

Serial Position Functions for Recognition of Olfactory Stimuli

Andrew J. Johnson and Christopher Miles

School of Psychology
Cardiff University
UK

Address for Correspondence: Christopher Miles
School of Psychology
Tower Building
Park Place
Cardiff
UK
CF10 3AT

Email: Miles@cardiff.ac.uk

Acknowledgements

We wish to thank Helen Beacham for her enthusiastic assistance in testing participants. This study was supported by a joint Medical Research Council and School of Psychology, Cardiff University postgraduate studentship awarded to the first author.

Abstract

Two experiments examined item recognition memory for sequentially presented odours. Following a sequence of six odours participants were immediately presented with a series of 2-alternative forced choice (2AFC) test odours. The test pairs were presented in either the same order as learning or the reverse order of learning. Method of testing was either blocked (Experiment 1) or mixed (Experiment 2). Both experiments demonstrated extended recency, with an absence of primacy, for the reverse testing procedure. In contrast, the forward testing procedure revealed a null effect of serial position. The finding of extended recency is inconsistent with the single-item recency predicted by the two-component duplex theory (Phillips and Christie, 1977). We offer an alternative account of the data in which recognition accuracy is better accommodated by the cumulative number of items presented between item learning and item test.

Keywords: Olfactory short-term memory; Recognition; Serial position function;

Introduction

Recognition of hard-to-name stimuli, e.g. unfamiliar faces, snow flakes, visual matrices is traditionally tested via a 2-alternative forced-choice (2AFC) recognition paradigm. In this paradigm a sequence of items is presented to the participant followed by a test probe comprising one item from the previous sequence paired with a novel item. The participant is required to judge which of the two test items was presented in the previous sequence. Unlike the bowed serial position function traditionally reported for the free recall of verbal stimuli (e.g. Glanzer and Cunitz, 1966; Craik, 1968; Ellis and Hope, 1968), the serial position function for recognition of these hard-to-name stimuli is typically characterised by single-item recency, with equivalent, but reduced recognition rates for pre-recency items e.g. Ward, Avons and Melling (2005).

This pattern of recognition was established originally by Phillips and Christie (1977) who presented participants with a sequence of patterns constructed within a 4x4 visual matrix. The sequence was followed by a series of 2AFC recognition test-pairs and testing was conducted in the reverse order of original presentation. This produced enhanced recognition for the final item in the sequence, with equivalent, but reduced performance for pre-recency items. The robustness of this particular pattern of data is evidenced by a number of replications, for example, with coloured light patches (Avons and Daley, 1990) and unfamiliar faces (Ward et al, 2005). This single-item recency effect has traditionally been interpreted via the two-component duplex theory, originally proposed by Phillips and Christie (1977). The theory proposes that attention is allocated to each item when presented serially within a sequence. When a new item in the sequence is presented attention is shifted from the previous item to the encoding of the new item. The previous item is then displaced from this accurate, but fragile, short-term attentional store and enters a more durable, but less accurate, long-term store. Phillips and Christie (1977) proposed that the recency effect reflects the immediate testing of the last item in the sequence whilst this item is held within the short-term store. With reverse testing, intervening items separate presentation and test of all pre-recency items; this causes displacement from the short-term store into the long-term store producing reduced, but consistent, recognition rates for pre-recency items.

However, the results of a number of more recent studies are difficult to accommodate within the duplex framework. For instance, Avons, Ward and Melling (2004, Experiment 1) employed a single 2AFC recognition test for sequences of 5-visual matrices and demonstrated evidence of a linear function with orthogonal contrasts showing extended recency. Such a linear trend is more consistent with a process account, such as temporal distinctiveness (e.g. Neath, 1993b), whereby memory for an item is linearly related to the time elapsed between original presentation and test.

Further evidence contradictory to the duplex account comes from Reed (2000). Groups of 10 participants were presented with sequences of odours and odour recognition was assessed via a single 2AFC task. Both primacy and recency effects were evident at both immediate testing and after a 60-second retention interval (Experiment 3). Reed argues that verbal labelling of the odours (and hence subsequent verbal rehearsal) was unlikely to underpin primacy due to (i) the low name-ability rate for the odours (31% correct name-ability) and (ii) the marginally significant main effect of articulatory suppression (Experiment 4). It is worth noting, however, that the Reed (2000) pattern of data have proved difficult to replicate. For instance, Miles and Hodder (2005) followed the methodology employed by Reed (2000) very closely. Across a series of seven experiments employing they failed to show any evidence for a primacy effect, despite mean recognition rates being broadly consistent with Reed's at approximately 70%. For two of the three experiments where serial position effects were apparent, Miles and Hodder (2005) reported some evidence of extended recency, a finding at odds with the duplex account.

Notwithstanding the differences in the pattern of findings between Reed (2000) and Miles and Hodder (2005), we argue that neither study represents an appropriate test of the duplex theory. At least two important methodological points differentiate them from the Phillips and Christie (1977) paradigm. First, both presented participants with a series of odours followed by a single 2AFC recognition test-pair. The single 2AFC test-pair probed recognition for only one of the sequence items per trial, with the order of testing for each position randomised across the experiment. This procedure contrasts with Phillips and Christie (1977) wherein each item in the sequence was tested within every trial alongside a novel item, and items were tested in the reverse order of original presentation. Second, in both Reed (2000) and Miles and Hodder

(2005) the same 8 odours were employed throughout the experiment. In contrast, Phillips and Christie (1977) employed novel patterns for each sequence.

Experiment 1

In the present study a stimulus set of 120 odours was employed in order to both limit the emergence of verbal labels through multiple presentations (verbal elaboration has been shown to aid odour recognition, e.g. Jehl, Royet and Holley, 1997) and imitate the original Phillips and Christie (1977) methodology. A sequence of 6-odours was presented followed by a series of 2AFC recognition tests, whereby each item from the previous sequence was tested alongside a novel odour. The series of 2AFC test-pairs was presented either in the original order of presentation (forward testing procedure) or in the reverse order of original presentation (backward testing procedure). Following duplex theory we predict single-item recency following the backward testing procedure since the final list-item is the only item present in the short-term store at test. For the forward testing procedure, based upon a duplex interpretation, we predict a flat serial position function, with all sequence items displaced into the long-term store.

Method.

Participants. Twenty-four (12 males, 11 females, 1 transsexual: mean age = 23 years 9 months, 18 non smokers) Cardiff University volunteer undergraduates from a range of disciplines participated. Each received a £5.00 honorarium upon completion of the study. Participants suffering from a blocked nose or cold were excluded.

Materials. One hundred and twenty non-food related odour pots supplied by Dale Air Limited, UK were utilised (see Appendix 1 for a complete list of the odours employed). Each odour was presented as a liquid soaked in cotton wool contained in a vortex cube. Each cube was blue in appearance with identical dimensions of 50mm by 50mm by 50mm. An odour name label was situated on the base of each pot. One face of the cube contained six perforations from which the odour was inhaled. The integrity of the odour within in each pot was maintained by a protective sticker placed over the perforations.

Design. A 2x6 within-subjects factorial design was adopted in a 2AFC) recognition paradigm. The first factor refers to the method of testing (forward versus backward) and the second factor refers to serial position (1-6). Method of testing (forward versus backward) was blocked and each block comprised 10 trials. Order of block presentation was counterbalanced across participants. Each of the 120 odours was presented twice: once in the forward testing procedure and once in the backward testing procedure. The order of odour presentation was randomised for each participant.

Procedure. The odour presentation procedure followed closely that reported by both Reed (2000) and Miles and Hodder (2005) and the testing procedure reflected that described for visual stimuli by both Phillips and Christie (1977) and Avons et al (2004, Experiment 1). Participants were tested individually in a well-ventilated, soundproofed laboratory and sat facing the experimenter with a fan blowing across their face. In order to minimise visual cues, the participant was instructed to fixate on a red spot located on the table 30cm in front of them throughout the trials. For each trial the participant was presented with a sequence of 6 odours. Each odour was presented over a wooden screen located 40cm in front of the participant and held under the nose of the participant for a period of 3-seconds. The participant was instructed to inhale deeply through both nostrils for the duration of each odour presentation. The odour was then replaced behind the screen during which time the participant exhaled. There was an inter-stimulus interval (ISI) of approximately 2-seconds after which the next odour was presented. This procedure continued to the presentation of the sixth odour.

A retention interval of approximately 3-seconds followed sequential presentation of the 6-odours. For the test phase the participant was presented with a series of 2AFC recognition tests, where one of the test odours was the target-probe odour taken from the previous sequence and the other was a novel odour. Both the rate of presentation of the test odours and the ISI between test-pairs were the same as those employed in the learning phase. The participant was required to state verbally whether the first or the second odour in the test-pair was familiar from the previous sequence by responding “first” or “second”. In the forward testing procedure the target odour in the first test-pair presented was the odour presented first in the previous sequence.

This procedure was repeated with the second test-pair which comprised the odour presented second in the previous sequence and a novel odour. This pattern of testing continued until each odour in the sequence had been tested against a novel odour. The order of testing was, therefore, identical to the order of presentation. The backward testing procedure followed that described for the forward testing procedure, with the exception that the sequence of test-pairs tracked backwards through the sequence previously presented. Thus, the first test-pair presented comprised the odour presented last in the preceding sequence paired with a novel odour. The position of the target odour (first or second) within test-pairs was randomly assigned with the proviso that it occurred an equal number of times in each position and that there was a maximum of two consecutive trials in which the position of the target odour was unchanged. Each trial was followed by an interval of approximately 12 seconds and the participant was given the option of a 2 minute rest after every 5 trials. The complete experiment lasted approximately one hour.

Results and Discussion.

Figure 1 shows the mean percentage correct recognition at each serial position for both the forward and the backward testing procedures. A 2-factor (2x6) within-subjects Analysis of Variance (ANOVA) was computed on the correct recognition scores where the first factor represents test procedure (forward versus backward) and the second factor represents serial position (1-6). A rejection criterion of $p < 0.05$ was adopted for this and all subsequent analyses. The ANOVA revealed a null effect of test procedure, $F < 1$, a main effect of serial position, $F(5,23)=2.88$, $MSe=2.07$, and, more importantly in the present context, the predicted interaction between test procedure and serial position, $F(5,115)=2.53$, $Mse=2.37$.

Figure 1 about here please

Further analysis of the interaction (Newman Keuls; $p < .05$) showed that for the forward testing procedure there were no significant differences in recognition rates between serial positions. In contrast, for the backward testing procedure, recognition at serial position six was significantly higher than that at serial positions one, two and four. In addition, recognition at serial position five was significantly higher than that at serial position one; indicating extended recency. Comparisons between the two

testing procedures at each serial position revealed a significant difference only at serial position six; recognition was significantly higher for the reverse testing procedure.

Following both Reed (2000) and Miles and Hodder (2005), trend analyses were computed. Trend analysis allows the functional relationship between serial position and correct recognition rates to be assessed. In contrast to computing comparisons between mean scores, trend analysis allows consideration of all the mean scores within a test procedure, focussing upon the overall shape or trend of the results (Keppel and Wickens, 2004). Trend analysis of the forward testing procedure showed both the linear and quadratic components to be absent, both $F_s < 1$. In contrast, trend analysis of the backward testing procedure revealed a significant linear component, $F(1,23)=17.28$, $Mse=2.72$ but a non-significant quadratic component, $F < 1$. The linear trend and extended recency evident following the backward testing procedure is contrary to the single-item recency predicted by the duplex account. Taken together these data offer no support for a duplex interpretation of 2AFC recognition for odour sequences.

Experiment 1 employed a blocked testing procedure, whereby participants received blocks of both forward and backward testing procedure trials. It is possible that in the backward testing procedure, participants learned that the items in the latter part of the sequence were always tested first and that such learning benefited recognition of those items. To test this possibility, an analysis comparing first half (mean recognition = 81.7%) and second half (mean recognition = 81.7%) last-item recognition accuracy for the backward testing procedure was computed and indicated no evidence of learning ($t(23)=0$, $P > 0.05$).

Experiment 2

In order to rule out learning directly, Experiment 2 replicated the previous experiment but adopted a mixed-order design. Throughout the experiment participants received both forward and backward test trials at random with the proviso that the same testing procedure was employed on no more than two consecutive trials. Such a design was devised to minimize the development of encoding strategies whereby participants learned to predict the particular (forward or backward) testing procedure.

Method.

Participants. Twenty-four (10 males, 14 females: mean age = 19 years and 11 months, 20 non-smokers) Cardiff University volunteer undergraduates from a range of disciplines participated and each received a £5.00 honorarium upon completion of the study. None had participated in Experiment 1. Participants suffering from a blocked nose or cold were excluded.

Materials. The materials were as those described for Experiment 1.

Design. The design was as that described for Experiment 1 with the exception that the forward and backward testing procedures were mixed within blocks, with the proviso that trials with the same direction of testing did not exceed two in succession. Two orders of trial presentation were employed and counterbalanced across participants.

Procedure. The procedure was as described for Experiment 1.

Results and Discussion.

Figure 2 shows the mean percentage correct recognition at each serial position in both the forward and the backward testing procedures. A 2-factor (2x6) within-subjects ANOVA was computed on the correct recognition scores and revealed a null effect of test procedure, $F < 1$, and a main effect of serial position, $F(5,23) = 2.34$, $MSe = 1.62$.

Figure 2 about here please

The predicted interaction between serial position and testing procedure failed to achieve significance, $F = 1.57$. Nevertheless, the pattern of correct recognition observed for Experiment 2 follows closely that observed for Experiment 1, that is, extended recency in the absence of primacy following the backward testing procedure. Further, trend analysis of the backward testing procedure revealed both linear and quadratic components to be significant, $F(1,23) = 5.84$, $MSe = 1.94$. In contrast, trend analysis of the forward testing procedure demonstrated both the linear and quadratic components to be absent (both $F_s < 1$). Thus, the data for Experiment 2

demonstrate that the extended recency effect observed in Experiment 1 for the backward testing procedure was not an artefact of the blocked design.

To check further that the effect of testing procedure was not a product of blocking trials, the data for Experiment 1 were combined with those for Experiment 2. Differential effects of blocking trials should be evidenced by an interaction between experiment, test procedure and serial position. A 3-factor (2x2x6) mixed design ANOVA (experiment × testing procedure × serial position) was computed on the correct recognition scores. The effect of experiment was non-significant, $F=1.46$, confirming the equivalent rates of recognition for both experiments (mean recognition rates were 71.6% and 68.6% for Experiments 1 and 2, respectively). Critically, the interaction between experiment, test procedure and serial position was non-significant, $F<1$.

Discussion

The two experiments were designed to provide an appropriate test of the duplex account of recognition memory for sequences of olfactory stimuli. Emulating the original Phillips and Christie (1977) methodology, participants received both novel items within each trial and a series of 2AFC probes at test. The finding of a linear function following backward testing contradicts the single-item recency function predicted by the duplex account. A duplex interpretation predicts equivalent, yet reduced, recognition rates for pre-recency items following transfer to the long term store. However, the present data indicate a general decline in recognition rates for pre-recency items coupled with evidence of extended recency. This pattern of data is, therefore, more consistent with process accounts of short-term memory (e.g. temporal distinctiveness, Neath, 1993b), whereby, in an equally distributed list, recently presented items are more distinct and, therefore, recognised with greater accuracy. However, the veridicality of temporal distinctiveness accounts has been questioned by Lewandowsky, Brown, Wright and Nimmo (2006). They demonstrated that increasing temporal isolation of sequence items i.e. increasing the inter-stimulus-interval and, thereby, the temporal distinctiveness of an item, did not, of itself, improve item recall. Rather, the authors proposed that output interference during recall influences the serial position function. Indeed, Lewandowsky (personal communication) has proposed that the data from the current experiments might be better explained in

terms of lag (defined as the number of intervening items occurring between presentation and test of an item). Under this account, as the number of items intervening between presentation and test increases, recognition accuracy reduces linearly.

In the light of this proposal, the present data were reanalysed with the number of items intervening between presentation and test as the independent variable. For example, following the backward testing procedure the last sequence item was presented in the first test-pair and therefore had, on average, 0.5 intervening items between presentation and test. However, with the forward testing procedure, the last sequence item was presented following five test-pairs and therefore had, on average, 10.5 intervening items between presentation and test. By this account, one would predict superior recall for the last sequence item following backward testing compared to forward testing. In addition, extended recency following backward testing is predicted as the number of intervening items increases cumulatively for pre-recency items. According to this account, combining data across testing procedures (forwards and backwards testing) should produce a line of best fit which demonstrates a linear decline in performance as the number of items intervening between presentation and test increases. This hypothesis is examined in Figure 3(a. and b.) where the relationship between number of intervening items and recognition accuracy for the combined forward and backward data is plotted for both Experiments 1 and 2.

Figures 3a and 3b about here please

A strong significant negative correlation between recognition accuracy and number of intervening items is found for both Experiment 1 ($R=-0.82$, $P<0.05$) and Experiment 2 ($R=-0.70$, $P<0.05$). Following the significant Pearson's correlation a linear regression was computed and confirmed that the number of items intervening between learning and test was a strong predictor of recognition accuracy for both Experiment 1 ($R^2=0.67$; $F(1,10)=20.02$, $MSe=8.33$, $P<0.05$) and Experiment 2 ($R^2=0.50$; $F(1,10)=9.79$, $MSe=8.64$, $P<0.05$). Thus, the results for both Experiments 1 and 2 are remarkably consistent with an intervening-item based account, whereby item recognition accuracy declines linearly as the number of intervening items increases; a finding clearly at odds with a duplex interpretation.

The current explanation broadens the output interference account (Lewandowsky et al, 2006) to encompass not only self-generated items i.e. those generated during the recall process, but also those items presented after a particular item in the learning phase and those items presented as recognition probes at test. One weakness of our account in this form is that it fails to differentiate empirically whether the observed effect is driven by interference from intervening items or by elapsed time. Recently, Lewandowsky, Duncun and Brown (2004) examined directly the extent to which elapsed time (lag) is responsible for producing the serial position function. In their study, the speed at which participants recalled a list of items was manipulated and no interaction between recall time and serial position was reported, indicating that elapsed time did not influence the serial position function. Therefore, the authors propose that interference from items output during test, rather than elapsed time per se, better accommodates the observed pattern of serial position recognition.

Recency following the backward testing procedure in both Experiments 1 and 2 further substantiates the finding that olfactory 2AFC recognition is characterised by extended recency when the last sequence item is tested first (Miles and Hodder, 2005; Reed, 2000). However, consistent with Miles and Hodder (2005), the current findings fail to replicate the presence of primacy as reported by Reed (2000) following 2AFC recognition. It is proposed that the present pattern of data is best explained in terms of an account whereby the number of items intervening between presentation and test predicts item recognition (as proposed by Lewandowsky, personal communication).

References

- Avons, S.E. (1998). Serial position and item recognition of novel visual patterns. *British Journal of Psychology*, 89(2), 285-308.
- Avons, S.E. and Daley, S.A. (1990). Serial position curves for coloured light patches. *Color Research and Application*, 15(5), 291-296.
- Avons, S.E., Ward, G. and Melling, L. (2004). Item and order memory for novel visual patterns assessed by two-choice recognition. *Quarterly Journal of Experimental Psychology*, 57A(5), 865-891.
- Craik, F.I.M. (1968). Two components of free recall. *Journal of Verbal Learning and Verbal Behaviour*, 7, 996-1004.

- Ellis, N.R. and Hope, R. (1968). Memory processes and the serial position curve. *Journal of Experimental Psychology*, 77, 613-619.
- Glanzer, M. and Cunitz, A.R. (1966). Two storage mechanisms in free recall. *Journal of Verbal Learning and Verbal Behaviour*, 5, 351-360.
- Jehl, C., Royet, J.P. and Holley, A. (1997). Role of verbal encoding in short- and long-term odour recognition. *Perception and Psychophysics*, 58(1), 100-110.
- Keppel, G. and Wickens, T.D. (2004). *Design and Analysis: A Researcher's Handbook*. Upper Saddle River, N.J.: Pearson Prentice-Hall.
- Lewandowsky, S., Brown, G.A., Wright, T., and Nimmo, L.M. (2006). Timeless memory: Evidence against temporal distinctiveness models of short-term memory for serial order. *Journal of Memory and Language*, 54, 20-38.
- Lewandowsky, S., Duncan, M., and Brown, G.A. (2004). Time does not cause forgetting in serial short-term recall. *Psychonomic Bulletin and Review*, 11(5), 771-790.
- Miles, C. and Hodder, K. (2005). Serial position effects in odour recognition memory for odours: A re-examination. *Memory and Cognition*. 33(7), 1303-1314.
- Neath, I. (1993b). Distinctiveness and serial position effects in recognition. *Memory and Cognition*, 21(5), 689-698.
- Phillips, W.A. and Christie, D.F.M. (1977a). Components of visual memory. *Quarterly Journal of Experimental Psychology*, 29, 117-133.
- Reed, P. (2000). Serial position effects in recognition memory for odours. *Journal of Experimental Psychology: Learning Memory and Cognition*, 26(2), 411-422.
- Ward, G., Avons, S.E. and Melling, L. (2005). Serial position in curves in short-term memory: Functional equivalence across modalities. *Memory*, 13(3/4), 308-3

Appendix 1

The stimulus set was provided by Dale Air UK Ltd and comprised 120 non-food related odour boxes. The odours are listed below. It should be noted that some of the odours are arbitrarily labelled. Therefore, the label will be of limited utility if one intended to independently reproduce the odour (e.g. 'dragon's breath').

Alpine Laundry Powder	Freesia	Pineapple Plantation
Baby Powder	Fresh Air	Pine/Heather/Peat
Beauty Soap	Frosty	Pit Ponies
Bergamot	Garden Shed	Polish-Wax
Boiler Room	Grass/Hay	Pot-Pourri
Bouquet	Gun Smoke	Riverbank
Brewery	Havana Cigar	Rope/Tar
Burning Peat	Hawaiian	Roses
Burnt Wood	Heather/Bracken	Rotten Egg
Camomile	Honeysuckle	Rubbish Acrid
Cannon	Hyacinth	Sandalwood
Carbolic Soap	Incense	Sea Breeze
Caribbean Holiday	Iron Smelting	Sea Shore
Cedar Wood	Jaguar Spray	Ships Canon
Christmas Tree	Jasmine	Smoke
Church Incense	Lavender	Sports Rub
Cinnamon	Leather	Stable/Horses
Clinic/Hospital	Leather/Hide	Stars Dressing Room
Cloisters	Lemon, Eucalyptus & Mint	Steam/Oil/Ships
Coal Face	Machine Oil	Steam/Oil/Trains
Coal Fire	Man-o-War	Street Bomb
Coal Gas	Methane	Sun, Sand & Coconut
Coal/Soot	Mahogany	Swamp
Cut Grass	Mixed Spice	Sweaty Feet
Deep Heat	Mountain Heather	Sweet Peas
Dentist-Clove Oil	Mummy	Tobacco Leaf
Dinosaur	Mustard Gas	Train Smoke

Dirty Linen	Musty	Tropical
Dragon's Breath	Oak	Tropical Rain Forest
Earthy	Old Drifter	Urine
Eau De Cologne	Old Inn	Victoria Lavender
Egyptian Mummy	Old Smithy	Violets
Eucalyptus	Old River	Volcano
Factory	Out At Sea	Vomit
Farmyard	Ozone	Wallflower
Fish Market	Peat	Washday
Flatulence	Pencil Shavings	Wild Stag
Flowery	Peppermint	Wine Cask-Oak
Forest	Phosgene Gas	Woodsmoke
Fox	Pine	Ylang Jasmine and Myrrh

Figure 1

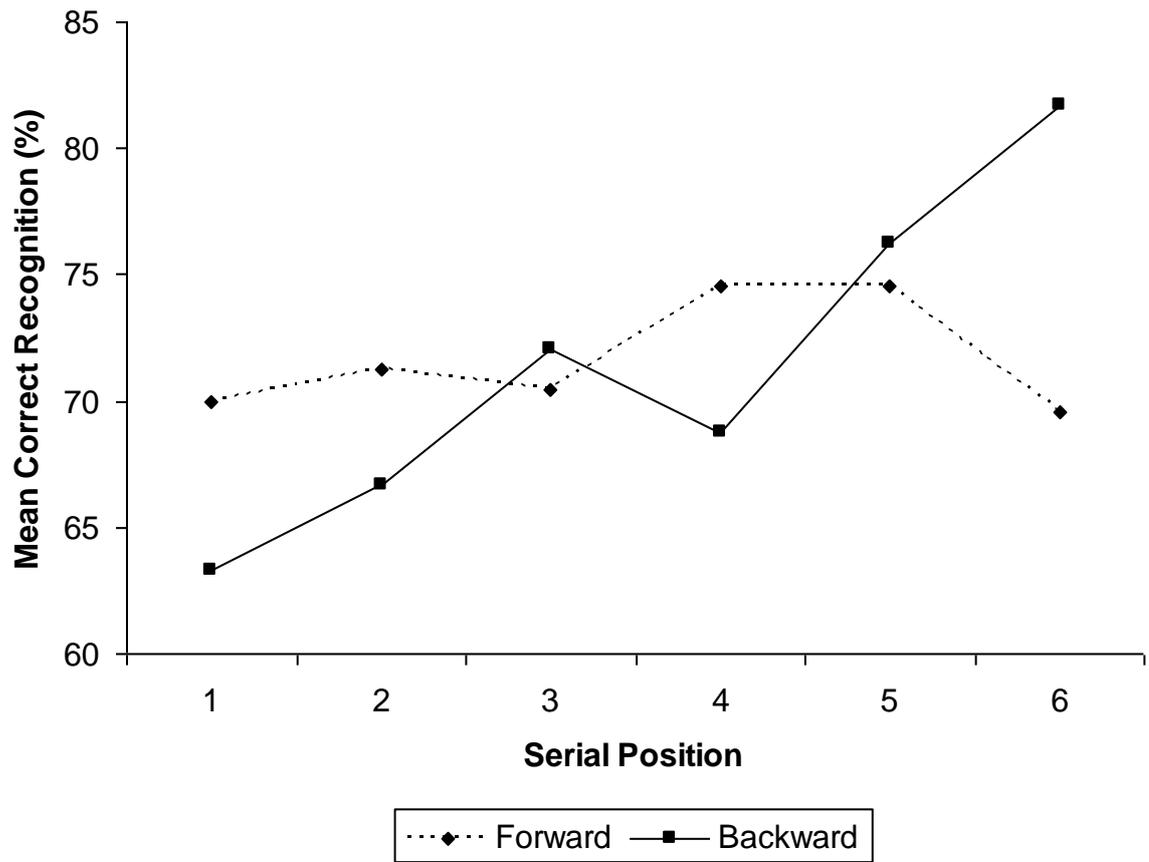


Figure 2

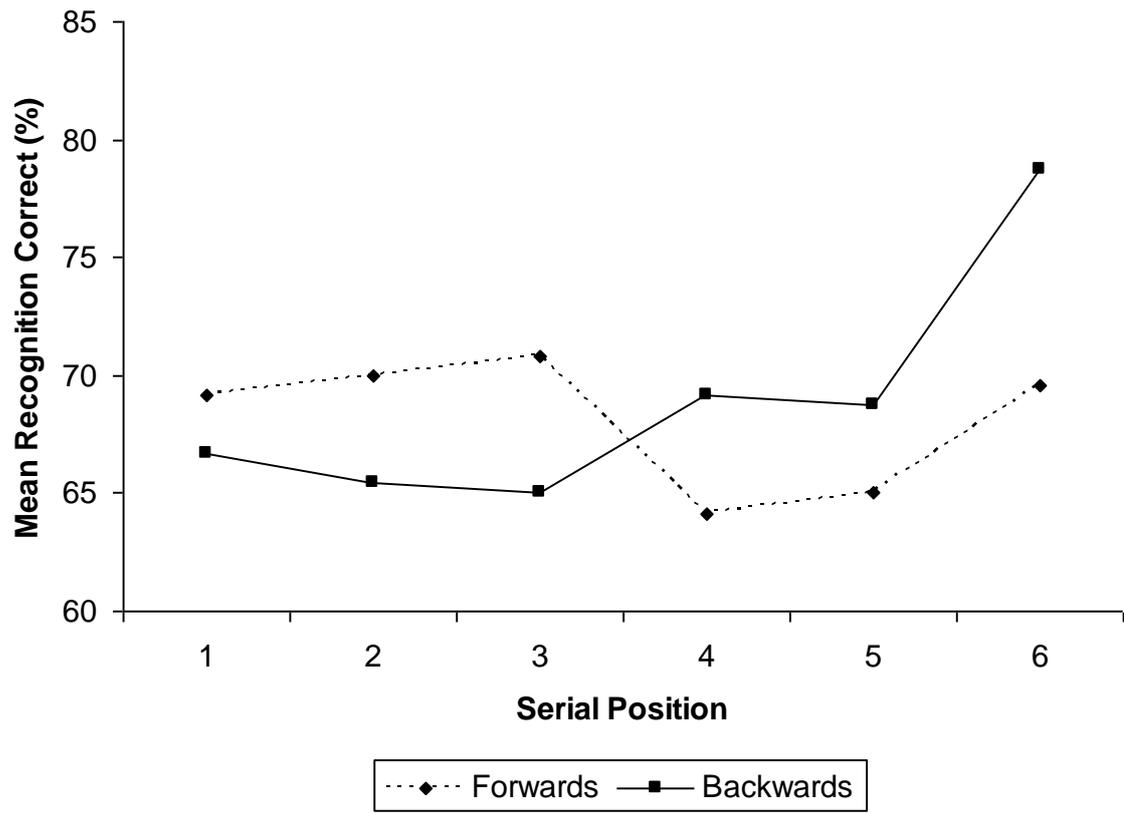
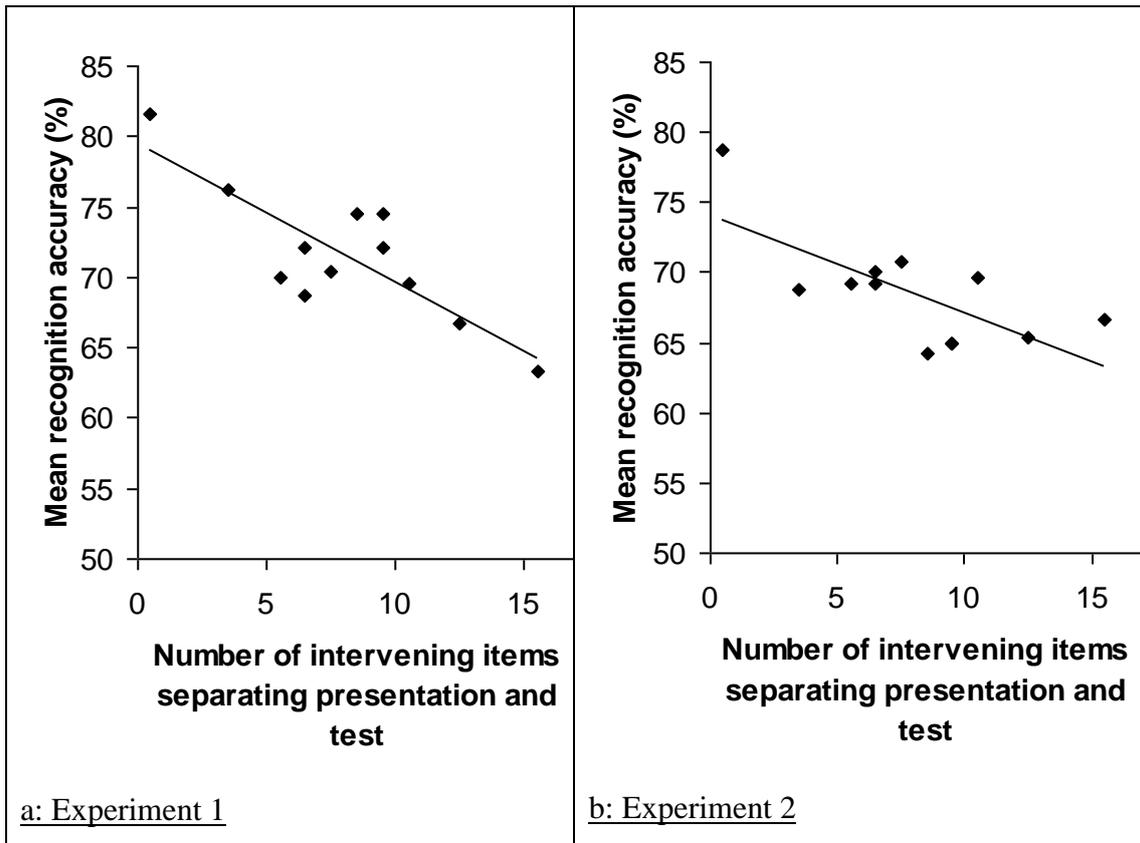


Figure 3(a-b)



Legends

Legend 1

Figure 1: Mean percentage correct recognition scores as a function of testing procedure (forward versus backward) and serial position.

Legend 2

Figure 2: Mean percentage correct recognition scores as a function of testing procedure (forward versus backward) and serial position.

Legend 3

Figure 3(a. and b.): Mean recognition accuracy for Experiments 1 and 2 in both the forward and backward testing procedure as a function of the number of intervening items between learning and test. The trend line indicates the line of best fit.