, S., & Koeppe, R. A. (1997). Verbal working memory load effects regional brain activation as measured by PET. *Journal of Cognitive Neuroscience, 9*, 462-475.

Jonides, J., Smith, E. E., Koeppe, R. A., Awh, E., Minoshima, S., & Mintun, M. A. (1993). Spatial working memory in humans as revealed by PET. *Nature, 363*, 623-625.

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working memory capacity. *Intelligence, 14*, 389-433.

Lakatos, I. (1968). Criticism and the methodology of scientific research programmes *Proceedings of the Aristotelian Society, 69*, 149-186.

Lewandowsky, S., & Murdock, B. B. (1989). Memory for serial order. *Psychological Review, 96*, 25-57.

Lewandowsky, S., Oberauer, K., & Brown, G. D. (2009). No temporal decay in verbal short-term memory. *Trends in Cognitive Sciences, 13*, 120-126. doi:10.1016/j.tics.2008.12.003.

Linck, J. A., Osthus, P., Koeth, J. T., & Bunting, M. F. (2014). Working memory and second language comprehension and production: A meta-analysis *Psychonomic Bulletin & Review, 21*, 861-883.

Logie, R. H. (1995). Visuo-spatial working memory. Hove, U.K.: Erlbaum.

Logie, R. H., Della Sala, S., Wynn, V., & Baddeley, A. D. (2000). Visual similarity effects in immediate serial recall. *Quarterly Journal of Experimental Psychology, 53A*, 625-646.

Logie, R. H., Zucco, G. M., & Baddeley, A. D. (1990). Interference with visual short-term memory. *Acta Psychologica, 75*, 55-74.

Luck, S. J., & Hollingsworth, S. (2008). *Visual Memory*. New York: Oxford University Press.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390*, 279-281.

Macken, W. J., Taylor, J. C., & Jones, D. M. (2014). Language and short-term memory: The role of perceptual-motor affordance. *Journal of Experimental Psychology: Learning , Memory and Cognition, 40*, 1257-1270. doi:10.1037/a0036845

Miles, T. R. (1993). *Dyslexia: The pattern of difficulties* (2nd ed.). London: Whurr.

Milner, B. (1966). Amnesia following operation on the temporal lobes. In C. W. M. Whitty & O. L. Zangwill (Eds.), *Amnesia* (pp. 109-133). London: Butterworths.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41*, 49-100.

Murray, D. J. (1968). Articulation and acoustic confusability in short-term memory. *Journal of Experimental Psychology, 78*, 679-684.

Nairne, J. S. (2002). Remembering over the short-term: The case against the standard model. *Annual Review of Psychology, 53*, 53-81.

Naveh-Benjamin, M., & Ayres, T. J. (1986). Digit span, reading rate, and linguistic relativity. *Quarterly Journal of Experimental Psychology, 38*, 739-751.

Nee, D. E., Brown, J. W., Askren, M. K., Berman, M. G., Demiralp, E., Krawitz, A., & Jonides, J. (2013). A meta-analysis of executive components of working memory. *Cerebral Cortex, 23*, 264-282. doi:10.1093/cercor/bhs007

Oberauer, K., & Lin, H.-Y. (2017). An Interference Model of Visual Working Memory. *Psychological Review, 124*, :21-59.

Page, M. P. A., & Norris, D. (1998). The primacy model: A new model of immediate serial recall. *Psychological Review, 105*, 761-781.

Papagno, C., Valentine, T., & Baddeley, A. D. (1991). Phonological short-term memory and foreign language vocabulary learning. *Journal of Memory and Language, 30*, 331-347.

Papagno, C., & Vallar, G. (1992). Phonological short-term memory and the learning of novel words: The effect of phonological similarity and item length. *Quarterly Journal of Experimental Psychology, 44A*, 47-67.

Papagno, C., & Vallar, G. (1995). Verbal short-term memory and vocabulary learning in polyglots. *Quarterly Journal of Experimental Psychology, 48A*, 98-107.

Paulesu, E., Frith, C. D., & Frackowiak, R. S. J. (1993). The neural correlates of the verbal component of working memory. *Nature, 362*, 342-345.

Peterson, L. R., & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology, 58*, 193-198.

Popper, K. (1959). *The Logic of Scientific Discovery*. London: Hutchison.

Postle, B. R., Idzikowski, C., Della Sala, S., Logie, R. H., & Baddeley, A. D. (2006). The selective disruption of spatial working memory with eye movements. *Quarterly Journal of Experimental Psychology, 59*, 100-120.

Rönnberg, J. (2003). Cognition in the hearing impaired and deaf as a bridge between signal and dialogue: a framework and a model. *International Journal of Audiology, 42*, S68-76.

Rönnberg, J., Söderfeldt, B., & Risberg, J. (2000). The cognitive neuroscience of signed language. *Acta Psychologica, 105*, 237-254.

Ryan, J. (1969). Grouping and short-term memory: Different means and pattern of grouping. *Quarterly Journal of Experimental Psychology, 21*, 137-147.

Sachs, J. S. (1967). Recognition memory for syntactic and semantic aspects of connected discourse. *Perception and Psychophysics, 2*, 437-442.

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

Shankweiler, D., Liberman, I. Y., Mark, L. S., Fowler, C. A., & Fischer, F. W. (1979). The speech code and learning to read. *Journal of Experimental Psychology: Human Learning and memory, 5*, 531-545.

Shallice, T., & Warrington, E. K. (1970). Independent functioning of verbal memory stores: A neuropsychological study. *Quarterly Journal of Experimental Psychology, 22*, 261-273.

Shute, V. J. (1991). Who is likely to acquire programming skills? . *Journal of Educational Computing Research, 7*, 1-2.

Smith, E. E., Jonides, J., Koeppe, R. A., Awh, E., Schumacher, E. H., & Minoshima, S. (1995). Spatial vs. object working memory: PET investigations. *Journal of Cognitive Neuroscience, 7*, 337-358.

Sperling, G. (1967). Successive approximations to a model for short-term memory. *Acta Psychologica, 27*, 285-292.

Toulmin, S. (1953). *The Philosophy of Science*. London: Hutchison.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task-dependent? *Journal of Memory and Language, 28*, 127-154.

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., & Baddeley, A. D. (1984). Fractionation of working memory. Neuropsychological evidence for a phonological short-term store. *Journal of Verbal Learning and Verbal Behaviour, 23*, 151-161.

Vallar, G., & Baddeley, A. D. (1987). Phonological short-term store and sentence processing. *Cognitive Neuropsychology, 4*, 417-438.

Vallar, G., & Papagno, C. (1993). Preserved vocabulary acquisition in Down's Syndrome: The role of phonological short-term memory. *Cortex, 29*, 467-483.

Vallar, G., & Papagno, C. (2002). Neuropsychological impairments of verbal short-term memory. In A. D. Baddeley, M. D. Kopelman, & B. A. Wilson (Eds.), *Handbook of memory disorders.* (2nd. ed., pp. 249-270). Chichester: Wiley.

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

Vogel. E., & Machizawa, M.G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature, 428*, 748-751.

Warrington, E. K., & Weiskrantz, L. (1968). New method of testing long-term retention with special reference to amnesic patients. *Nature, 217*, 972-974.

Wilson, B. A., & Moffat, N. (1992). *Clinical management of memory problems* (2nd ed.). London: Chapman and Hall.

Wilson, M. (2001). The case for sensorimotor coding in working memory. *Psychonomic Bulletin & Review, 8*, 44-57.