Testing a positional model of the Hebb effect.

Nick Cumming

Mike Page

Dennis Norris

MRC Cognition and Brain Sciences Unit,

15 Chaucer Road,

Cambridge CB1 2EF

nick.cumming@mrc-cbu.cam.ac.uk

ABSTRACT

In two experiments, we investigate the hypothesis that a strengthening of position –item associations underlies the improvement seen in performance on an immediate serial recall task, when a given in list is surreptitiously repeated every third trial. Having established a strong effect of repetition, performance was tested on transfer lists in which half the items held the same position as in the repeated list (S-items), the remainder moved (D-items). In Experiment 1, S-items showed a small advantage over control and D-items, in order errors. A second experiment tested whether a design element in Experiment 1 underlay this advantage. When the experimental design was better controlled, no improvement was shown for either S- or D-items over controls. These data were shown to be inconsistent with the results of computer simulations of a positional model. An alternative model is outlined.

INTRODUCTION

Motivation

The working memory (WM) theory, proposed by Baddeley and his colleagues (Baddeley, 1986; Baddeley, 1992; Baddeley & Hitch, 1974) has proved a productive qualitative framework within which to account for a great deal of data relating to short-term memory for serial order. The immediate serial recall (ISR) task in particular has been used extensively in research into the phonological loop component of WM. In ISR, participants are asked to recall a list of items such as letters or words in their correct order immediately after presentation. The WM framework gives a good qualitative account of many of the effects seen in ISR, such as the word-length effect (Baddeley, Thompson, & Buchanan, 1975), the list-length effect, the effect of articulatory suppression (Murray, 1968), and the phonological similarity effect (Conrad, 1964) (see Baddeley, 1986, for a review).

There is growing evidence, both neuropsychological (e.g. Baddeley, Papagno, & Vallar, 1988; Papagno, 1996; Papagno, Valentine, & Baddeley, 1991; Papagno & Vallar, 1992; Trojano, Stanzione, & Grossi, 1992; Vallar & Baddeley, 1984a; Vallar & Baddeley, 1984b; Vallar, DeBetta, & Silveri, 1997; Vallar, Papagno, & Baddeley, 1991; Warrington & Shallice, 1969) and developmental (e.g. Baddeley, Gathercole, & Papagno, 1998; Gathercole, 1995; Gathercole, Service, Hitch, Adams, & Martin, 1999; Gathercole, Willis, & Baddeley, 1991), that phonological working memory plays an important role in the acquisition of novel phonological forms during vocabulary learning. It has been found that patients, such as PV and SC (Baddeley, 1993; Baddeley, Papagno, & Vallar, 1988; Papagno, Valentine, & Baddeley, 1991; Vallar & Baddeley, 1984a), who have a very low auditory span, around two or three items, have great difficulty learning new vocabulary. In terms of developmental evidence, Gathercole, Baddeley and colleagues (see above) have found that, in children, the ability to repeat nonwords predicts later vocabulary size. Indeed, Baddeley, Papagno and Gathercole (1998) have suggested that vocabulary acquisition is one of the primary roles of phonological short-term memory.

The working memory framework has two major shortcomings when it comes to addressing the issue of how long-term ordered phonological representations are formed. First, it does not deal directly with the problem of serial order (Lashley, 1951), in that it has no formal description of how the order of items is represented in phonological working memory. Second, it suggests no mechanism for learning the information it stores. Baddeley et al. (1998) have argued that the primary function of the phonological store is to permit the learning of new phonological representations. They base this claim on evidence that children's ability to repeat nonsense words, essentially lists of phonemes or syllables, is correlated with later vocabulary size. In order to repeat a nonsense word, the listener must be able to maintain an ordered representation of the phonological forms contained, and learning of this short-term representation is required for a long-term phonological representation to be formed. Given this claim, detailed models of these processes are required.

There are now several computational models of performance in the ISR task, all of which have explicit ordering mechanisms. These provide fits of varying quantitative precision to empirical data from many aspects of ISR data. They include Lee and Estes's perturbation model (Lee & Estes, 1977; Lee & Estes, 1981), the TODAM model (Lewandowsky & Murock, 1989; Li & Lewandowsky, 1995; Murdock, 1996; Murdock, 1992; Murdock, 1993; Murdock, 1995), ACT-R (Anderson, Bothell, Lebiere, & Matessa, 1998), Burgess and Hitch's phonological loop model (Burgess & Hitch, 1992; Burgess & Hitch, 1996; Burgess & Hitch, 1999), the OSCAR model, (Brown, Preece, & Hulme, 2000), the Start-End model (Henson, 1998) and the Primacy Model

(Page & Norris, 1998). The implications of this research for some of these models will be discussed in more detail later.

There is now considerable converging evidence that the phonological store plays an important role in establishing the long-term representations required for vocabulary learning. In the present paper, we attempt to discover whether the phonological store plays a more general role in long term learning of serial order. We do this by studying the Hebb effect (Hebb, 1961) which can be seen as a model of the acquisition of long-term representations of ordered phonological material. In the Hebb effect, memory for serial order develops gradually with repeated presentations of lists. The Hebb effect enables us to test the predictions of at least one current model of ordered memory.

The Hebb effect

Hebb presented subjects with 24 lists of nine digits, with every third list being identical without the participants' prior knowledge of the repetition. The other lists were non-repeating. Contrary to his predictions (Hebb, 1949), participants showed significantly increased performance on the repeating list compared with the other lists. This recall advantage for the repeating lists over the non-repeating lists that are interleaved between them is known as the Hebb repetition effect. It has been replicated several times (Cunningham, Healy, & Williams, 1984; Fendrich, Healy, & Bourne, 1991; Melton, 1963; Schwartz & Bryden, 1971; see Seger, 1994, for a review). Although the primary task in the Hebb-effect paradigm is explicit recall, the Hebb effect itself is generally assumed to be an implicit learning task. Most participants typically report being unaware of the repetition of the critical list. In the original experiment by Hebb (1961), participants showed no non-specific practice effect; that is, there was no improvement in performance on the nonrepeating lists while the performance on the repeating list improved. Melton (1963), replicated the experiment, but used a larger set of lists (80 as opposed to 24 lists), and showed a non-specific practice effect. In the experiments described below, close to 100 lists are shown, so some general improvement is expected over the course of an experiment.

There are, however, some experimental conditions under which a Hebb-effect might be expected but none has been observed. For example, if at least two items at the start of the repeating list are changed on each repetition (Schwartz & Bryden, 1971), or if there are more than five intervening lists between each repetition of the critical list (Melton, 1963), no Hebb effect is found. Cohen & Johansson (1967) found that rehearsal without an overt response did not result in learning of the repeating sequence. This finding was replicated by Cunningham (1984), who used an experiment where subjects were told to rehearse the whole of an 8-item list, grouped into two four-item chunks, but to recall only one chunk. When it was not indicated which chunk was to be recalled until after presentation of the entire list, there was still no evidence of learning for the chunk which had not been repeatedly recalled, compared with the chunk which had been repeatedly recalled.

Positional coding in models of serial order memory

Several influential models of serial order memory use position-item associations to encode order (e.g. Brown et al., 2000; Burgess & Hitch, 1999, Henson, 1998; Lee and Estes, 1981). Learning is explained in these models in terms of the strengthening of these positional codes. Each time an item is presented in a certain position, the strength of the association between that item's representation and a positional code is incremented slightly. The phonological loop model of Burgess and Hitch (1992; 1996; 1999) has implemented this concept most explicitly in terms of simulating the Hebb effect (Burgess and Hitch, 1999). It uses the strengthening of long-term weights between items and the state of a context, or timing, signal to simulate Hebb effect learning. The primary purpose of this research is to test the predictions this model makes about the long-term representation of order, thereby constraining the types of model that should be used to simulate the operation of the phonological loop component of WM.

Other types of models that use alternative means of representing order exist, some of which will be discussed later. However, for the purposes of understanding the empirical work and simulations carried out, only the concept of the strengthening of position-to-item associations employed by positional models should be borne in mind.

Experimental task

In the two experiments described in this paper, there is a common structure to the procedure used. In both, Hebb-effect learning, as described above in relation to the Hebb (1961) experiment, is induced in the usual manner. Once a repeated Hebb list has undergone learning, performance is measured on a "transfer" list that is constructed from the Hebb list in a particular way. More specifically, the transfer list is derived from the learned Hebb list by keeping alternate list items in the same serial position while randomly re-arranging the remaining items. Figure 1 shows an example re-ordering of the Hebb list. Before experimental predictions regarding performance on the transfer list can be examined, the Hebb effect itself must be replicated. Performance must be shown to have increased for a repeated list relative to nonrepeated controls (henceforth called "filler" lists). The Hebb effect is indexed by a comparison between the repeating Hebb lists and the nonrepeating filler lists. Once this is done, the experimental design allows a comparison between four types of list item. The items from the final presentation of the Hebb list (H) can be compared with the items from filler lists matched for practice (F) as a check on the main result. The transfer lists comprised of alternating items that either retained the position they occupied in the Hebb lists (S-items) and items that were randomly repositioned (D-items). In half of the experimental blocks, the transfer list started with an S-item and in the other half, it started with a D-item. To simplify statistical analysis, reconstructed "lists" of S-items (S) and D-items (D) are compared with performance on H and F when examining error patterns over the list as a whole.

/* * * * * * * * * * * * * * * Insert figure 1 about here please * * * * * * * * * * * * /

Predictions of positional models on the experimental task.

A pure positional model would predict a recall advantage for the items that stay in the same position as they occupied in the repetition-learned list over items that have moved positions. More recent models, such as the Burgess and Hitch (1992; 1996; 1999) model, do not necessarily make such predictions. In the Burgess and Hitch model, items are not associated directly with a single position as such, but with the activation pattern of the time-based "context" layer described above. The slowly changing context layer is associated at presentation, through Hebbian learning, with the item nodes. Importantly, there is overlap in the context layer activation patterns corresponding to adjacent and near-adjacent positions. Thus, when the context layer is "replayed" at recall, the correct item is maximally activated, and the nearby item nodes activate to a lesser extent. As described earlier, the items compete for output selection through a noisy-choice procedure, resulting in some recall errors. These errors will therefore happen most often between adjacent items in the list.

According to the Burgess and Hitch model, when a transfer list is presented in which alternate items have been repeated in the Hebb lists, these items have a large long-term weight component to the relevant pattern on the context layer. The items on either side of the strengthened item are therefore less likely to be selected in error. Thus, on transfer lists, one of two things might be predicted: either the level of recall will be generally higher on alternating lists than on non-repeating lists because fewer one-apart transpositions are being made on both S- and D-items; or alternatively, items that stay at the same position will receive a large-enough activation boost from the long-term context-item weights that fewer errors will be made on the S-items than on the D-items. It is also possible that the strengthening of context-to-item weights for alternate items will cause these items to be recalled too early, due to their "inadvertently" strong connections with the overlapping context pattern corresponding to the position preceding that in which they have been learned. Because the balance of each of these effects is difficult to predict qualitatively, these issues are explored quantitatively in computational simulations presented later in the paper. First, however, we present the experimental findings.

EXPERIMENT 1

In this experiment, we examined the error patterns observed on a transfer list itself derived from a list that has been previously presented eight times under a traditional Hebb-repetition paradigm. The transfer list was presented on the trial immediately following the eighth presentation of the Hebb list, as illustrated in the top panel of Figure 2. Figure two illustrates the slight difference in design between the two experiments. The reasons for this minor difference will be explained in the discussion of this experiment.

/* * * * * * * * * * * * * * * Insert figure 2 around here please * * * * * * * * * * * */

Method

Participants

Twenty-four members of the CBU panel of paid volunteers, 13 females and 11 males aged between 17 and 42, took part in the experiment.

Materials and design

All lists were 10-item combinations of the digits 0 to 9, using no repeats and subject to the following constraints: there were no runs of three or more consecutive digits in ascending or descending order; there were no items in the filler lists that shared a serial position with the same digit in the current repeating Hebb list; no filler list was used twice. As in Hebb (1961), every third list in a given block was identical. Each block contained eight repetitions of the Hebb list and, hence, 24 lists in total. The first list of the second block was the transfer list derived from the first-block's repeating Hebb list, and so on.

Four blocks of trials were used, using one of two manipulations to derive the transfer list. In one manipulation, the items of the Hebb list at odd-numbered serial positions were presented at the same serial position in the transfer list, while the items at even-number serial positions were randomly re-assigned to a different even-numbered serial position in the transfer list. In the other, the items in even-numbered serial positions were presented at the same serial position in the transfer list, while the items at odd-number serial positions were randomly re-assigned to a different odd-numbered serial position in the transfer list (see Figure 1 for a possible construction of a transfer list). Thus, each participant received two Hebb-repetition blocks at the end of which the transfer list was constructed in an odd-same, even-different manner (SDSD), and two where the construction was even-same, odd different (DSDS). The total number of lists presented to each subject was 97, comprising four Hebb blocks of 24 lists plus a transfer list for the final Hebb list.

Procedure

The lists were presented one digit at a time in large (72-point font) black type in the centre of a computer screen. Each digit was visible for 700ms, followed by a 180ms period where no digit was visible. The presentation of each list was initiated by the participant using a key-press. Response was manual; participants typed their responses into the computer number pad, the responses appearing in a horizontal array of boxes on the screen. Participants were told to omit a response rather than guess. They indicated an omission by pressing the enter key, and a dash appeared on the screen in the appropriate response-box.

The participants' responses were recorded. These were scored not only such that the number of correct responses was counted, but also so that when an error was made the type of error and the serial position of the error was recorded. This permitted a detailed analysis of error types across serial position. As the set of possible responses was limited to the digits 0 to 9 and all members of that set were presented in every trial, it was possible to classify all errors as one of two types, as follows:

Omission errors: The subject pressed the enter key to indicate not knowing what the item was.

Order errors: The participant made a non-omission response, but the digit was recalled in the wrong serial position.

Transfer lists were of the structure outlined in Figure 1 above, such that S-items fell either on the odd-numbered serial positions, or on the even-numbered serial positions. Each participant saw two of each type of transfer list, on alternate blocks in the experiment, and the type of transfer list seen first was counter-balanced across participants. Thus, performance for both S-items and D-items could be calculated at all serial positions for all subjects.

Results.

The Hebb effect.

As discussed in the introduction, it is necessary to determine whether or not there was a Hebb effect present in the data before proceeding with further analyses.

Before analysing the transfer-list data, it is necessary to demonstrate that recall performance on Hebb lists improves above and beyond that observed in the filler-list data. To do this, the gradients of improvement in number of correct responses were calculated using least-squares linear regression for the Hebb lists (lists 3, 6, 9, etc., in each block) and the filler list immediately preceding each Hebb list (lists 2, 5, 8, etc.). This results in a data point for each condition, which can be expressed as an items-per-presentation change in recall performance. These data were analysed using a paired-sample t-test, which showed a greater mean improvement gradient across the presentations of the Hebb lists than the immediately preceding fillers, $t\ (23) = 4.84$, $p < 0.001$. The Hebb lists showed a mean improvement of 0.3 items correct per presentation, compared with an improvement of 0.02 for the filler lists.[1]

[1] Learning gradient data were also analysed for differences in learning rates across experimental blocks, using a four (blocks) by two (list types) repeated measures ANOVA. This showed no main effect of block and a significant main effect of list type, $F(1, 23) = 18.93$, $p < 0.001$.

Error types during Hebb list learning

The error types made during Hebb list learning were analysed using a two (list types) by two (error types) by eight (Hebb list repetitions) ANOVA. These data are shown in Figure 3. As this shows, it appears that order errors and omission errors decrease at approximately the same rate in the Hebb condition, but remain relatively constant in the filler condition.

As expected, there was a significant effect of list type, $F(1, 23) = 29.67$, $p < 0.001$, such that Hebb lists were recalled significantly better than the filler lists (mean errors of each type per list were 1.83 and 2.42 respectively). There was a significant main effect of repetition of the Hebb list, $F(7,161) = 10.18$, $p < 0.001$. There was no main effect of error type (but see experiment 2, where omission errors decreased more than order errors). There was a significant interaction between list type, i.e. Hebb lists vs. filler lists, and repetition of the Hebb list, $F(7,161) = 5.71$, $p < 0.001$, such that the number of errors of both types fell with repetition of the Hebb list, while the number of each type of error made per filler list stayed relatively constant over the same period of eight repetitions. There were no other significant effects or interactions.

/* * * * * * * * * * * * * * * * * Insert figure 3 about here please * * * * * * * * /

Serial position curves

Positional models predict that a list in which half the items retain the serial positions they held in a previously learned list, should show a recall advantage over a randomly-ordered list in which no items share serial positions with the learned list. The serial position curves of the final presentation of the Hebb list in each block, the immediately preceding filler lists and the transfer lists were compared. A three (list types) by 10 (serial positions) repeated measures ANOVA was carried out on the proportion-correct data.

Mauchley's test of sphericity showed a significant skew in error distribution on serial position, $W(44) = 0.018$, $p < 0.001$, giving an epsilon value of 0.55 for the Huyn-Feldt correction. Thus the serial position degrees of freedom reported below are altered using this correction.

There was a significant main effect of list type; $F(1.96, 45) = 17.62$, $p < 0.001$, a significant effect of serial position, showing a typical serial position curve with normal primacy and recency portions; $F(4.96, 114) = 37.38$, $p < 0.001$, and no interaction.

Planned comparisons (multiple comparisons corrected for using Tukey's studentized range statistic, q) showed that recall accuracy on the final Hebb lists was significantly greater than either filler lists ($p < 0.001$) or transfer lists ($p < 0.01$).

Figure 4 shows that the serial position data displays the usual extended primacy portion and a small one-or-two item recency portion, and that performance on the final presentation of the Hebb list is higher across all serial positions than either the filler lists or the transfer lists. Recall performance plots are shown separately in figure 4 for the two types of transfer lists; those starting with an S-item (SDSD) and those starting with a S-item (DSDS). As can be seen, the S-items seem to be more accurately recalled than filler items at early serial positions. In order to investigate further, the types of errors made on the four item types outlined earlier.

/*********** Insert figure 4 about here please **************/

The error types made during transfer list recall were analysed. The mean number of errors per list for each item type are shown in Table 1. A four (item types) by two (error types) repeated measures ANOVA was conducted on the data from the four transfer trials per subject, the final repetition of the Hebb list from each block and the immediately preceding filler list. The item-type data did not meet sphericity requirements, $W(5) = 0.42$, $p < 0.01$, therefore the degrees of freedom in this analysis were corrected using the Huyn-Feldt correction with an epsilon value of 0.8. There was a significant main effect of item type; $F(2.4, 52.6) = 10.87$, $p < 0.001$, but not of error type, and a significant interaction between the two, $F(2.58, 56.7) = 4.14$, $p < 0.05$. Planned comparisons showed that there were significantly fewer errors made on Hebb-list items than on any other item type.

/ * * * * * * * * * * * * * * * Insert Table 1 about here please * * * * * * * * * * * * /

Inspection reveals a rather different pattern of order errors for S-items than for D- or F-items that could be contributing to the interaction between item type and error type. In order to investigate this further, the levels of the error type variable were analysed separately. Analyses of variance were carried out on the number of order errors and omission errors made for each item type. In the order errors analysis, using a Huyn-Feldt correction for nonsphericity of variance where $\varepsilon = 0.811$ there was a main effect of item type, $F(2.43, 55.95) = 4.09$, $p = 0.01$ and planned comparisons using Tukey's studentized range statistic (q) show fewer errors made on the S-items than on the D-items ($p < 0.05$) and a near-significant trend towards fewer order errors on the S-items than the filler items. There were also fewer order errors on the H-items than on either the D-items or F-items . In the omission errors analysis, using a Huyn-Feltd correction for nonsphericity of $\varepsilon = 0.725$, there was a main effect of list type, $F(2.41, 55.45) = 12.64$, $p < 0.001$

In order to test the hypothesis that the S-items show fewer order errors than D-items and F-items because of short-term positional protrusions of responses from one list to the next, protrusion distances were calculated for every order error. These calculations were made separately for those filler trials that followed Hebb list trials and those that followed other filler trials. In other words, for each order error in a list, the position in which that digit appeared in the previous list was compared to the position in which the item was given as an erroneous response. For example, if the list "9,7,4,1,6,3,5,2,8,0" was presented and the previous list had been "6,2,0,4,9,1,7,8,3,5" and the participant's response was "9,7,1,4,6,5,–,2,8,0", the order errors in the list are "1" in position 3, "4" in position 4 and "5" in position 6. These appeared in the previous list in positions 6, 4 and 10 respectively, giving protrusion distances of -3, zero and -4. The normalized distributions of these protrusion distances are shown in Figure 6. All these errors are normalized for the number of opportunities an error has to occur. That is, there are ten opportunities in a list for an erroneous response to have been in same serial position in the previous list, but only two for an item to have been eight items earlier in the list. Dividing the observed number of errors by the number of opportunities for that error to occur effectively controls for the triangular distribution that would occur by chance. As Figure 5 shows, there is a pronounced peak in the distribution at zero-distance, that is, the order error was most likely to have "come from" the same position in the previous list. In order to test the hypothesis that this zero-peak is significantly different from zero, a t-test was carried out on the distribution data, comparing the difference between the number of zero-distance protrusions with the average of all other protrusions. The peak is statistically reliable in both cases; after Hebb lists, $t(23) = 3.33$ $p < 0.01$, after filler lists $t(23) = 2.64$, $p < 0.05$, and there was no reliable difference in the size of the peaks when analysed using a two (filler after Hebb lists vs filler after filler lists) by two (peak vs average of the rest) repeated measures ANOVA, $F(1,23) < 1$.

What this peak shows is that if a similar proportion of protrusions occur between the final Hebb list and the transfer list (the immediately following list) then there will be a reduction in the overall number of order errors made on the S-items, since a zero-distance protrusion in these circumstances happens to be a correct response.

/ * * * * * * * * * * * * * * insert Figure 5 about here please * * * * * * * * /

Discussion

These data showed a strong Hebb effect; an average increase of around 2 items correct per list over a block containing eight repetitions. It is interesting to note that the two error types identified seem to fall at roughly the same rate through learning. Obviously, the degree to which participants make omission errors is very dependent on the amount of freedom they have to omit responses. In many immediate serial recall experiments, the participant must make as many responses as there were items, guessing to fill in items they do not know.

Although there are no differences found between the serial position curves of the S-item, D-item and F items, there was a difference found in the total numbers of order errors made between S-items and F- and D-items, with fewer order errors being made on S-items than either D-items or F-items. This result would appear to be supportive of the positional account of serial recall: It is one of the predictions of this type of model that the strengthening of a position-item association will result in fewer order errors than non-strengthened position-item associations, such as those involving F-items or D-items. However, the data collected here gave cause for scepticism. The sources of this scepticism will now be outlined.

First, the number of omission errors made was the same for the S-items as both the D-items and F-items, and significantly greater than the number made on the H-items. This is not predicted by the Burgess and Hitch (1992; 1996; 1999) model, or by the class of positional models in general. If the long-term component of the connection weight between context signal and list item has been strengthened, as it should have been in the case of S-items, the activation that item node receives should cause it to be omitted less often than either F-items or D-items. F- and D-items will not receive as much activation as S-items, hence F- and D- items should fall below the hypothesised omission threshold more frequently than S-items. This is made doubly clear when one notes that the strong Hebb effect observed in these data is based just as much on a decrease in omission errors as it is on a decrease in order errors[2].

Second, there was a concern with the design, as follows. McNicol (1978) noted that items that appear in the same serial position in consecutive lists in an ISR task are recalled slightly better than items that change position. McNicol hypothesized that strengthening of some positional coding caused the increase in recall accuracy on the following list. These short-term positional effects are a potential source of confounding effects in the current design (see figure 2, top panel). The idea, in the current experimental context, is that under the influence of some short-term positional context, items from the previous list occasionally "protrude" into recall of the current list and, when they do so, have a tendency to preserve their previous within-list position. When this happens under normal ISR conditions, i.e. on a filler list, this results in an error, since the target item (the item presented in the current list) is different from the recalled item (the item presented in the previous list). However, when the same positional context exists an items in successive lists, the result is a small increase in the proportion of correct responses.

[2] However, note that in experiment 2, the decrease in errors was concentrated more in omission errors than in order errors.

Thus there is a potential confound between the existence of a short-term positional code that results in this tendency for items to protrude into the same serial position, and the mechanism of long-term order learning. While these data demonstrate a small effect of positional context, our assessment of the Hebb effect's contribution to performance on the transfer list is compromised. Note that in Experiment 1, the transfer list comes immediately after the last repetition of the Hebb list; five of its ten items maintain the position they had in the previous list, thereby increasing the opportunities for inadvertent correct answers. Thus, a performance advantage will be expected on S-items regardless of any improvement due to Hebb repetition. The lack of any difference in the size of the effect between filler lists that followed a Hebb list and filler lists that followed other filler lists supports this conclusion. If there had been a greater number of zero-distance protrusions following a Hebb list, this might have argued in the favour of the strengthening of positional codes.

It was decided to test whether the apparent support of the order-error data for the positional model of Burgess and Hitch is due to this confound in the design of the first experiment or is a genuine effect which will be present whether or not the transfer list is the list immediately following the final Hebb list. If it is the case that the repeating Hebb list is learned by the long term strengthening of position-item or context-item connections, then the S-items of the transfer list should still be recalled better than F- or D-items even after several intermediate lists. This was tested in Experiment 2.

EXPERIMENT 2

The experimental design was changed slightly so as not to present the transfer list immediately after the final Hebb list. The transfer list, in this experiment, was presented on the fourth list after the final Hebb list. The Hebb list was then re-presented a further three lists later, to check that performance on, and therefore the long-term representation of, the Hebb list had not decayed to baseline levels. The design of this experiment, in terms of the progression of trials, can be seen in the lower panel of Figure 2, above. Only two blocks are shown for convenience. As can be seen from this, the transfer list occurs immediately after the first presentation of the repeating list from the following block, which, since it has not been presented before is essentially a novel (filler) list following another filler list, thus removing the confounding influence of the short-term positional protrusions from the preceding Hebb list list. Any positional protrusion errors will still be errors, rather than "inadvertently" correct responses.

Method

Participants

Forty members of the CBU panel of paid volunteers, aged between 17 and 40 participated in the study.

Materials, Design and Procedure

The materials, design and procedure were identical to those of Experiment 1, with the exception that, after each block of eight Hebb lists, if the final repetition of the Hebb list is said to be list N, the transfer list was presented at list N + 4, and the Hebb list was presented again at list N + 7. The final Hebb lists appeared at trials 24, 48, 72 and 96, the transfer lists on trials 28, 52, 76 and 100 and the control Hebb lists on trials 27, 51, 79 and 103. This control Hebb list added a further item type to those available for analysis, giving five: H – final Hebb list items; CH – control Hebb list items; F – filler list items; S –

transfer list items that retain the same serial position in the transfer list as the Hebb list; and D – transfer list items that have different serial positions in the Hebb list and transfer list.

Results

Hebb effect

The gradients of the lines of best fit through the recall data for both Hebb lists and the immediately preceding filler lists were calculated for each block. A two (list type: Hebb vs. filler) by four (block) repeated measures ANOVA was carried out on these gradient data. There was a highly significant main effect of list type, $F(1,39) = 33.79$, $p < 0.001$, and no other significant effect or interaction, thus showing a highly reliable increase on the Hebb lists above and beyond any non-specific practice effect. The mean gradient collapsed across groups were 0.32 and 0.05 items correct per list per presentation respectively.

Error types during Hebb list learning

The error types made during Hebb list learning data were analysed using a two (list types) by two (error types) by eight (Hebb list repetitions) repeated measures ANOVA.

This showed a significant main effect of list type, $F\ (1, 39) = 70.30$, $p < 0.001$ such that more errors were made on filler lists than Hebb lists and a significant main effect of error type, $F\ (1, 39) = 4.50$, $p < 0.05$, such that more omissions than transposition errors were made overall. There was a reliable main effect of Hebb list repetitions, $F\ (7, 273) = 13.2$, $p < 0.001$, and a significant downward linear trend in number of errors made with each repetition of the Hebb list, $F(1, 39) = 53.36$, $p < 0.001$. There was a significant interaction between Hebb list repetition and list type, $F(7, 273) = 6.74$, $p < 0.001$, showing that there is a reduction in errors made with repetitions of the Hebb list, but no commensurate reduction in filler list errors over the same period. A significant three way interaction, $F(7, 273) = 2.89$, $p < 0.01$, shows that this reduction in Hebb list errors is greater for omission errors than order errors, in contrast to the result of the same analysis in experiment 1. The difference in the error-type profiles of the experiments presented here are most likely attributable to participant adherence to instructions. These interactions can be seen easily in Figure 6.

/* * * * * * * * * * * * * * * * insert Figure 6 about here please * * * * * * * * * * /

Serial position curves

As with Experiment 1, the serial position data were analysed for differences in recall performance between the eighth presentation of the Hebb list, the filler list immediately preceding the eighth Hebb list and the transfer list. The additional, delayed presentation of the Hebb list was also included in this analysis as a check on whether the long-term representation of the list built up over the first eight presentations decays rapidly. These serial position curves are shown in Figure 7, below. A four (list types) by ten (serial positions) ANOVA was carried out on these data. Due to a violation of the assumption of sphericity in the serial position data, $W(44) = 0.001$, p 0.001, the Huyn-Feldt correction was applied to the degrees of freedom in these data, $\varepsilon = 0.381$. The analysis of variance showed significant main effects of list type, $F(2.58, 98.08) = 26.29$, $p < 0.001$ and serial position, $F(3.43, 130.36) = 51.54$, $p < 0.001$. There was no significant interaction. Figure 7 shows the serial position curves, with the transfer lists split into those that started with an S-item and those that started with a D-item. As this illustrates, there is no difference in recall between the filler lists and the transfer lists, and no difference between the eighth presentation of the Hebb list and the delayed control Hebb list. The zig-zag pattern in the serial position curves predicted by the This is confirmed by planned comparisons between the levels of the list-type variable, using the Bonferroni correction for multiple comparisons. These showed that recall accuracy on transfer and filler lists did not differ, that recall accuracy on the eighth presentation of the Hebb list and the control Hebb list did not differ, and that recall was significantly more accurate on both the eighth Hebb list presentation and the control Hebb list than on either the transfer list or the filler list.

/* * * * * * * * * * * * * * * * * insert figure 7 about here please * * * * * * * * * */

Error types comparison

As with the previous experiment, the errors were classified as order errors or omission errors and the different item types compared in terms of both error types. These are shown in table 2, below. A two (error types) by five (item types) repeated measures ANOVA was carried out on these data. There was a non-significant trend towards more omission errors than order errors; $F\ (1, 39) = 3.24$, $p = 0.08$, and a significant main effect of item type; $F\ (2.88, 112.3) = 36.52$, $p < 0.001$ (Degrees of freedom corrected using the Huyn-Feldt correction at $\varepsilon = 0.72$ for a violation of the assumption of sphericity).

/* * * * * * * * * * * * * * * * * insert Table 2 about here please * * * * * * * * * */

Planned comparisons between item types, using the Newman-Keuls studentized range statistic, showed no difference between the H-items and C-items ($p > 0.4$) and no differences between the S-items, D-items and F-items (all $p > 0.5$), and large (all $p < 0.001$) differences between all combinations of H- and C-items and S-, D- and F-items.

As the values in table 2 show, the apparent effect found in Experiment 1 is not evident once the confounding factor has been removed; that is, once the transfer list is not subject to protrusion errors from the final Hebb list, there is no longer any advantage for the S-items in order errors.

Discussion

Experiment 2 was a replication of Experiment 1 with one crucial difference; the confound introduced by having the transfer list immediately following the final repetition of the Hebb list was removed. In this second experiment, there were no differences between the S-items and the D-items of the transfer list, and neither the S-items nor the D-items differed in performance from items in filler lists. There were no differences apparent between the S-items, D-items or F-items in either number or types of error. This is inconsistent with a model that uses positional codes to store information about serial order and uses strengthening of those positional codes to simulate the Hebb effect, that is, to learn long-term information about serial order.

From a qualitative examination of how the Burgess and Hitch (1992,1996,1999) model works, it appears that the model could not give the pattern of results shown in this experiment. Qualitative interpretation of complex models is, however, a risky business. In order to both test whether, as claimed by Burgess and Hitch (1999), the Hebb effect is handled adequately by their model, and whether it can replicate the data from this experiment, a computational implementation of their model was required.

Introduction

The Burgess and Hitch (1999) model of serial order memory uses a localist connectionist framework to simulate the operation of the phonological loop component of working memory (Baddeley, 1986). It uses four layers of nodes connected as shown in Figure 8.

/* * * * * * * * * * * * * * * * * insert figure 8 about here please * * * * * * * * * */

Heavy lines indicate full connectivity. Narrow lines indicate one-to-one connections.

Each node in the phoneme layers is a localist representation of a phoneme. Activation of an input phoneme causes automatic activation of the corresponding output phoneme and vice-versa. For visual presentation, an item node is activated, suppressing all other item nodes. The output phoneme nodes associated with that item are activated, thus activating a subset of the input phonemes. Weighted connections differentially activate the item nodes, which compete for selection, in a “winner takes all” race. The context layer, which has a slowly changing pattern of activation, is associated with the selected item through fast-learning, fast-decaying weight change. At recall, the progression of context layer states is “played back” activating items that then compete in a winner-takes-all selection process. The winner’s long-term connection weights with the current pattern of the context layer and input phoneme layer are strengthened slightly. It is this strengthening of long-term weights that Burgess and Hitch use to model the Hebb effect and other repetition-learning effects (see Burgess and Hitch, 1999, for a full description).

In Burgess and Hitch (1999), the Hebb effect is simulated by manually increasing the long-term context-item association weights of the lists. Unsurprisingly the model is more accurate on these “Hebb” lists than on simulation runs where the long-term context-item association weights are fixed at zero.

Two sets of simulations will be presented here. First, we will discuss an attempt to directly model the Hebb effect through repeated presentation of the same list interleaved with non-repeating trials. Second, the results of simulations that used the method employed by Burgess and Hitch (1999) will be presented.

Direct Modelling of the Hebb effect

Method

Simulations were carried out in an attempt to model the Hebb effect directly. That is, by setting the model’s long-term learning rate parameter, ΔW_{lt}, to a small positive value, and allowing the long-term context-item weights to be altered by a small amount each cycle. Burgess and Hitch state that the model had been implemented without long-term decay because this would be negligible over the single-trial time-course used in all their simulations, but that slow decay was assumed to act on all long-term weights. Thus, for these simulations, an additional parameter was introduced to the model, namely the rate of decay of the long-term weight component. This was defined as the power to which the short-term weight decay rate was raised and was a small positive value between zero and one. To illustrate this, if the short-term decay rate was 0.75, the value used in all the current simulations, a long-term decay constant of 0.1 would mean long term weights decayed at a rate of $0.75^{0.1} = 0.97$. In this way, long-term context-item associations could decay slowly over time, avoiding the potential complication of the long-term context-item associations saturating, preventing further learning.

The normal Hebb-effect design was implemented, such that on every third trial, the model was presented with the sequence 1 to 9, and all other lists were random permutations with no repetition or substitution of those digits.

Simulations

To explore the space of parameter combinations thoroughly optimisation techniques were employed on the output from the model. The factors manipulated were the long-term decay factor, the long-term learning rate, ΔW_{lt} and the maximum the long-term weight, W_{lt} (max). The model's performance was marked such that for each parameter setting, the gradient of the least-squares line of best fit was calculated separately for performance on repeating lists and non-repeating lists, as was done for the data from experiments 1 and 2. By comparing these values with each other and with zero, a measure of the model's performance was derived, such that negative gradient values were penalised (performance should not get worse), as were very low performance values at any point (performance should not start at floor). Simulations where performance on repeating lists improved more than performance on non-repeating lists scored lowest. A gradient-descent algorithm was employed in an attempt to find optimal parameter values that would result in a significant Hebb effect. Several starting points were chosen for the optimisation to reduce the chance of the optimisation merely finding a local minimum. The results of the best-performing model are shown below in Figure 9. It is clear that the performance chart in Figure 9 is not a good simulation of real Hebb effect learning. The optimisation routines have chosen values of the free parameters such that performance decreases sharply initially, then recovers, with the repeated presentation of items at the same serial position on the Hebb lists allowing faster recovery on those lists than on the filler lists.

/ * * * * * * * * * insert Figure 9 about here please * * * * * * * * * * * * * * * /

Simulation of the Hebb effect following Burgess and Hitch (1999)

Given the failure of our attempts to make the model to produce a Hebb effect directly, in the following simulations, therefore, the Hebb effect was not simulated by updating the long-term weights on a trial-to-trial basis. Rather, the effect was simulated in the same, artificial manner as it was in Burgess and Hitch (1999), that is the parameter in the model that fixes the number of times a list is presumed to have been seen was manipulated to simulate final Hebb lists, transfer lists and filler lists. The failure to replicate the basic Hebb effect, except under conditions where the program's parameters are specifically manipulated to instantiate it, casts doubt on this model as an adequate description of the performance improvement found with Hebb repetition.

Method

To simulate the Hebb effect as implemented by the Burgess and Hitch model, the model parameter that sets the number of assumed previous repetitions of the list, N(Hebb), was set to 8 and the maximum value of the long-term context-item weight, W_{lt}(max), was set to 0.25. This value was chosen because it was high enough to allow some weight-change before saturation, but not so high that the long-term component could come to dominate the short-term, decaying weights. The value of the increment in long-term context-item weight per repetition (ΔW_{lt}) was varied from near zero to a value that would take the long-term weight for an item seen repeatedly in the last eight lists close to W_{lt}(max). Ten thousand trials at each value of ΔW_{lt} were carried out.

To simulate filler list performance, 10000 trials were carried out with N(Hebb) set to zero, and with the other variables manipulated in the same way.

A final set of simulations were carried out where, at recall only even-numbered serial positions received any activation from the long-term component of the context-item association weights (W_{lt}), thus simulating transfer trials where D-items fall on odd-numbered serial positions.

Simulations and discussion

As expected, given that the Hebb effect is being simulated through parameter setting, lists whose items have been given an activation boost equivalent to having been presented with, and having recalled, a list 8 times, showed a recall advantage over lists whose items had not been so boosted. This advantage increased with the value of ΔW_{lt}. Unsurprisingly, filler performance did not alter with increasing ΔW_{lt}. The top panel of Figure 10 shows the gain in Hebb list performance across the parameter set used in these simulations.

/* * * * * * * * * * * * * * Insert Figure 10 about here please * * * * * * * * * * * */

When the parameters relating to the maximum weight a context-item node can take and the size of the increment in long-term weight per presentation were set so as to give a Hebb effect comparable to the size of effect found in empirical data (about 12 – 13 % recall advantage - $W_{lt}(max) = 0.25$ and $\Delta W_{lt} = 0.013$), a saw-tooth pattern is seen in those simulated serial position curves for which only even-numbered serial positions are given the activation-boost associated with the model's implementation of the Hebb effect, as shown in the lower panel of figure 10. When the rate of long-term weight learning is increased beyond this level, performance on these lists becomes worse. This is because as the strength of the S-items becomes too large they tend to be recalled too early in the list and, moreover, they are often then repeated later in the list when they overcome the decaying response repetition inhibition. This performance is in line with the predictions made earlier, that in the Burgess and Hitch model S-items would be recalled better than filler items, and as it turns out, D-items until they reach a point at which their strength causes them to be recalled too early. Both patterns are inconsistent with the data collected in Experiment 2.

GENERAL DISCUSSION

In two experiments and with computer simulations, we have shown the inadequacy of the Burgess and Hitch (1999) model of the phonological loop with respect to its ability to account for Hebb-effect learning. Moreover, by showing a robust Hebb effect in the absence of any advantage for items that maintain their Hebb positions in a transfer list, we have called into question all those theories which posit that the Hebb effect results from the strengthening of position-item associations. That is not to say that that we believe there to be no role for position-item associations in models of serial recall. The effect of protrusions from the previous list evinced here in Experiment 1 and by McNicol (1978) indicates that positional effects can be seen in ISR experiments. This supports other studies that have found similar positional effects (as discussed and extended in Henson, 1998). While such positional effects are well documented, they appear to be short-lived, not contributing greatly, if at all, to the observed Hebb effect, since the small advantage in the number of order errors for items occupying the same serial position that they were learned in disappears when they do not appear in that position in a list immediately subsequent to the learned list. We believe (Page

& Norris, 1998) that these effects originate outside of the phonological loop, the system primarily involved in standard immediate serial recall tasks.

The question has been raised of what is being learned in the Hebb effect, as presented in the current study. The response method in this study is manual keying of responses into a keypad. Therefore there remains the possiblility that learning here is not of a long-term phonological representation of the repeating list, but of a motor sequence, analogous to serial reaction time learning (e.g. Nissen and Bullemer, 1987; Stadler, 1992). In serial reaction time experiments, implicit learning of a sequence is hypothesised to occur, since there is no evidence of an explicit representation of the sequence of key-presses, yet a speeding of responses and a reduction of errors is observed. It seems unwise to discount any possible motor-learning component to the Hebb effect, indeed, evidence (Cumming, 2001) suggests that there is indeed some contribution of response-learning in the Hebb effect. When responses are made such that a different motor pattern is required for correct recall of each presentation of the Hebb list, learning is slower. However, a purely implicit learning explanation seems unlikely, as there is evidence (Hebb, 1961; McKelvie, 1987) that the majority, but by no means all, participants in this task are aware of the repeating sequence.

If we assume for the moment that the Hebb effect is an experimental analogue of those phonological memory processes that underlie the learning of phonological word forms (i.e., vocabulary) then it is perhaps not surprising that a model based on position-item associations does not fair too well. It is rather difficult to see how a position-item association model could usefully be applied to vocabulary learning. First, the beginning and ends of words, and hence the within-word position of any

given sublexical unit, are not reliably marked in speech input. In those experimental analogues of word-form learning carried out with infants by Saffran et al. (1996; 1997) any markers to word boundaries (e.g., stress pattern, etc.) were deliberately avoided. Nonetheless, infants as young as 7 months old were able to recognise on a later test those pseudowords that had been repeated within the otherwise featureless stream of syllables that was used as a training stimulus. Second, even if within word position were marked, the result of training with stimuli containing multiple words would presumably be position-item associations that represented an amalgam of the full set of position-item correspondences, that is, the relative frequencies of different items in different positions. But such an encoding would have nothing to offer with regard to the learning of individual vocabulary items. In order to have any bearing on vocabulary learning proper, each word would have to engage its own, unique, positional context vector with which to encode the within-word positions of the sublexical items of which it was composed. But the idea of having one context vector per word raises all sorts of questions about how such an arrangement might be established, questions that would take us far enough away from the current experiments that we will refrain from discussing them further here.

If, as we believe, a position-item association model is unsuited to the modelling of the Hebb effect and, ultimately, of vocabulary learning, what sort of system might better fulfil these functions. Our view is that the Hebb effect proceeds by combining list items together into chunks (c.f. Miller, 1956), maybe of size equal to that of the list itself, but more likely smaller. The repetition of a list aids in the formation of chunks comprising subsequences of the list itself, the longer list subsequently being remembered as a chunk or, more likely, a list of chunks rather than as a long list of individual items. Memory for lists of pre-established chunks, such as the list FBIPHDUSA, is more accurate than for lists such as AFBIPHDUS, which has virtually the same item-to-item transitions but with the chunks broken up or "disguised" (Bower & Springston, 1970). In our view, the Hebb effect involves the establishment of new chunks in LTM, which can be used to expedite the recall of a long letter sequence.

Such an account of the Hebb effect is entirely consistent with the data presented here. When a transfer list is derived from the Hebb list by leaving only alternate items in place (e.g., transforming the list 0123456789 to 0529476381) the chunks out of which the list is comprised are completely changed. One would expect, therefore, that repeated presentation of the former list would not assist in recall of the latter. The experiments presented above confirm this expectation.

As was noted earlier in this paper, the primacy model of immediate serial recall was derived from previous work into the long-term memory of item sequences (Nigrin, 1993; Page, 1993; 1994). In this earlier work, long-term memory for short item sequences or chunks was implemented using an unsupervised learning mechanism that constructed localist representations of those subsequences made familiar by repetition. Briefly, these localist representations of short item-sequences were constructed using the primacy gradient in short-term memory in such a way that the connection between the (connectionist) node representing a list item and a node corresponding to a chunk of which it forms part, is stronger the earlier in the chunk the item appears. Thus for the learned chunk ABC, the connection to the chunk node from the node representing item A (henceforth, the A-node) would be stronger than that from the B-node which, in turn, would be stronger than that from the C-node. Thus, there would be a primacy gradient in connection weights that would mirror, and would be learned via, the proposed primacy gradient in short-term memory activations. Once such a chunk node is established, it is deemed to activate best when its items activate in an order consistent with the primacy gradient in long-term weights. Thus the ABC-node is activated best by the sequence ABC and worse by, for example, the sequences ACB, BAC, etc.. This sensitivity to correct order can be established in various ways (see Nigrin, 1993; Page, 1993; 1994), some developments of which are the subject of current work. Suffice to say here that the reactivation of learned chunks during the presentation of a familiarised list can be of assistance in correct recall of that list, just as previously learned knowledge can assist in the recall of the list FBIPHDUSA, as noted above. Clearly, if sequence chunks are represented in such a way that their reactivation depends tightly on the correct items' arriving in the correct order, then any experimental manipulation of a familiarised Hebb list that leaves alternate items in position while randomly placing other items, would not be expected to yield an increased level of performance relative to filler controls.

The data we have presented here are a challenge for models that seek to account for the Hebb effect in terms of a strengthening of position-item associations and at least consistent with a model based on the establishment of order-sensitive chunks. Further work will investigate whether Hebb-effect learning is a good analogue of the learning of phonological word forms. If it proves to be so, this will help place experimental and theoretical work relating to the phonological loop in a broader and more ecologically valid context.

REFERENCES

Anderson, J. R., Bothell, D., Lebiere, C., & Matessa, M. (1998). An integrated theory of list memory. Journal of Memory and Language, **38**(4), 341-380.

Baddeley, A. (1993). Short-term phonological memory and long-term learning - a single-case study. European Journal of Cognitive Psychology, **5**(2), 129-148.

Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. Psychological review, **105**(1), 158-173.

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.

Bower, G. H., & Springston, F. (1970). Pauses as recoding points in letter series. Journal of Experimental Psychology, **83**(3), 421-430.

Brown, G. D. A., Preece, T., & Hulme, C. (2000). Oscillator-based memory for serial order. Psychological Review, **107**(1), 127-181.

Burgess, N., & Hitch, G. J. (1992). Toward a network model of the articulatory loop. Journal of Memory and Language, **31**(4), 429-460.

Burgess, N., & Hitch, G. J. (1996). A connectionist model of STM for serial order. In S. E. Gathercole (Ed.), Models of Short Term-Memory (pp. 51-72). Hove: Psychology Press.

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.

Cohen, R. L., & Johansson (1967). The activity trace in immediate memory: A re-evaluation. Journal of Verbal Memory and verbal Behavior, **6**, 139-143.

Cumming, N. (2001). The Hebb effect: Investigating long-term learning from short-term memory. Unpublished Doctoral Thesis, Cambridge University, Cambridge.

Cunningham, T. F., Healy, A. F., & Williams, D. M. (1984). Effects of repetition on short-term retention of order information. Journal of Experimental Psychology:Learning, Memory and Cognition, **10**(4), 575-597.

Ebbinghaus, H. (1885/1964). Memory: A contribution to experimental psychology. New York: Dover. (Originally published 1885, translated 1913).

Estes, W. K. (1972). An associative basis for coding and organization in memory. In A. Melton & E. Martin (Eds.), Coding processes in human memory (pp. 161-190). New York: Halstead Press.

Fendrich, D. W., Healy, A. F., & Bourne, L. E. (1991). Long-term repetition effects for motoric and perceptual procedures. Journal of Experimental Psychology: Learning memory and Cognition, **17**(1), 137-151.

Gathercole, S. E. (1995). Is nonword repetition a test of phonological memory or long-term knowledge - it all depends on the nonwords. Memory & Cognition, **23**(1), 83-94.

Gathercole, S. E., Service, E., Hitch, G. J., Adams, A.-M., & Martin, A. J. (1999). Phonological short-term memory and vocabulary development: further evidentce on the natyre of the relationship. Applied Cognitive Psychology, **13**, 65-77.

Gathercole, S. E., Willis, C., & Baddeley, A. D. (1991). Nonword repetition, phonological memory, and vocabulary - a reply to snowling, chiat, and hulme. Applied Psycholinguistics, **12**(3), 375-379.

Grossberg, S. (1978a). Behavioral contrast in short term memory: Serial binary memory models or parallel continuous memory models? Journal-of-Mathematical-Psychology, **17**(3), 199-219.

Grossberg, S. (1978b). A theory of Human memory: Self organization and performance of sensory-motor codes, maps and plans. In R. Roden & F. Snell (Eds.), Progress in theoretical biology (Vol. 5,). New York: Academic Press.

Hebb, D. O. (1949). The Organization of Behavior. New York: Wiley.

Hebb, D. O. (1961). Distinctive features of learning in the higher animal. In J. F. Delafresnaye (Ed.), Brain mechanisms and learning (pp. 37-46). Oxford: Blackwell.

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

Henson, R. N. A. (1998). Short-term memory for serial order: The start-end model. Cognitive Psychology, **36**(2), 73-137.

Keppel, G., & Saufley, W. H. (1964). Serial position as a stimulus in serial learning. Journal of Verbal Learning and Verbal Behaviour, **3**, 335-343.

Lashley, K. A. (1951). The problem of serial order in behaviour. In L. A. Jeffress (Ed.), Cerebral mechanisms in behaviour (pp. 112-136). New York: Wiley.

Lee, C. L., & Estes, W. K. (1977). Order and position in primary memory for letter strings. Journal of Verbal Learning and Verbal Memory, **16**, 395-418.

Lee, C. L., & Estes, W. K. (1981). Item and order information in short-term memory: evidence for multilevel perturbation processes. Journal of Experimental Psychology: Human Learning and Memory, **7**, 149-169.

Lewandowsky, S., & Murock, B. B. (1989). Memory for serial order. Psychological Review, **96**(1), 25-57.

Li, S. C., & Lewandowsky, S. (1995). Forward and backward recall - different retrieval processes. Journal of Experimental Psychology : Learning Memory and Cognition, **21**(4), 837-847.

McKelvie. (1987). Learning and awareness in the Hebb digits task. The Journal of General psychology, 114(1), 75-88.

McNicol, D. (1978). Initial codes for order in memory. Quarterly Journal of Experimental Psychology, **30**, 681-691.

Melton, A. W. (1963). Implications of short-term memory for a general theory of memory. Journal of Verbal Learning and Verbal Behavior, **2**, 1-21.

Miller, G. A. (1956). The magical number seven, plus or minus two, some limits on our capacity for processing information. Psychological Review, **63**(2), 81-97.

Murdock, B. (1996). Item, associative and serial order information in TODAM. In S. E. Gathercole (Ed.), Models of short-term memory (pp. 239-266). Hove: Psychology Press.

Murdock, B. B. (1992). A chunking model for serial organization. International Journal of Psychology, **27**(3-4), 103.

Murdock, B. B. (1993). Derivations for the chunking model. Journal of Mathematical Psychology, **37**(3), 421-445.

Murdock, B. B. (1995). Developing todam - 3 models for serial-order information. Memory & Cognition, **23**(5), 631-645.

Nairne, J. S., & Neath, I. (1994). Critique of the retrieval/deblurring assumptions of the Theory of Distributed Associative Memory. Psychological Review, **101**(3), 528-533.

Nigrin, A. L. (1993). Neural networks for pattern recognition. Cambridge, Massachusetts: MIT Press.

Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. Cognitive Psychology, 19, 1-32.

Page, M. P. A. (1993). Modelling aspects of music perception using self-organizing neural networks, University of Wales.

Page, M. P. A. (1994). Modelling the perception of musical sequences with self-organizing neural networks. Connection Science, **6**, 223-246.

Page, M. P. A., & Norris, D. (1998). The Primacy Model: a new model of immediate serial recall. Psychological Review, **105**(4), 761-781.

Papagno, C. (1996). Contribution of the phonological loop to phonological learning of new vocabulary. Perceptual and Motor Skills, **82**(3 Pt1), 769-770.

Papagno, C., Valentine, T., & Baddeley, A. (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. The Quarterly Journal of Experimental Psychology, **44A**(1), 47-67.

Schwartz, M., & Bryden, M. P. (1971). Coding factors in the learning of repeated digit sequences. Journal of Experimental Psychology, **87**(3), 331-334.

Seger, C. A. (1994). Implicit Learning. Psychological Bulletin, **115**(2), 163-196.

Shallice, T., & Vallar, G. (1990). Impairment of auditory-verbal short-term storage. In G. Vallar & T. Shallice (Eds.), Neuropsychological impairments of short-term memory (pp. 11-53). Cambridge: Cambridge University Press.

Slamecka, N. J. (1964). An inquiry into the doctrine of remote associations. Psychological Review, **71**(1), 61-67.

Slamecka, N. J. (1967). Serial learning and order information. Journal of Experimental Psychology, **74**(1), 62-66.

Stadler, M. A. (1992). Statistical structure and implicit serial-learning. Journal of Experimental Psychology-Learning Memory and Cognition, 18(2), 318-327.

Trojano, L., Stanzione, M., & Grossi, D. (1992). Short-term-memory and verbal-learning with auditory phonological coding defect - a neuropsychological case-study. Brain and Cognition, **18**(1), 12-33.

Vallar, G., & Baddeley, A. D. (1984a). Fractionation of working memory - neuropsychological evidence for a phonological short-term store. Journal of Verbal Learning and Verbal Behavior, **23**(2), 151-161.

Vallar, G., & Baddeley, A. D. (1984b). Phonological short-term store, phonological processing and sentence comprehension - a neuropsychological case-study. Cognitive Neuropsychology, **1**(2), 121-141.

Vallar, G., DeBetta, A. M., & Silveri, M. C. (1997). The phonological short-term store rehearsal system: Patterns of impairment and neural correlates. Neuropsychologia, **35**(6), 795-812.

Vallar, G., Papagno, C., & Baddeley, A. D. (1991). Long-term recency effects and phonological short-term memory. A neuropsychological case study. Cortex, **27**(323-326).

Warrington, E. K., & Shallice, T. (1969). The selective impairment of auditory verbal short-term memory. Brain, **92**, 885-896.

Young, R. K. (1961). The stimulus in serial learning. American Journal of Psychology, **74**, 517-528.

Young, R. K. (1962). Tests of three hypotheses about the effective stimulus in serial learning. Journal of Experimental Psychology, **63**(3), 307-313.

AUTHOR NOTES

The authors would like to thank Neil Burgess and Graham Hitch for allowing access to the source code for their model.

TABLE CAPTIONS

Table 1: Mean number of errors per list split into error types (standard deviations shown in parentheses). Note S-items and D-items are reconstructed “lists” of items taken from alternate blocks of the experiment.

Table 2: Mean number of errors per list split into error types (standard deviations shown in parentheses), Experiment 2.

FIGURE CAPTIONS

Figure 1: An example transfer list ordering. The top row of circles represents the order of the items in the Hebb list, while the bottom row is the relative order of the transfer-list items.

Figure 2: Top panel: progression of experimental trials used in experiment 1. For simplicity, only two experimental blocks are shown here. As can be seen, the transfer lists are presented on the trial immediately following the eighth presentation of each blocks' Hebb list. Bottom panel: trial-type progression used in experiment 2. In this experiment, the transfer list is 4 lists after the eighth presentation of the Hebb list.

Figure 3: Error types during Hebb list learning, Experiment 1.

Figure 4: Serial position curves for all list types, experiment 1. Note the slight zig-zag pattern in the early portion of the transfer lists.

Figure 5: Protrusion distribution for filler lists following Hebb lists and following other filler lists, expressed as a normalised proportion of times an erroneous response appeared at a distance of x away in the previous list from its position in the current list.

Figure 6: Order and omission errors by Hebb list repetition for both Hebb lists and error lists, Experiment 2.

Figure 7: Serial position curves for all list types in experiment 2. Transfer lists are shown separately depending on whether they were SDSD or DSDS type.

Figure 8: Diagram of the Burgess and Hitch (1999) model. Heavy lines show full interconnectivity with modifiable weights. Narrow lines indicate one-to-one mapping with no alterable weights.

Figure 9: Performance of the optimised gradient performance model.

Figure 10: Top panel: Serial position curves for the range of parameter values investigated here (ΔWlt = 0.011 to 0.03). Serial position curves of the filler lists, transfer lists and Hebb lists (8 simulated repetitions)

Item type	Error type	
	omission	order
Hebb	1.156 (1.151)	1.729 (1.572)
filler	2.146 (1.623)	2.365 (1.414)
same	2.480 (1.529)	1.854 (1.211)
different	2.229 (1.602)	2.542 (1.421)

Table 1: Mean number of errors per list split into error types (standard deviations shown in parentheses), Experiment 1. Note S-items and D-items are reconstructed "lists" of items taken from alternate blocks of the experiment.

Item type	Error type	
	omissions	order
Hebb	1.431 (1.173)	1.369 (1.141)
Control	1.725 (1.479)	1.400 (1.259)
Same	2.700 (1.358)	2.175 (1.542)
Different	2.725 (1.266)	2.188 (1.492)
Filler	2.938 (1.174)	2.156 (1.345)

Table 2: Mean number of errors per list split into error types (standard deviations shown in parentheses), Experiment 2.

Significance levels for order error comparisons.

	S-item	D-item	H-item	F-item
S-item	//////	0.001	n.s.	0.05
D-item		//////	0.05	n.s.
H-item			//////	0.01
F-item				//////

Table 2 : Comparisons of number of order errors made on each item type.

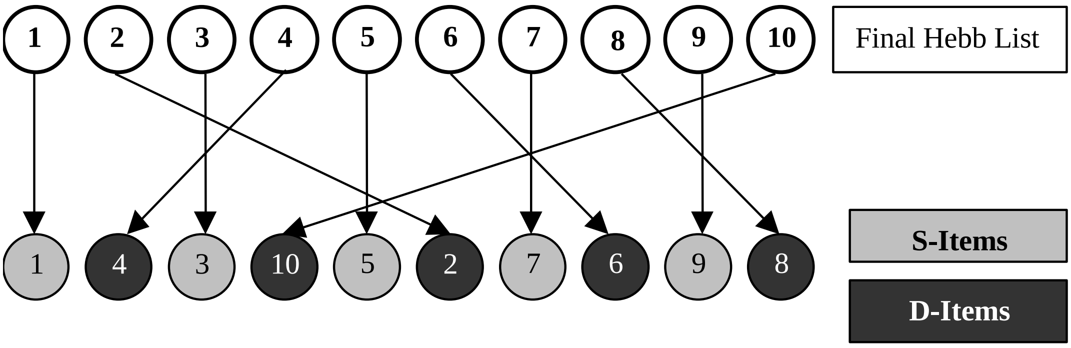
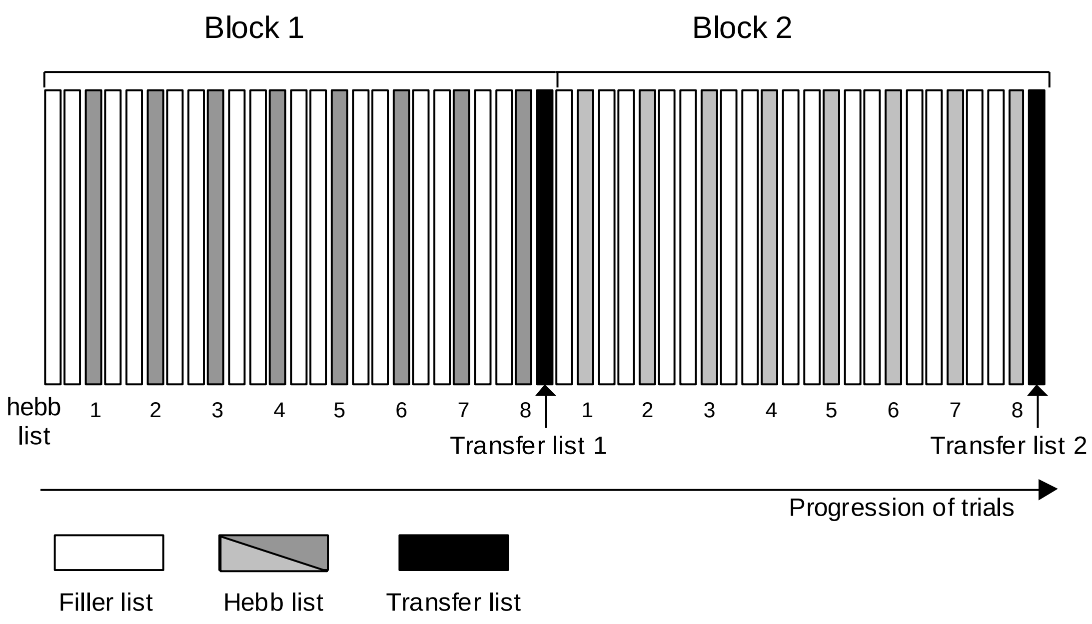

1 2 3 4 5 6 7 8 9 10
Final Hebb List
1 4 3 10 5 2 7 6 9 8
S-Items
D-Items

Figure 1: An example transfer list ordering. The top row of circles represents the order of the items in the Hebb list, while the bottom row is the relative order of the transfer-list items.


Block 1
Block 2
hebb list
1
2
3
4
5
6
7
8
Transfer list 1
1
2
3
4
5
6
7
8
Transfer list 2
Progression of trials
Filler list
Hebb list
Transfer list

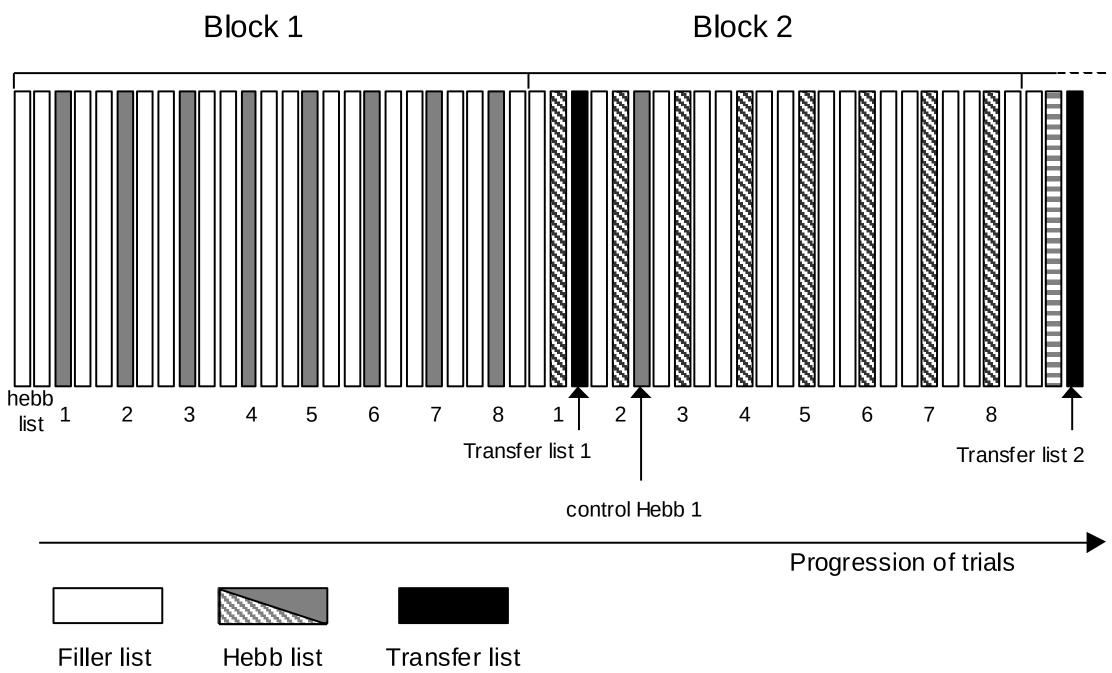
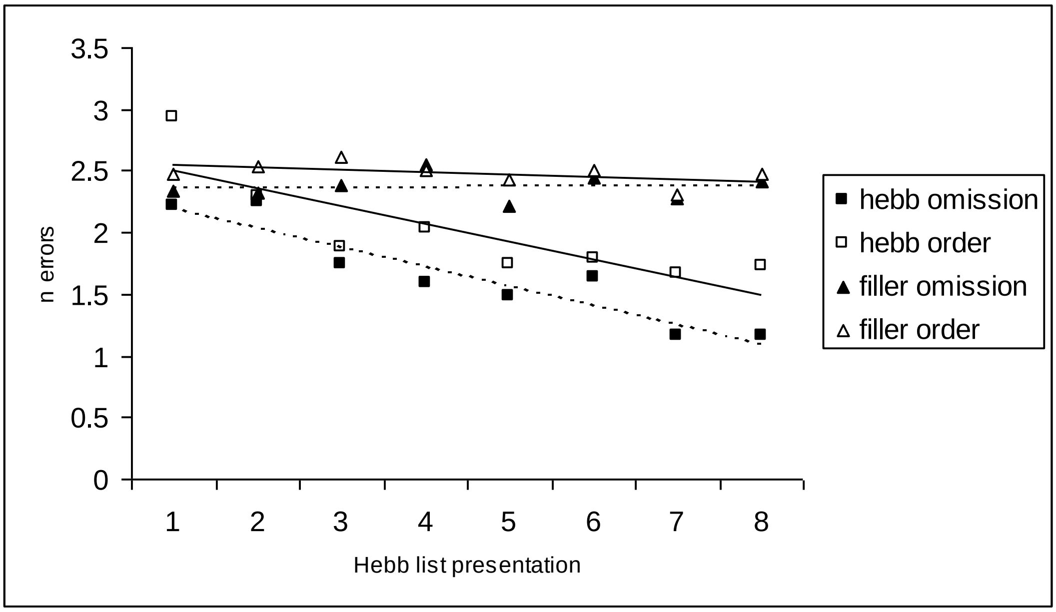

Block 1
Block 2
hebb list
1
2
3
4
5
6
7
8
1
Transfer list 1
2
control Hebb 1
3
4
5
6
7
8
Transfer list 2
Progression of trials
Filler list
Hebb list
Transfer list

Figure 2: Top panel: progression of experimental trials used in experiment 1. For simplicity, only two experimental blocks are shown here. As can be seen, the transfer lists are presented on the trial immediately following the eighth presentation of each blocks' Hebb list. Bottom panel: trial-type progression used in experiment 2. In this experiment, the transfer list is 4 lists after the eighth presentation of the Hebb list.



3.5
3
2.5
2
1.5
1
0.5
0
n errors
1
2
3
4
5
6
7
8
Hebb list presentation
hebb omission
hebb order
filler omission
filler order

Figure 3: Error types during Hebb list learning, Experiment 1.

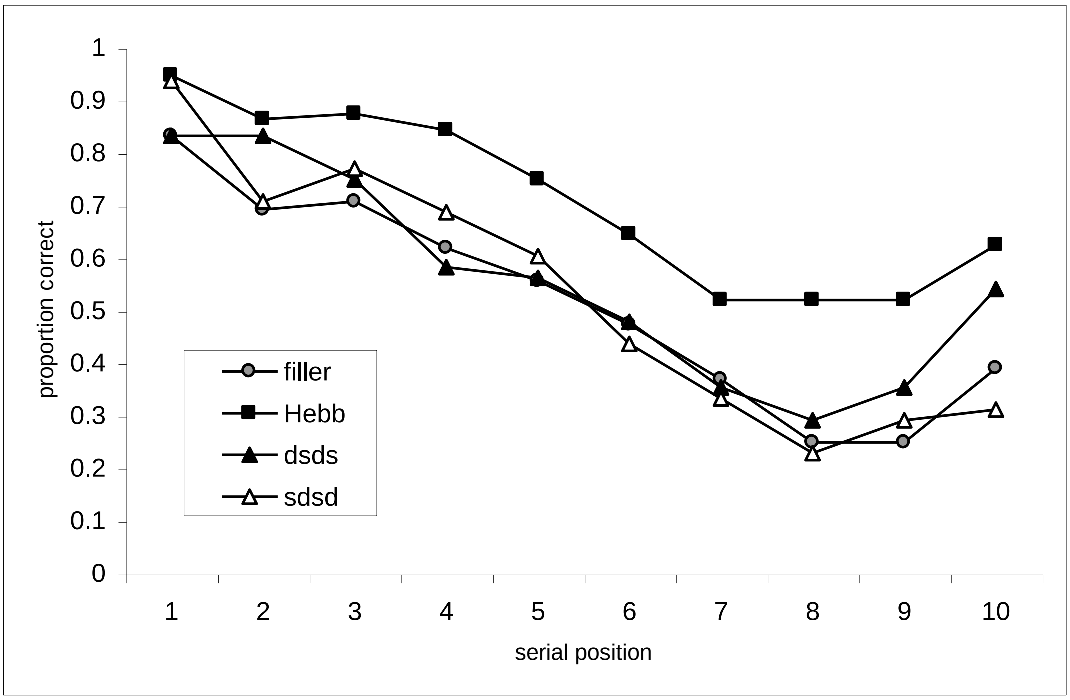

proportion correct
serial position
filler
Hebb
dsds
sdsd
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1 2 3 4 5 6 7 8 9 10

Figure 4: Serial position curves for all list types, experiment 1. Note the slight zig-zag pattern in the early portion of the transfer lists.

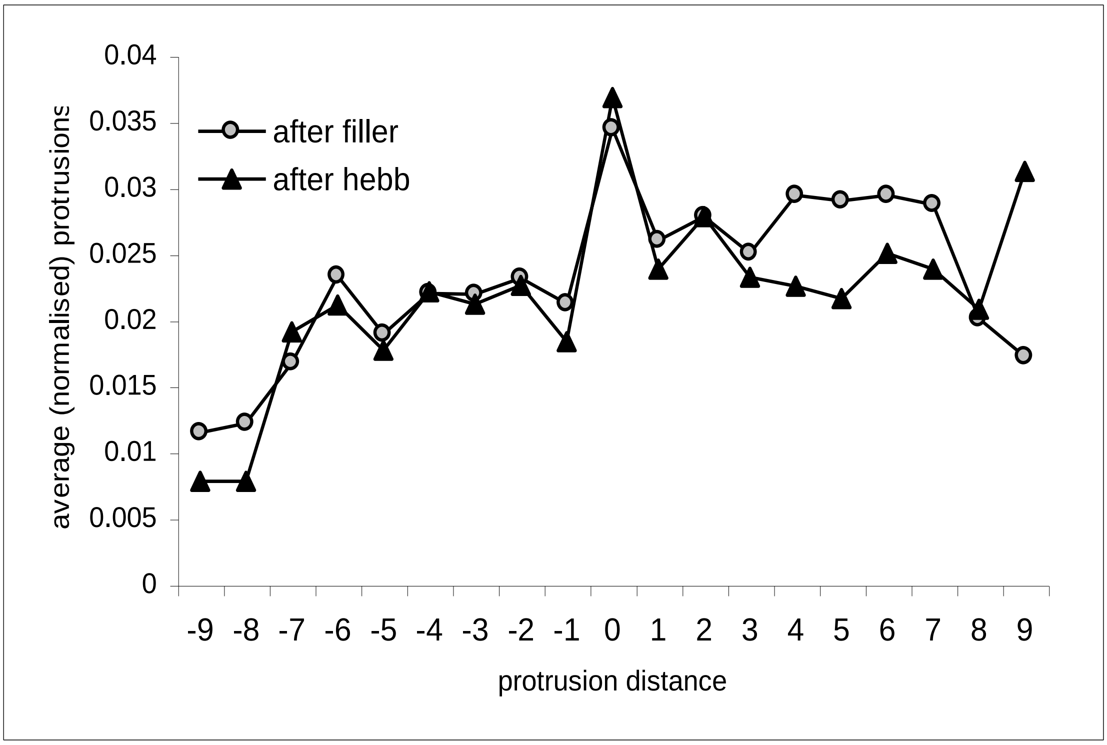

0.04
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
average (normalised) protrusions
after filler
after hebb
-9 -8 -7 -6 -5 -4 -3 -2 -1 0 1 2 3 4 5 6 7 8 9
protrusion distance

Figure 5: Protrusion distribution for filler lists following Hebb lists and following other filler lists, expressed as a normalised proportion of times an erroneous response appeared at a distance of x away in the previous list from its position in the current list.

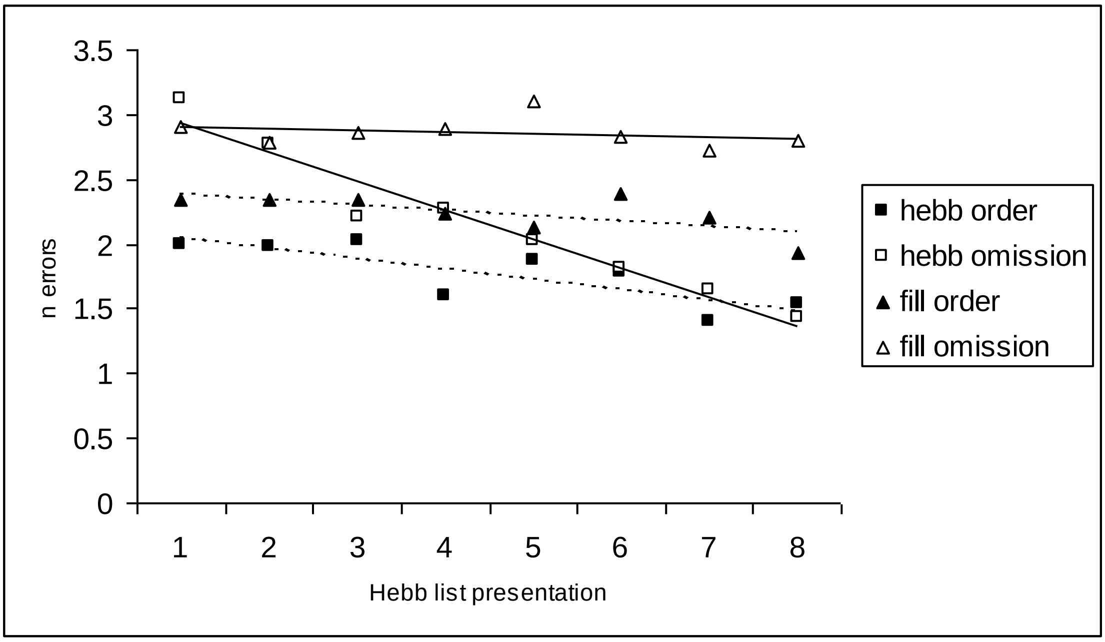
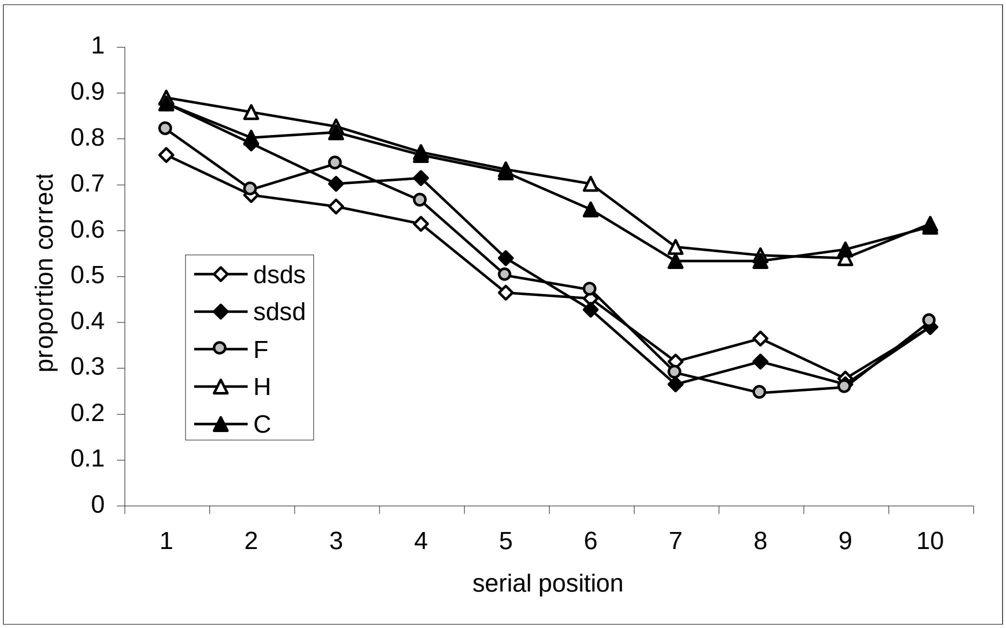

3.5
3
2.5
2
1.5
1
0.5
0
n errors
1
2
3
4
5
6
7
8
Hebb list presentation
hebb order
hebb omission
fill order
fill omission

Figure 6: Order and omission errors by Hebb list repetition for both Hebb lists and error lists, Experiment 2.



proportion correct
serial position
dsds
sdsd
F
H
C
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1 2 3 4 5 6 7 8 9 10

Figure 7: Serial position curves for all list types in experiment 2. Transfer lists are shown separately depending on whether they were SDSD or DSDS type.

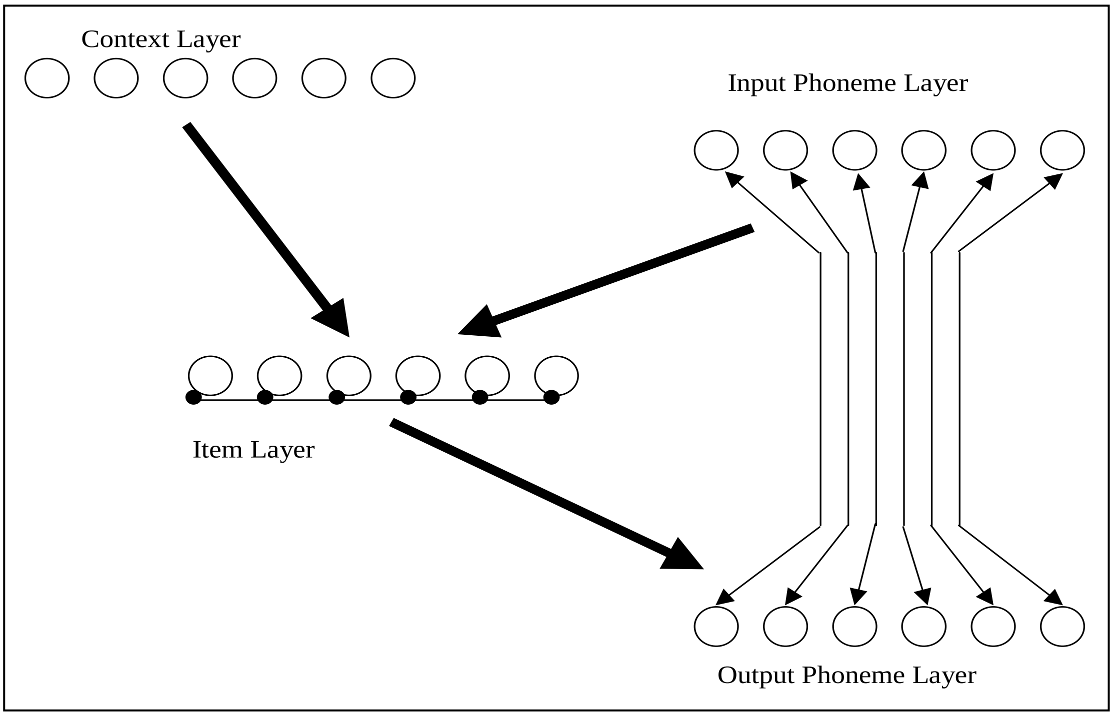

Context Layer
Input Phoneme Layer
Item Layer
Output Phoneme Layer

Figure 8: Diagram of the Burgess and Hitch (1999) model. Heavy lines show full interconnectivity with modifiable weights. Narrow lines indicate one-to-one mapping with no alterable weights.

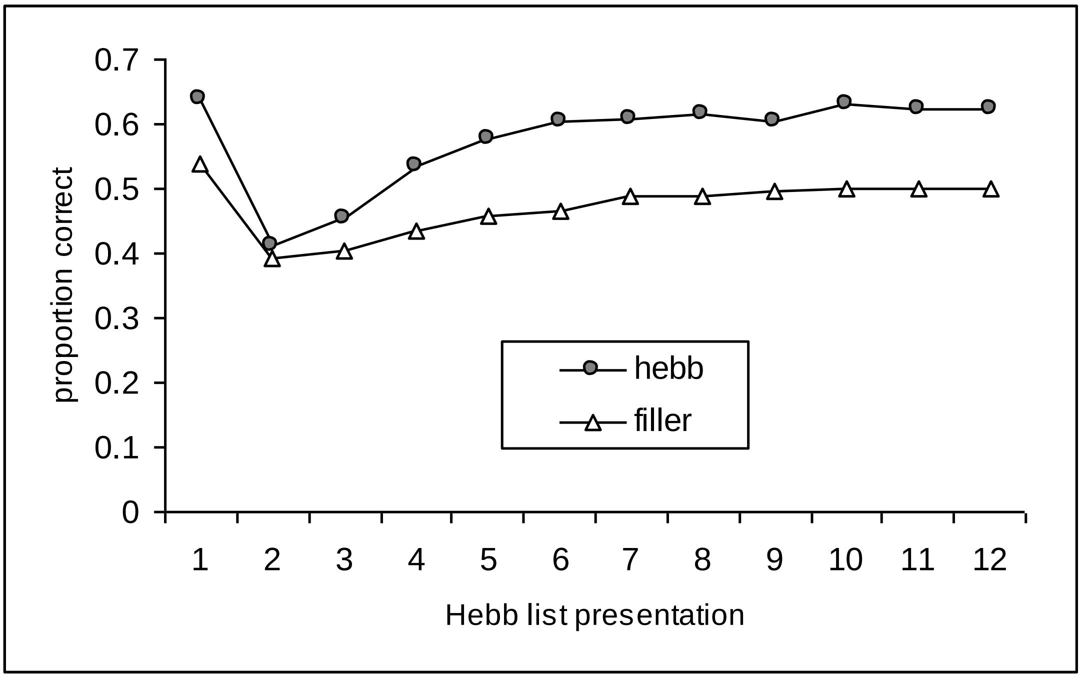

0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
proportion correct
hebb
filler
1
2
3
4
5
6
7
8
9
10
11
12
Hebb list presentation

Figure 9: Performance of the optimised gradient performance model.

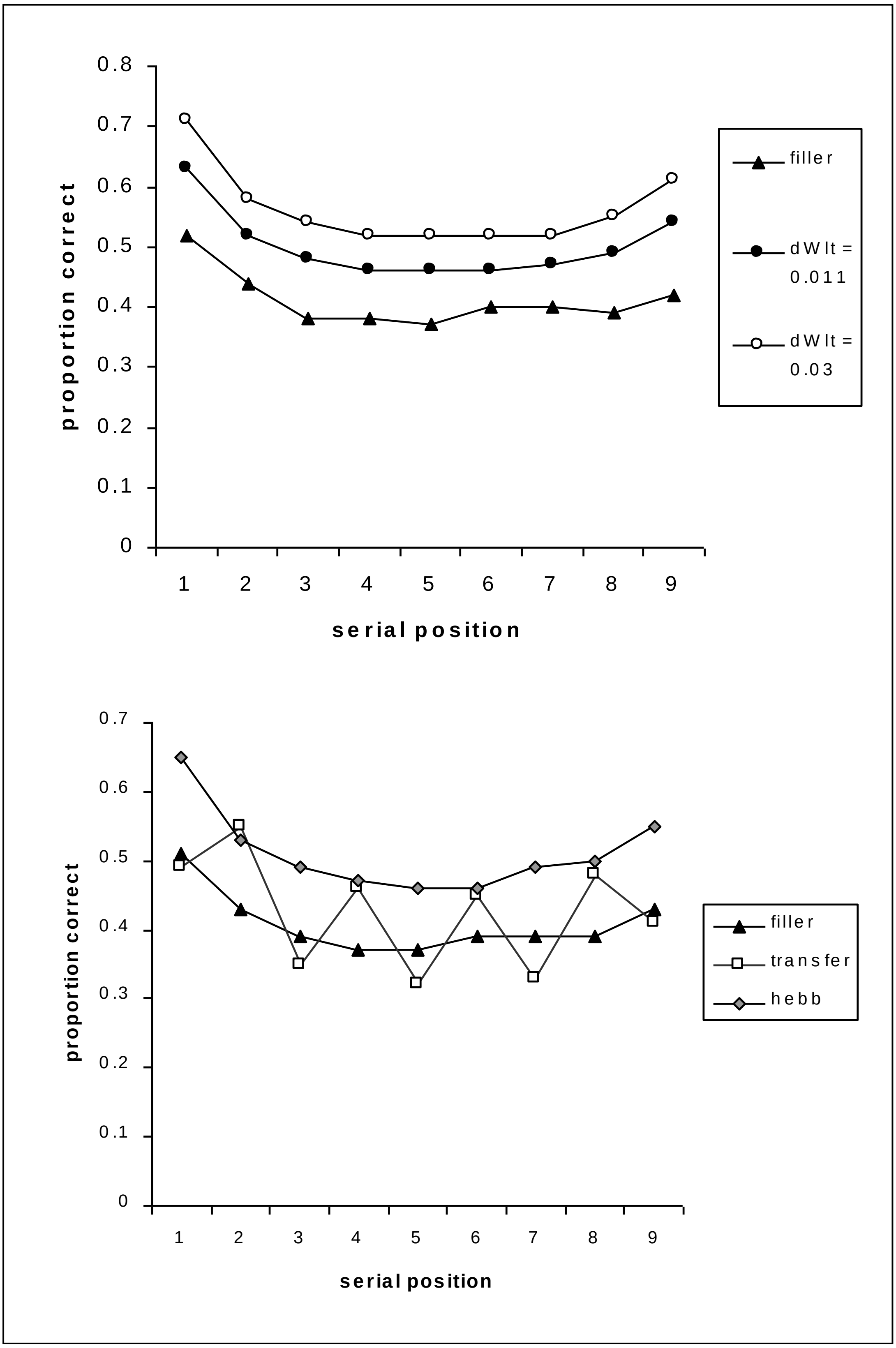

0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
proportion correct
1 2 3 4 5 6 7 8 9
serial position
filler
dWlt = 0.011
dWlt = 0.03
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
proportion correct
1 2 3 4 5 6 7 8 9
serial position
filler
transfer
hebb

Figure 10: Top panel: Serial position curves for the range of parameter values investigated here ($\Delta Wlt = 0.011$ to 0.03). Serial position curves of the filler lists, transfer lists and Hebb lists (8 simulated repetitions)