

## Linking Associative and Serial List Memory: Pairs Versus Triples

Jeremy B. Caplan, Mackenzie G. Glaholt, and Anthony R. McIntosh  
The Rotman Research Institute, Baycrest Centre for Geriatric Care

Paired associates and serial list memory are typically investigated separately. An “isolation principle” (J. B. Caplan, 2005) was proposed to explain behavior in both paradigms by using a single model, in which serial list and paired associates memory differ only in how isolated pairs of items are from interference from other studied items. In the present study, 2 experiments identify a critical dissociation between the 2 paradigms, challenging this unified account. Specifically, forward and backward probes were highly correlated for pairs and less so for short lists (triples). The authors asked whether the isolation principle could quantitatively accommodate this type of dissociation. A simulation confirmed that a single model incorporating the isolation principle can adequately explain this and other dissociations, supporting the common processes view.

Paired associates and serial list memory paradigms have much in common. Memory for associations tests whether participants can remember that two items are linked (usually by having been presented at adjacent times during a study episode). Memory for serial lists also requires participants to recall the relationships among items, but serial lists include more than two items. One important difference is that with a pair of items, AB, order is irrelevant for many types of memory probes. For example, consider cued recall. When presented with one item of the pair, the participant need only recall the other item, regardless of whether it was the first or second item. In contrast, when presented with a single item from a list of three items, ABC, the participant may require some order information to recall the actual target item and rule out the third, incorrect item. Memory investigators have generally asked questions about associative memory and serial list memory in separate studies, underlining their differences. But models that account for a broad range of data have more explanatory value. If it could be shown that there are common mechanisms underlying behavior in both associative and serial list memory, then researchers could not only explain behavior in those two paradigms, but also potentially explain behavior in related and hybrid paradigms. By investigating associative and serial list memory together, we identify how performance on the two paradigms is similar and how it differs (Experiments 1 and 2), and test whether a single model can account for both the similarities and differences in behavior by using identical model processes (Simulation).

---

Jeremy B. Caplan, Mackenzie G. Glaholt, and Anthony R. McIntosh, The Rotman Research Institute, Baycrest Centre for Geriatric Care, Toronto, Ontario, Canada.

This research was supported by the Canadian Institutes of Health Research. We thank Michael Kahana and Kelly Addis for helpful discussions, Morris Moscovitch and his group for their generosity in the use of their laboratory, and Esther Fujiwara and Debbie Talmi for comments on the manuscript.

Correspondence concerning this article should be addressed to Jeremy Caplan, who is now at the Psychology Department, Biological Sciences Building, University of Alberta, Edmonton, Alberta T6G 2E9, Canada. E-mail: jcaplan@ualberta.ca

The notion that associative and serial list memory are closely related dates back to Ebbinghaus’s (1885/1913) *chaining* model. In chaining, the participant learns a list by encoding item-to-item associations between neighboring items in the list (and sometimes between pairs of items separated by a lag):

$$\{A-B, B-C, C-D, D-E, E-F, \dots\}, \quad (1)$$

where letters denote items (e.g., nouns) and the dash symbol denotes a learned association. The participant could retrieve the list by starting at the first item, A, using the A–B association to retrieve B, then the B–C association to retrieve C, and so on until the end of the list or some other stopping criterion. Lewandowsky and Murdock (1989) translated this concept into mathematical form. Using this model, a participant can also easily learn a list of pairs by excluding or weakening the between-pair associations (Murdock & Franklin, 1984):

$$\{A-B, C-D, E-F, \dots\}. \quad (2)$$

Caplan (2005) introduced a complementary strategy for unifying paired associates and serial list memory: adapting a positional coding (or order coding) model (e.g. Brown, Preece, & Hulme, 2000; Burgess & Hitch, 1999; Conrad, 1965; Henson, 1998; Lee & Estes, 1977) of serial list learning to account for data on paired-associates learning. As initially suggested by Conrad (1965) for serial lists, a list is acquired by learning associations between list items and an abstract representation of position:

$$\{A-1, B-2, C-3, D-4, E-5, F-6, \dots\}. \quad (3)$$

A list of pairs, A–B, C–D, E–F . . . , would be learned by associating the A and B items of each pair with very similar positions:

$$\{A-1, B-1.1, C-3, D-3.1, E-5, F-5.1, \dots\}. \quad (4)$$

Thus, there are simple ways of treating associative and serial list memory within the same theoretical framework. However, empirical evidence is needed to test whether associative and serial list memory have anything in common or whether they should be treated separately. On one hand, behavioral dissociations between paradigms can reflect distinct underlying mechanisms, suggesting separate modeling frameworks. On the other hand, sometimes a

single model can produce such dissociations between paradigms even while the underlying mechanisms are the same. Determining how to model the two paradigms together and nonetheless account for the dissociation often results in a more constrained model and could lead to insights as to how memory works in both association and list paradigms. Some studies have already revealed behavioral dissociations between associative and serial list memory. Rather than indicating that the two types of memory are distinct, such dissociations could provide powerful clues as to how the paradigms should be unified, what they have in common, and exactly how they differ (cf. Caplan, 2004, 2005).

We first review the existing dissociations between the two paradigms. We then suggest that within-experiment comparisons of cued recall of pairs (associations) and triples (lists of three items) will provide key data regarding exactly how associative and serial list memory differ, and provide constraints on how to unify them. We then present two experiments comparing cued recall of pairs and triples and show how the results, including apparent dissociations, support unifying models in which paired-associates learning and long-list learning represent extremes on a continuum of paradigms.

### Dissociations Between Associative and Serial List Memory

One of the most striking findings in paired-associates learning is that cued recall performance does not depend on probe direction. If a participant is probed on a pair AB in the forward direction (the question A?) and subsequently on the same pair in the backward direction (?B), accuracy is highly correlated, nearly as much as when a participant is probed twice in the same direction (Kahana, 2002; Rizzuto & Kahana, 2001). In other words, the experimenter would obtain the same result (in terms of accuracy) whether the participant was probed for a particular pair in the forward or backward direction. These authors used this high forward-backward correlation to support the notion that pairs are learned as holistic units, a notion that was termed *associative symmetry* (Köhler 1947; Asch & Ebenholtz, 1962). In contrast, in cued recall of 19-word lists, previously learned to a criterion of perfect serial recall (Kahana & Caplan, 2002), the correlation between forward and backward probe accuracy took on a lower, midrange value (Caplan, 2005).

Complementary to the correlation between forward and backward probes, researchers have long observed that the mean accuracy on backward probes of pairs is nearly equal to that on forward probes of pairs (Asch & Ebenholtz, 1962; Horowitz, Brown, & Weissbluth, 1964; Horowitz, Norman, & Day, 1966; Köhler 1947; Murdock, 1962, 1965, 1966; reviewed by Kahana, 2002). In contrast, Kahana and Caplan (2002) found that cued recall of both 3-word (triples) and 19-word lists showed an overall accuracy and response time advantage for forward probes. Although this stands as an important dissociation between memory for associations and memory for lists, it does not speak directly to the question of associative symmetry (Kahana, 2000). Furthermore, as we demonstrate below, findings of symmetry or asymmetry in mean performance do not constrain the unified framework account of associative and serial list memory (see General Discussion; see also Caplan, 2005).

### The Isolation Principle.

Dissociations between associative and serial list learning could be used to argue that the cognitive processes underlying learning and recall of associations are fundamentally different than those relevant to lists. However, Caplan (2004, 2005) proposed that one could account for these dissociations parsimoniously within a single model (either a chaining model or a positional coding model) by invoking the *isolation principle*. According to this principle, the only difference between memory for lists and memory for associations is the degree to which nearest-neighbor pairs are isolated from the remaining studied items.

For the chaining model, consider the pair [C–D] in the list of pairs in Equation 2. Because between-pair associations are weak, performance on the forward probe, C–?, depends chiefly on the encoding strength of the C–D association. Similarly, the backward probe, ?–D, taps memory for this same C–D association strength. Because the pair C–D is largely isolated from the rest of the list, there are only very low levels of competition from other list items. Therefore, these two probe questions will be highly correlated. This assumes that the encoded association from C to D is highly correlated to the association from D to C, a necessary assumption to fit the high correlation between forward and backward probes (Rizzuto & Kahana, 2001). In contrast, in probing the pair {C–D} in the serial list (Equation 1), the forward probe depends not only on the encoded strength of the C–D association, but also on the strength of competing associations (e.g., B–C and remote associations if present). The backward probe still relies on the strength of the C–D association, but now its competing associations are different (e.g., D–E). Because competition may come from different parts of the list depending on probe direction, forward and backward probes will be less correlated than was the case with pairs.

The positional coding model, although it operates on different study and retrieval processes, can be analyzed in an analogous way. A critical assumption of positional coding (and order coding) models is that interference or competition between list items is a function of the proximity of their encoded positions. In a list of pairs (Equation 3), paired items have similar positional codes, whereas other list items are coded with relatively distant codes. Thus, a forward probe (e.g., C–?) relies on the encoded strength of C–3 and the strength of D–3.1, and a backward probe (?–D) relies on the same item–position associations. Interference from other list items is minimal, keeping the forward–backward correlation high. In contrast, for lists, more interference comes from earlier list items (e.g., B–2) than later list items for the forward probe, whereas the backward probe faces more interference from later list items (e.g., E–5). Like the chaining model, this differential interference based on probe direction tends to reduce the forward–backward correlation.

Thus isolation, whether associative or positional, controls which other list items compete at time of test, which leads to specific predictions related to this interference.

### Comparing Pairs and Triples

Kahana and Caplan (2002) underlined the importance of studying triples as an intermediate between associations and lists. Triples are the smallest possible serial lists. As such, they could have properties similar to either lists or pairs. For example, Caplan

(2004) implemented a simulation of a chaining model with the isolation principle and had it learn lists containing both pairs and triples, anticipating the experiments reported here. The simulation predicted a less-than-perfect correlation between forward and backward probes for triples, even in the same simulated conditions that simultaneously showed a perfect correlation for pairs. Thus, the dissociation as measured by this correlation should be present even at the very shortest possible list length. For triples, the interference comes primarily from the item of the triple that is neither the probe item nor the target item, whereas in pairs the interference is the same for forward and backward probes. However, to date, no one has measured the correlation between forward and backward probes of triples.

In the present study, we directly compared pairs with triples. Our main measure of interest was the forward–backward correlation in a successive testing paradigm. Our main memory test was cued recall. For tests of list memory, this type of cued recall probe is known as the *sequential probe technique*, wherein the participant has to recall the word in the list that immediately followed the probe word (Detterman, 1977; Murdock, 1968; Posnansky, 1972; Woodward, 1970; Woodward & Murdock, 1968) or the list word that immediately preceded the probe (Kahana & Caplan, 2002). We present two experiments in which we controlled as much as possible for study and test conditions between pairs and triples. In Experiment 1, each list in a session comprised either pairs or triples (“pure lists”), but participants studied both pair lists and triple lists, intermixed throughout the experimental session. Participants were given cued recall questions twice for each tested pair or triple (“successive testing”) so that the correlation between forward and backward probes could be measured. All study and test conditions apart from the groupings of the words at study were matched between pairs and triples, including the total list length (total number of words is 18 for both pair lists and triple lists). Experiment 2 tested how the findings of Experiment 1 generalized to a paradigm in which each list contained both pairs and triples (“mixed lists”).

Finally, intrusions can give us important clues as to the sources of competition at time of test. To make the isolation principle work (i.e., to be able to treat memory for pairs and lists in the same model), both positional coding and chaining models predict a contiguity effect in lists of pairs, in which items studied near the probe and target item intrude more than words studied at very different times. This is in contrast to most existing models of paired associates learning, which do not explicitly encode associations from one pair to another (i.e., no remote associations). The same prediction applies to triples, with one exception: The bulk of within-list intrusions should come from the “triple-lure” item, that is, the item of the triple that is neither the probe nor the target item; this item is not associatively or positionally isolated from the probe and target items.

## Experiment 1: Pure Lists of Pairs or Triples

### Method

#### Participants

Forty-five volunteers participated for monetary compensation. Three participants were excluded because they had ceiling or floor performance in at least one condition (AB/AB\_/\_BC  $\times$  Forward/Backward), leaving

42 participants included in the analyses. Of these, 11 were male and 31 were female (mean age  $\pm$   $\sigma$  = 23.6  $\pm$  7.3 years). All participants spoke English as their primary language.

#### Materials

The fixation consisted of seven asterisks presented in the center of the screen, displayed for 3,750 ms and then erased for 250 ms.

Lists consisted of nouns from the Toronto Word Pool (Friendly, Franklin, Hoffman, & Rubin, 1982), sampled randomly without replacement. Each noun was presented visually in the center of the screen. The lists were grouped into either nine pairs or six triples, which kept the total list length constant at 18 words. The order of pair lists and triple lists was chosen randomly, with the constraint that each set of three lists had to include one pair list and two triple lists. This ensured that an equal amount of data was collected on pairs, AB\_\_ triples, and \_\_BC triples (because there are twice as many ways of probing triples). Each noun was displayed for 1,750 ms followed by a 250-ms blank interstimulus interval (ISI). An additional interval of 4,000 ms was inserted between pairs and triples. During this interblock interval, the participant viewed either \*\*\*2\*\*\* (lists of pairs) or \*\*\*3\*\*\* (lists of triples). This interblock cue served to remind the participant of whether they were studying a list of pairs or a list of triples.

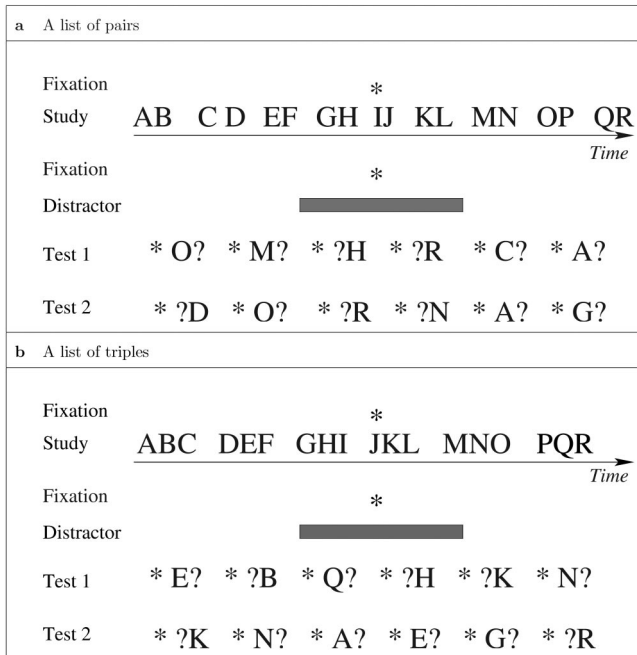
The distractor consisted of four equations of the form  $A + B + C = ?$ , where A, B, and C were randomly selected digits from 0 to 6, with the restriction that the identical distractor could not be used twice in succession. Each equation remained on the screen for 3,750 ms and then was erased for 250 ms. The participant was asked to respond vocally with the correct answer to the equation and was given a fixed interval of 4,000 ms within which to respond.

Cued recall consisted of a probe word with six question marks, either to the left or to the right of the word. The participant was instructed to recall the word that followed or preceded the probe item, depending on whether the question marks were placed to the right or left of the probe word, respectively. Each probe was preceded by a fixation. The probe remained on the screen for 7,000 ms and then was erased for 1,000 ms. The participant was asked to respond vocally and was given an interval of 8,000 ms within which to respond. All the pairs and triples were either probed twice or, for some pairs, not at all. All the probed pairs and triples were probed once in random order (Test 1); then, without additional delay, all pairs and triples were probed once again in a new random order (Test 2). The probe order was constrained such that the first probe on Test 2 could not be on the same pair or triple as the last (sixth) probe on Test 1. Vocal responses to the distractor and cued recall probes were recorded digitally and scored later for accuracy and response time by replaying the response and marking its onset manually in the sound file.

If a triple was probed for its AB subpair on Test 1, it would always be probed for the AB subpair on Test 2 (denoted AB\_\_ triples), and likewise for the BC subpair (denoted \_\_BC triples). The first and second probes of AB\_\_ triples could be in the forward (A?) or backward (?B) direction. Similarly, \_\_BC triples could be probed in the forward ( B?) or backward (?C) direction. The set of probe directions for Tests 1 and 2 was counterbalanced separately for pairs and triples across sets of two lists. Whether the first (AB) or second (BC) subpair of a triple (ABC) was probed was independently counterbalanced across sets of two lists, with equal overall proportions of AB and BC probes. All responses were recorded and subsequently scored for accuracy and latency. All triples were probed once each in Tests 1 and 2. A subset of six of the nine pairs from each pair list was selected to be successively tested, to control for test-order effects.

#### Procedure

Figure 1 illustrates the procedure for a single trial. Participants first viewed a fixation, then studied the list in a single study trial. Next, they performed the distractor task, and finally, they answered cued recall



*Figure 1.* Procedure for Experiment 1 (pure lists). a: A single pair-list trial. Each capitalized letter represents a unique noun. The asterisk (\*) denotes a fixation presented for 3,750 ms and then erased for 250 ms. In a study trial, words are presented grouped in pairs. A mathematical distractor follows (denoted by the gray rectangle), and then two complete probes of the list. Six of the nine pairs are probed once in Test 1 and then once again in Test 2; the remaining three pairs are not probed at all. b: Procedure for a single triple-list trial. Note that all triples are probed. For the hypothetical example shown here, the pairs AB, OP, and QR and the triples DEF, JKL, and MNO were probed in the same direction on Tests 1 and Test 2, whereas the pairs CD, GH and MN and the triples ABC, GHI and PQR were probed in the opposite direction on Test 1 and 2. The pairs EF, IJ, and KL were not probed.

questions based on the list. A session consisted of 20 lists. For the first list, self-paced instructions preceded each of the study, distractor, and cued recall phases of the task. The first two lists included one pure list of pairs and one pure list of triples; the order of these was randomized across participants. These two practice lists were excluded from all analyses.

### Analyses

All analyses of variance (ANOVAs) are reported with the Greenhouse-Geisser correction for nonsphericity to correct degrees of freedom when levels of a factor are not completely independent. Effects are considered significant based on an alpha level of .05, and post hoc pairwise comparisons are always Bonferroni-corrected. Only effects that reach significance are reported.

Correlation in accuracy was measured in terms of Yule's  $Q$ , which allows one to measure correlation between discrete, two-level variables. Yule's  $Q$  is equivalent to a discrete correlation for a  $2 \times 2$  contingency table (for a review, see Kahana, 2002). For each  $2 \times 2$  contingency table, there are four different tallies of the possible Test 1/Test 2 outcomes: (a) = correct/correct, (b) correct/incorrect, (c) incorrect/correct, and (d) incorrect/incorrect (Yule's  $Q = [ad - bc] \div [ad + bc]$ ). Differences in  $Q$  were evaluated using the chi-square test for the interaction between the two contingency tables based on the log-odds ratio transform of  $Q$  as discussed by Bishop, Fienberg, and Holland (1975) and Hayman and Tulving (1989).

For response time correlation analyses, response times for cases in which Test 1 and Test 2 were both correct were percentile-transformed (rank-transform divided by the total number of response times) for each participant individually. Spearman correlations were then computed on the transformed values compiled across participants. Because response times had to be collapsed across participants for this analysis, in evaluating differences between conditions we used a permutation test (Good, 2000) in which the response times from the conditions to be compared had their condition labels shuffled at random 1,000 times. This resampling method evaluates the obtained correlation values against a null hypothesis that there is no systematic difference between a pair of conditions, and controls for subject variability.

### Results and Discussion

First, we report mean performance measures for accuracy and response times to identify potential sources of variance and confounds for the correlational analyses. These also serve to connect the present findings to prior data on mean performance and, in particular, to the question of whether forward and backward probes are of equal average difficulty. Second, we present correlational analyses, which are the main focus of this article. The correlations directly test the relationship between forward and backward probes at the level of individual pairs. Finally, we analyze error-response patterns, testing further predictions of the isolation principle account.

#### Mean Performance Measures

Accuracy and response time were analyzed separately. The effects of response time largely reinforced those found for accuracy; for the sake of brevity, we report findings for accuracy followed by only the most relevant findings for response time. We wanted to assess the effects of serial position of the probed pair or triple (input position in the presented list) and probe position (the order in which a pair or triple was tested within the first or second set of tests). Breaking down the data by serial position and probe position resulted in high levels of missing data; for this reason, we performed separate analyses for serial position and probe position. In this experiment, lists of pairs had nine serial positions (nine pairs) and lists of triples had six (six triples). To include pairs and triples in the same factorial design, we collapsed serial position into three bins (three pairs or two triples per bin). Accuracy on the probes is plotted as a function of serial position bin (Figure 2) and probe position (Figure 3).

Serial position bin and probe position could be important sources of variability that could serve to increase correlations overall. Test number (Test 1 vs. Test 2; first/second probe of the pair/triple, respectively) effects could reveal the effects of output encoding or forgetting, which are important for accounting more fully for the measured patterns of correlations. Pair/triple type (pairs vs. AB\_\_ triples vs. \_\_BC triples) can provide clues as to the overall difficulty of the three conditions.

Probe direction (forward vs. backward probes) effects address the question of whether probed recall is symmetric on average. But more important, any interactions between factors (especially involving pair/triple type) and probe direction could represent confounds for comparing forward-backward correlations across conditions.

*Serial position bin analysis.* We performed ANOVAs on Serial Position Bin (3)  $\times$  Test Number (2)  $\times$  Pair/Triple Type (3)  $\times$  Probe Direction (2). Main effects of test number,  $F(1, 41) = 98.71$ ,

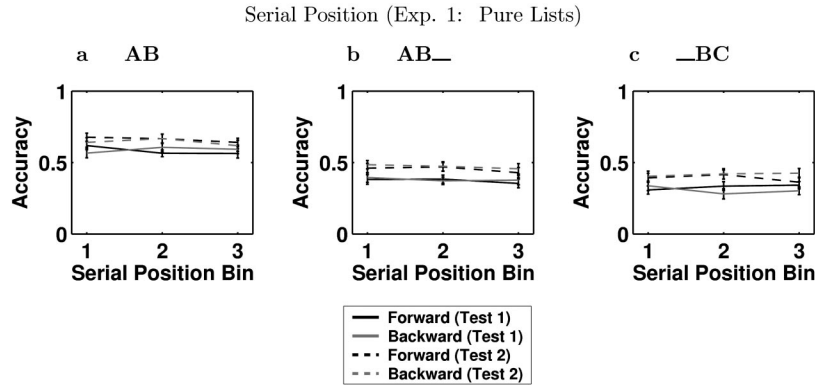


Figure 2. Experiment 1, serial position curves: Accuracy as a function of serial position, probe type (columns), Test 1 versus Test 2 (solid vs. dashed lines, respectively), and forward versus backward probe direction (black vs. gray lines, respectively). Serial position represents binned serial positions (1st, 2nd, and 3rd third of the list). Error bars are *SEM* across participants, corrected for between-subjects variability (Loftus & Masson, 1994).

$MSe = 0.024$ ,  $p < .001$ , and pair/triple type,  $F(2, 69) = 106.95$ ,  $MSe = 0.102$ ,  $p < .001$ , were significant. Post hoc pairwise  $t$  tests showed that participants were more accurate on the second test than the first, which suggests the use of output encoding (Figure 2a–c, dashed vs. solid lines), and more accurate on pair probes than on triple probes. Accuracy was greater for AB\_ triples than \_BC triples ( $ps < .001$ ; compare Panels a, b, and c, respectively, in Figure 2). Response time effects largely paralleled accuracy effects. In addition, the three-way interaction Serial Position Bin  $\times$  Pair/Triple Type  $\times$  Probe Direction was significant,  $F(4, 152) = 2.77$ ,  $MSe = 0.471 s^2$ ,  $p < .05$ . Simple effects analysis found only trends toward significant forward advantages in two conditions: for pairs, Serial Position Bin 2, and for \_BC triples, Serial Position Bin 3 ( $p < .1$ ).

**Probe position analysis.** Note that *probe position* refers to the probe order at the time of test, regardless of serial position bin during study. We performed ANOVAs on Probe Position (3)  $\times$  Test Number (2)  $\times$  Pair/Triple Type (3)  $\times$  Probe Direction (2). Since all other factors in the serial position bin and probe position ANOVAs were the same, it is to be expected that any effects not

involving serial position bin should trivially replicate here. Indeed, for accuracy, the only significant effects in this analysis were redundant with those found in the serial position bin analyses, plus a main effect of probe position,  $F(5, 180) = 3.00$ ,  $MSe = 0.103$ ,  $p < .05$ , with no significant post hoc pairwise comparisons. For response time, the three-way interaction did replicate.

**Summary.** The serial position and probe position curves are remarkably flat (Figures 2 and 3). The lack of a main effect of serial position bin and interactions suggests that the short distractor task was sufficient in eliminating recency, and that primacy was negligible. There was no main effect of probe position on response time, and no interactions involving probe position. Although the main effect of probe position was significant for accuracy, no corrected pairwise differences were significant. Performance was best for pairs and better for the AB\_ triples than for the \_BC triples (Figure 2). The total number of words in the pair lists was designed to be identical to the total number of words in the triple lists (namely, 18 words). Therefore, the study-set complexity cannot account for these differences. In addition, because the within-pair ISI was equated to the within-triple ISI, and the between-pair

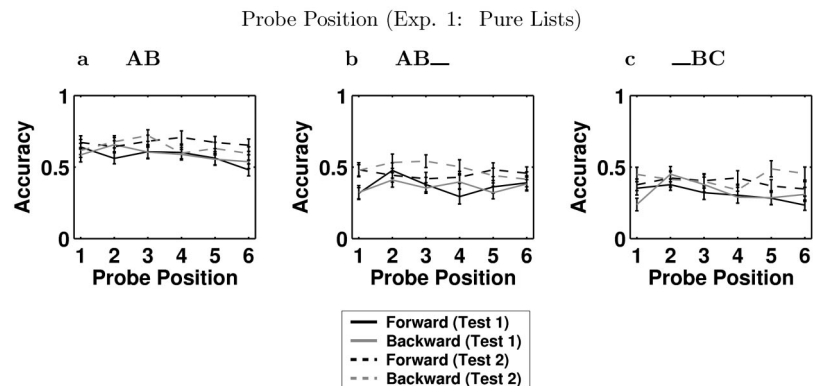


Figure 3. Experiment 1, probe position curves: Accuracy as a function of probe position, probe type (columns), Test 1 versus Test 2 (solid vs. dashed lines, respectively), and forward versus backward probe direction (black vs. gray lines, respectively). Error bars are *SEM* across participants, corrected for between-subjects variability (Loftus & Masson, 1994).

ISI was equated to the between-triple ISI, the total presentation time for lists of pairs was greater than that for triples. Thus, study–test lag was, on average, greater for pairs than for triples, which should tend to hurt pair performance, the reverse of what we observed. On the other hand, each pair may have received a greater number of rehearsals per pair word than per triple word, which may have boosted pair recall, a difference that would no longer be a confound in Experiment 2. A performance advantage was found for Test 2 over Test 1 (dashed vs. solid lines in Figures 2 and 3). This may be evidence of effective output encoding during cued recall. Nonetheless, accuracy and response times were relatively close in value between the two tests; thus these effects, although significant in this sample, are small in magnitude.

Finally, the relative equality of cued recall accuracy and response time (not plotted) between forward and backward probes is striking (black vs. gray lines in Figures 2 and 3). This replicates prior findings of symmetric accuracy in pairs (Kahana, 2002; Murdock, 1962, 1965, 1966) and extends this result to symmetry of response times, a result that was suggested for pairs by Waugh (1970).

On the other hand, the finding of symmetric performance in triples contrasts with prior findings. Kahana and Caplan (2002) found that cued recall of both triples and 19-noun lists was asymmetric, showing an advantage for forward probes in both accuracy and response time. Thus, whereas symmetric mean performance in cued recall of pairs is remarkably general, asymmetric mean performance in cued recalls of lists (including very short lists) appears to depend on particular experimental conditions. This could be an effect of degree of learning: prior findings of forward-probe advantages occurred in paradigms with multiple presentations of the lists, whereas our experiment involved only a single presentation of each triple. The suggestion that directionality emerges only after multiple presentations was put forward by Waugh (1970). Furthermore, multiple presentations may result in additional covert practice with the studied materials, potentially resulting in greater experience in covertly producing later items rather than earlier items, along the lines of Horowitz and colleagues' view (Horowitz et al. 1964, 1966).

### Correlations on Successive Tests

We now turn to the analysis that is the main focus of this article, namely, the correlation between the two successive tests as a function of probe direction on Tests 1 and 2. If this correlation is nearly as high as the correlation between test and retest effects, this would suggest that forward and backward probes tap the same stored information, consistent with associative symmetry. If the forward–backward correlation is substantially lower, there must be additional sources of variability that differ depending on probe direction. Figure 4a shows Yule's  $Q$ , the correlation for discrete, two-level data (Bishop et al. 1975), as a function of probe type and direction. "Same" refers to Test 1/Test 2 in the same direction (forward/forward or backward/backward). "Diff" refers to Test 1/Test 2 in the opposite direction (forward/backward or backward/forward). "Control" is an estimate of correlation due to list-to-list variability; it is the correlation between Test 1 taken from one pair or triple, and Test 2 taken from a different pair or triple from the same list (computed identically for AB and BC). All permutations of pairs or triples, respectively, on each list were used to compute the control correlation.

All  $Q_{\text{Diff}}$  values were significantly different from  $Q_{\text{Control}}$  (pairs:  $z = 7.1$ , AB:  $z = 9.4$ , BC:  $z = 9.4$ , all  $ps < 10^{-8}$ ). All  $Q_{\text{Diff}}$  values were also significantly different from  $Q_{\text{Same}}$  (pairs:  $z = 3.5$ ,  $p < .0005$ ; AB:  $z = 5.3$ ,  $p < 10^{-6}$ ; BC:  $z = 5.6$ ,  $p < 10^{-7}$ ). For the critical comparisons,  $Q_{\text{Same}} - Q_{\text{Diff}}$  was significantly smaller for pairs than for both types of triples (AB:  $z = 2.8$ ,  $p < .01$ , BC:  $z = 3.5$ ,  $p < .0005$ ). Finally, the quantity  $Q_{\text{Same}} - Q_{\text{Diff}}$  did not differ between AB and BC triples ( $z = 0.7$ ,  $p > .1$ ). Thus, cued recall of triples taps directionally sensitive information to a greater degree than cued recall of pairs. Although  $Q_{\text{Control}}$  was greater for pairs than for triples, these differences were not significant. Nonetheless, we follow up on this curious effect in the Simulation section.

It is possible that forward and backward probes tap different information, but that this is evident only when one looks at response times in a kind of speed–accuracy tradeoff. To complement the Yule's  $Q$  analyses, we analyzed the Pearson correlation ( $r$ ) for response times for correct responses. Note that the comparison of response time correlation between pairs and triples should be interpreted with caution due to the three-way interaction found in mean response time analyses, but the pattern for pairs alone is straightforward. Because it was not clear how to calculate a within-list control correlation, control correlations were not computed. Figure 5 plots these correlations for pairs, AB\_\_ triples, and \_\_BC triples. First,  $r_{\text{Diff}}$  was not significantly different from  $r_{\text{Same}}$  for pairs or \_\_BC triples, but  $r_{\text{Diff}}$  was significantly lower than  $r_{\text{Same}}$  for AB\_\_ triples ( $\Delta r = .20$ ,  $p < .05$ ). Although  $r_{\text{Diff}}$  was not significantly different for pairs versus each type of triple individually, in comparison of  $r_{\text{Diff}}$  for pairs versus both types of triples combined,  $r_{\text{Diff}}$  was significantly greater for pairs than for triples ( $\Delta r_{\text{Diff}} = .15$ ,  $p < .05$ ), corroborating findings for accuracy.

**Summary.** Successive testing allowed us to examine the correlations between two different ways of probing pairs and triples—in the forward and backward directions. Prior studies have found that the correlation between forward and backward successive probes of pairs is near perfect (Kahana, 2002; Rizzuto & Kahana, 2000, 2001), but is only moderate for probed recall of long, 19-word serial lists (Caplan, 2005). According to the isolation principle framework introduced by Caplan (2005), pairs and long lists represent extremes of a continuum, wherein what differs is only the degree to which adjacent pairs of items are isolated from the rest of the list. According to this account, triples should begin to show (albeit only slightly) lower correlations between forward and backward probes. This pattern is evident in that, although  $Q_{\text{Diff}}$  was reduced relative to  $Q_{\text{Same}}$  for pairs and triples, the magnitude and statistical robustness of this effect was larger for both types of triples than for pairs (Figure 4). The more direct test of the isolation principle, that this correlation difference should differ between pairs and each type of triple, was supported for both AB and BC triple probes. The response time correlations (Figure 5) show a similar distinction between pairs and triples, wherein response times for correct forward and backward recalls are less correlated for triples than for pairs. This helps to rule out an alternate account—that participants trade off accuracy and response time differently for pairs and triples. Instead, accuracy and response time analyses are consistent with one another, rather than contrasting. This is the first study to analyze  $r_{\text{Same}}$  versus  $r_{\text{Diff}}$  for either pairs or triples.

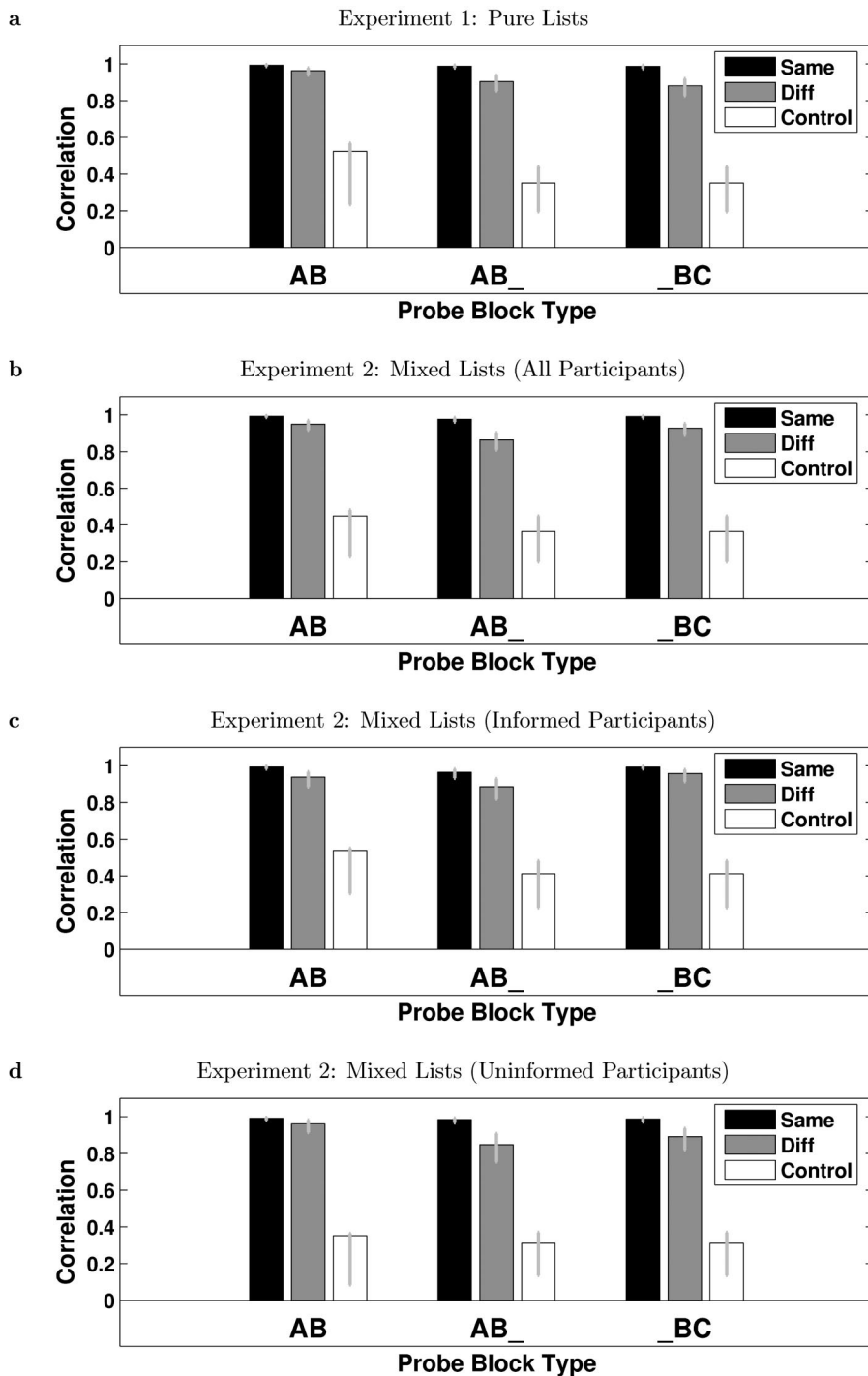


Figure 4. Accuracy correlations for both experiments. Yule's  $Q$  between Test 1 and Test 2 for each of the block types for pure lists in Experiment 1 (a), mixed lists in Experiment 2 for all participants (b), precued participants (c), and postcued participants (d). Same: both probes in the same direction; Diff: probes are in opposite directions; Control: Test 1 and Test 2 are taken from different pairs or triples from the same list, collapsed across probe directions. Error bars denote 95% confidence intervals.

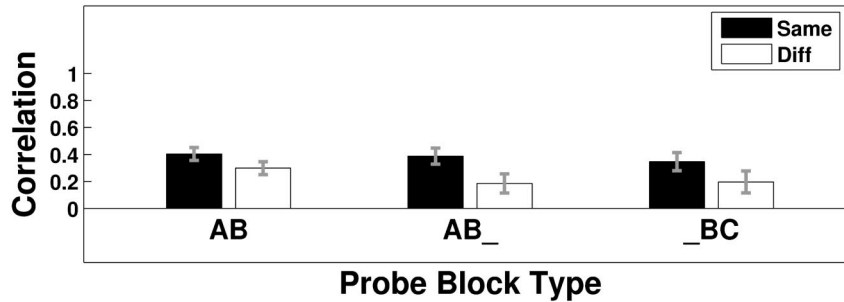


Figure 5. Response time correlations for Experiment 1. Pearson correlation between Test 1 and Test 2 (when both tests were recalled correctly) for each of the block types for pure lists in Experiment 1. Same: both probes in the same direction; Diff: probes are in opposite directions. Error bars denote *SEM* based on bootstrap resampling (1,000 resamples with replacement).

### Intrusions

Finally, examining the patterns of error responses could allow us to further understand the nature of differences between memory for pairs and triples. Many of the errors committed were omission errors—participants either explicitly passed (vocalizing the word *pass*) or made no response within the allotted 8 s. However, participants also made a total of 1,739 intrusions, or  $38 \pm 36\%$  (mean  $\pm$  standard deviation across participants) of all probes. Each

participant made at least one intrusion. Some intrusions were to words that had not been presented to the participant, whereas others were items presented in the same or prior study lists.

“Triple-lure” intrusions. Of the 1,739 intrusions, 411 were the unprobed item of triple blocks, namely, the item of a triple that was neither the target nor the probe item of the cued recall probes. These are plotted as dashed error bars in Figure 6b and are corrected for the fact that not all lags were available as intrusions

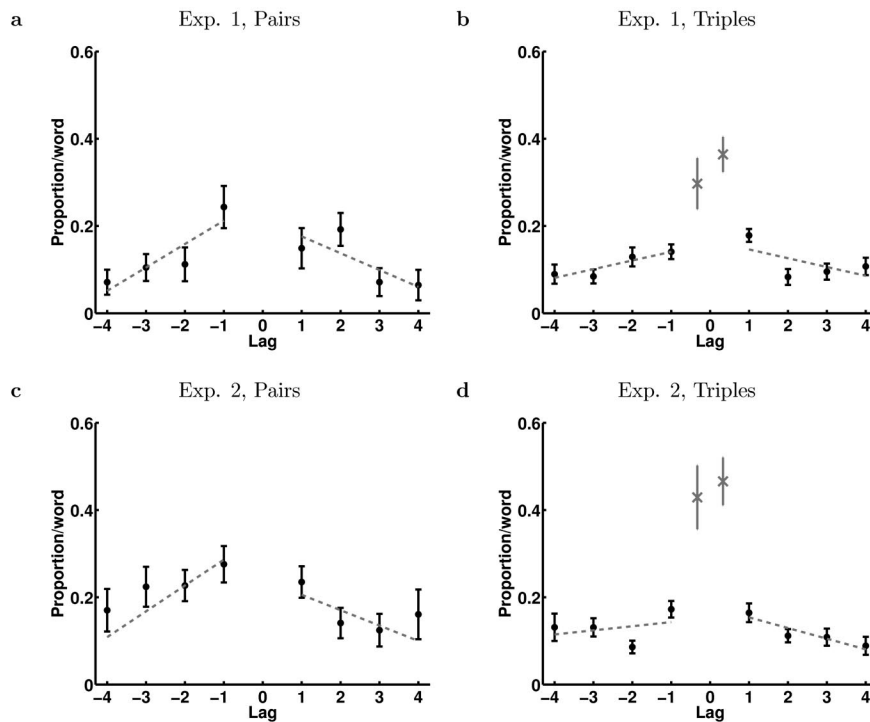


Figure 6. Within-list intrusions as a function of list lag. Lag refers to pair/triple number. Proportion is the ratio of the number of intrusions at a particular lag normalized by the total number of intrusions committed. These are corrected for the fact that not all lags were available for intrusions on each probe. a: Experiment 1, pairs. b: Experiment 1, triples. c: Experiment 2, pairs. d: Experiment 2, triples. Points marked with X (gray error bars) denote the rates of intrusions to the triple-lure items (at lag =  $-1/3$ : intrusions to the A item on probes of the BC subpair; at lag =  $+1/3$ : intrusions to the C item on probes of the AB subpair). Dashed gray lines denote the average regression fit. Error bars denote *SEM*.



on each probe. This was done by considering each probed recall question individually; if a lag was available, the availability-counter tally for that lag was incremented; if unavailable (i.e., either the probe or target item, or off the edges of the study set), the availability-counter tally for that lag was not incremented. The lag of the intruded item was incremented in a separate response counter, and the resulting response-counter tally was divided by the availability-counter tally at each lag. The first of these two points plots the rate of intrusions to A for BC probes, and the second plots the rate of intrusions to C for AB probes; both collapse across probe direction. This represents responses to  $14 \pm 12\%$  of all probed triple blocks. All but 2 participants committed one or more such triple-lure intrusion.

**Prior-list intrusions.** Intrusions came from prior studied lists 181 times, or  $1.0 \pm 1.1\%$  of responses to all probes. Thirty-five of the 42 participants committed one or more prior-list intrusion. We asked whether prior-list intrusions to pairs tended to come from pairs and whether intrusions to triples tended to come from triples. Table 1 shows the rates of each possible transition (pair–pair, pair–triple, triple–pair, triple–triple). Yule’s  $Q$  gives an estimate of the degree to which intrusions come from the same block type compared with the opposite block type. The mean value of  $Q$  is in fact negative, and the confidence intervals do not overlap zero. This indicates that participants intruded an item from a pair to a triple probe and vice-versa at higher than chance levels, arguing against the notion that they had access to contextual information about whether probe items came from lists of pairs or lists of triples. Note that the numbers of pair lists previously presented were not equal to the numbers of triple lists previously presented. Thus, if participants were drawing prior-list intrusions at random, the marginal probabilities (overall probability of recalling pairs and overall probability of recalling triples) would be different. However, Yule’s  $Q$  controls for this in a manner analogous to that of Pearson’s  $r$ .

**Within-list intrusions.** Intrusions came from other pairs or triples from the same list 751 times, or  $4.1 \pm 4.1\%$  of all probes. All participants committed one or more within-list intrusion. By inspecting the characteristics of these within-list intrusions, we might understand better how items within the list compete at retrieval.

**Contiguity and asymmetry in within-list intrusions.** We asked whether within-list intrusions came from nearby list lags (contiguity) and from later versus earlier list lags (asymmetry). Figure 6

shows the proportion of intrusions as a function of lag, for within-list intrusions, for probes of pair blocks (Panel a) and triple blocks (Panel b). These plots include only participants with enough data (a minimum of four intrusions overall and data on at least two lag values for negative and positive lags, respectively). Similar to studies of free recall (Kahana, 1996; Kahana, Howard, Zaromb, & Wingfield, 2002) and probed recall of long, 19-word lists (Kahana & Caplan, 2002, Figure 3), errors in both pairs and triples decrease with increasing list lag. We tested the significance of the contiguity effect by performing a linear regression over negative lags ( $-4$  to  $-1$ ) and positive lags ( $1$  to  $4$ ) for each participant and then computing a two-tailed  $t$  test between fit slopes (averaging together the slopes for positive and negative lags).

For pairs, the average slope was  $5.4 \pm 1.9\%$  per lag (here a lag represents a whole pair) and was significantly nonzero,  $t(16) = 3.64$ ,  $p < .005$ . For triples, the average slope was  $2.0 \pm 0.6\%$  per lag (here a lag represents a whole triple) and was also significantly nonzero,  $t(33) = 3.33$ ,  $p < .005$ .

We tested for asymmetry by comparing the value of the average regression line at lag =  $+1$  with that at lag =  $-1$ . This test was not significant for pairs,  $t(16) = -0.66$ , or for triples,  $t(33) = 0.27$ . Thus, contiguity effects were found, but no asymmetries.

According to the isolation principle, in addition to these small contiguity effects, the strongest effects of contiguity should be found within the triple itself. As is evident in Figure 6b, intrusions to triple-lure items were far more frequent than intrusions to items from adjacent triples, for AB probes compared with items from the subsequent triple,  $t(70) = 6.9$ ,  $p < 10^{-7}$ , as well as for BC probes compared with the prior triple,  $t(71) = 4.1$ ,  $p < .0005$ . Within-triple asymmetry effects were not found, in that the C item was not more likely to be intruded for AB probes than was the A item for BC probes, collapsing across probe direction,  $t(80) = 0.94$ . There was also no greater likelihood of triple-lure intrusions depending on probe direction: A? vs. ?B,  $t(76) = 0.11$ ; B? vs. ?C,  $t(78) = 0.33$ , *ns*. These analyses found strong evidence for contiguity effects in intrusion patterns within triples, but more weak effects between triples or pairs, consistent with the notion that pairs and triples are relatively isolated from other list items, and thus are sources of interference outside the pair or triple itself.

**Summary.** Intrusions to probed recall of serial lists (Kahana & Caplan, 2002) tend to come from nearby list positions rather than from remote list positions (contiguity), and tend to come more

Table 1  
Probabilities of Intrusions Transitioning Between Pair and Triple Blocks

Probed block	Intruded item type					
	Experiment 1: Prior-list <sup>a</sup>		Experiment 2: Prior-list <sup>b</sup>		Experiment 2: Within-list	
	Pair	Triple	Pair	Triple	Pair	Triple
Pair	14	34	17	57	53	290
Triple	53	80	47	138	195	483
Yule’s $Q$	-0.23 (-0.40, -0.06)		-0.07 (-0.22, 0.10)		-0.38 (-0.45, -0.30)	

Note. The Yule’s  $Q$  values reflect the degree to which intrusions come from the same type (pair or triple) of block. 95% confidence intervals are in parentheses.

<sup>a</sup> Pure lists. <sup>b</sup> Mixed lists.

from later serial positions than from earlier serial positions (asymmetry). This phenomenon may be related to contiguity and asymmetry effects that are commonly found in analyses of recall transitions in free recall; that is, given that a participant has just recalled an item from a list, they tend to choose to recall nearby items next, and to predominantly choose later serial positions compared with earlier serial positions (Howard & Kahana, 1999; Kahana, 1996; Kahana et al. 2002). Although the isolation principle is neutral with respect to the asymmetry, it is consistent with contiguity effects, but it predicts that these should be smaller than intrusions to the triple-lure items because isolation between pairs or triples causes positional similarity to decrease rapidly in positional coding models, or stored associative strength to decrease rapidly in chaining models. Contiguity effects were found for both pairs and triples, and the triple-lure items were intruded far more often than any other list item (Figures 6a and 6b), supporting the isolation principle. Asymmetry was absent, perhaps obscured by contiguity effects, perhaps for the same reason as for mean accuracy and response time: the lack of repeated presentations of pairs and triples.

Finally, we wanted to test whether participants were switching strategies between pair and triple lists (cf. Murdock & Franklin, 1984). If this were the case, we reasoned that this “task set” might be accessible as contextual information. Thus, participants might be able to distinguish whether candidate words were studied in pair lists or in triple lists. What follows from this is that in probes of pairs, intruded items should come more often from pairs than from triples, and more often from triples than from pairs for probes of triples. However, the analysis suggests that this is not the case (Table 1). Thus, it is unlikely that participants were drastically switching study strategies for the two list types, although we cannot rule this out entirely.

### Experiment 2: Mixed Lists of Pairs and Triples

In Experiment 1 we found that the patterns of accuracy, response times, and within-list intrusions were similar between lists of pairs and lists of triples. The exceptions are consistent with predictions one would make as a result of unifying associative and serial list memory via the isolation principle. We asked whether the differences between cued recall of pairs and triples would still be present when participants were forced to treat pairs and triples more similarly. Thus, in Experiment 2 we followed up with mixed lists, containing both pairs and triples. Although the findings of Experiment 1 rule out between-experiment differences, it is still possible that participants substantially altered their strategies between lists, and those alterations might have produced our observed differences. Thus, in a within-list manipulation, the distinction between pairs and triples might vanish.

We also conjectured that if participants had the opportunity to switch strategies between study of pairs and triples, they might employ extra study processes that would help to shield triples from triple-lure interference effects. To this end, Experiment 2 includes two participant groups, one informed regarding whether the next word cluster would be a pair or a triple, and the other not informed. We expected that the reduction in forward–backward correlation for triples (compared with pairs) would be greater for the uninformed group.

## Method

### Participants

Fifty-five volunteers participated for monetary compensation. One participant gave no correct responses (perhaps not having understood the instructions) and was excluded on that basis. Four participants were excluded because they had ceiling or floor performance in at least one condition (AB/AB/BC  $\times$  Forward/Backward). Of the 50 remaining participants, half were randomly assigned to the informed group (17 women, 8 men, mean age  $\pm$   $\sigma$  = 27.9  $\pm$  9.6) and half to the uninformed group (18 women, 7 men, mean age  $\pm$   $\sigma$  = 29.0  $\pm$  11.1). All participants spoke English as their primary language.

### Materials

The materials were identical to those in Experiment 1, with the following exceptions: First, all lists contained two pairs and four triples. This equalled the number of data points for each condition (pairs, AB\_\_ triples, and \_\_BC triples); because there were twice as many ways of probing triples, this required twice the number of triples per list. Second, we wanted to be able to track whether participants changed their study strategies on the basis of whether the words being presented were part of a pair or part of a triple.

Therefore, the two randomly formed groups were scored by an experimenter who was blind to condition. The informed group viewed an informative string before each pair (\*\*2\*\*) or triple (\*\*3\*\*). In this condition, analogous to Experiment 1, participants knew from the onset of the first word in a block whether the block was a pair or a triple. Thus, they had the opportunity to shift study strategy from one pair or triple to the next. The uninformed group saw an uninformative string (\*\*\*\*\*) before each pair or triple. Thus, these participants may not have known for certain whether the first two words in each block comprised a pair or part of a triple, until the third word (or an uninformative string) appears. The uninformed participants were expected to have a more consistent study strategy, at least during presentation of the A and B words.

### Procedure

As illustrated in Figure 7, the procedure for a single trial was identical to that in Experiment 1, except that the lists were constructed differently. Also, in Experiment 1, some pairs were not probed at all; in Experiment 2, all pairs and all triples were tested. The first two lists are considered practice and are excluded from all analyses.

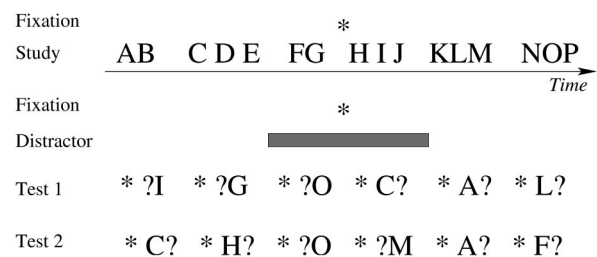


Figure 7. Procedure for a single trial, Experiment 2 (mixed lists). Each capitalized letter represents a unique noun. The asterisk (\*) denotes a fixation presented for 3,750 ms and then erased for 250 ms. In a study trial, words are presented grouped in pairs or in triples. A mathematical distractor follows (denoted by the gray rectangle), and then two complete probes of the list. Each pair/triple is probed once in Test 1 and then once again in Test 2. For the hypothetical example shown here, the pair AB and the triples CDE and NOP were probed in the same direction on Test 1 and Test 2, whereas the pair FG and the triples HIJ and KLM were probed in the opposite direction on Test 1 and 2.

### Results and Discussion

We analyzed mean performance (accuracy and response times) to assess potential sources of confounds and overall difficulty differences, correlations between successive tests to evaluate our chief predictions, and intrusions to test for further dissociations and similarities between memory for pairs and memory for triples.

#### Mean Performance Measures

Accuracy and response time were analyzed separately as were serial position (binned into thirds to compare with Experiment 1) and probe position. Accuracy is plotted as a function of serial position bin (Figure 8) and probe position (Figure 9) for both groups.

*Serial position bin analysis.* Experiment 2 has the additional between-subjects factor group (informed vs. uninformed). We performed ANOVAs on Serial Position Bin (3)  $\times$  Test Number (2)  $\times$  Pair/Triple Type (3)  $\times$  Probe Direction (2)  $\times$  Group (2).

For accuracy, main effects of serial position bin,  $F(2, 84) = 7.6$ ,  $MSe = 0.132$ ,  $p < .05$ ; test number,  $F(1, 48) = 99.5$ ,  $MSe = 0.022$ ,  $p < .001$ ; and pair/triple type,  $F(2, 95) = 54.3$ ,  $MSe = 0.11$ ,  $p < .001$ , were found. Post hoc pairwise  $t$  tests found greater accuracy on probes of early (first third) serial positions than middle or late (middle and last third) serial positions (sloped curves in Figure 8). In Experiment 1, participants were more accurate on Test 2 than on Test 1, suggesting

output encoding (dashed vs. solid lines in Figures 8a–8c, and 8g–8i). They were also more accurate on pair probes than on both triple probes, on AB\_\_ triples than on \_\_BC triples (cf. Figure 8), and on AB triple probes than on BC triple probes ( $p < .05$ ). The following interactions were significant: Serial Position Bin  $\times$  Test Number,  $F(2, 93) = 4.54$ ,  $MSe = 0.011$ ,  $p < .05$ ; and Test Number  $\times$  Probe Direction,  $F(1, 48) = 7.04$ ,  $MSe = 0.058$ ,  $p < .05$ . Simple effects analysis found a significant effect of probe direction on Test 2,  $F(1, 48) = 6.04$ ,  $MSe = 0.059$ ,  $p < .05$ , with greater accuracy for backward than forward probes, but no effect of probe direction on Test 1. This provides some evidence of a violation of symmetric mean performance, even for pairs, on Test 2. The Test Number  $\times$  Pair/Triple Type  $\times$  Probe Direction interaction was also significant,  $F(2, 96) = 5.4$ ,  $MSe = 0.032$ ,  $p < .01$ . Simple effects revealed a significant effect of probe direction for AB\_\_ triples in Test 1,  $F(1, 48) = 4.44$ ,  $MSe = 0.046$ ,  $p < .05$ , with an advantage for forward probes, and at Test 2,  $F(1, 48) = 10.73$ ,  $MSe = 0.051$ ,  $p < .01$ , with a backward-probe advantage.

Response time effects mostly paralleled those for accuracy, absent a main effect of serial position bin. The four-way interaction of Serial Position Bin  $\times$  Test Number  $\times$  Probe Direction  $\times$  Group was significant,  $F(2, 90) = 3.3$ ,  $MSe = 0.35 s^2$ ,  $p < .05$ ; no simple effects reached significance. The three-way interaction of Pair/Triple Type  $\times$  Probe Direction  $\times$  Group was also significant,  $F(2, 89) = 5.39$ ,  $MSe = 0.69 s^2$ ,  $p < .01$ . Simple effects found a

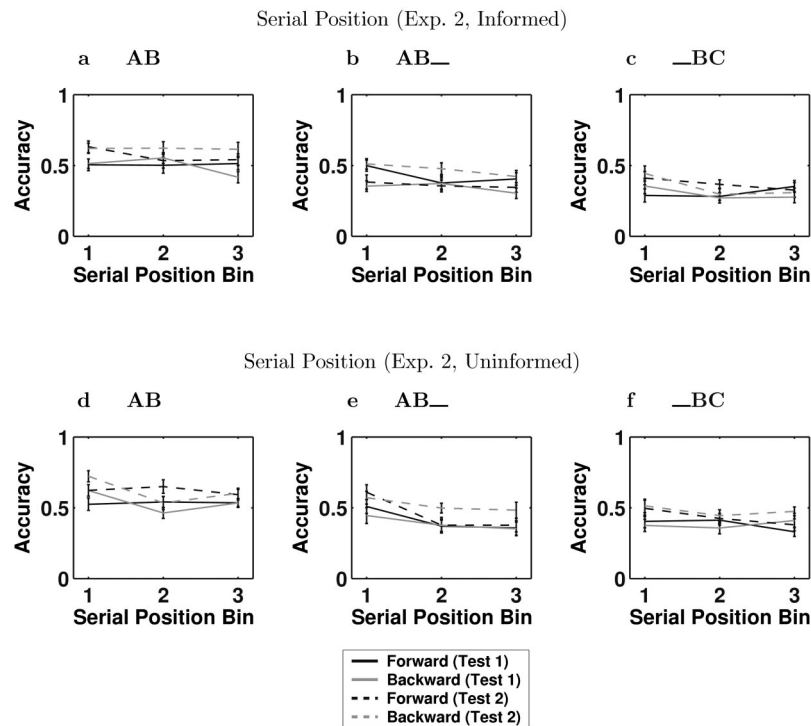


Figure 8. Experiment 2, serial position curves: Accuracy as a function of serial position, probe type (columns), Test 1 versus Test 2 (solid vs. dashed lines, respectively), and forward versus backward probe direction (black vs. gray lines, respectively). Serial position represents binned serial positions (1st, 2nd, and 3rd third of the list). Panels a–c are for the informed group; Panels d–f are for the uninformed group. Error bars are *SEM* across participants, corrected for between-subjects variability (Loftus & Masson, 1994).

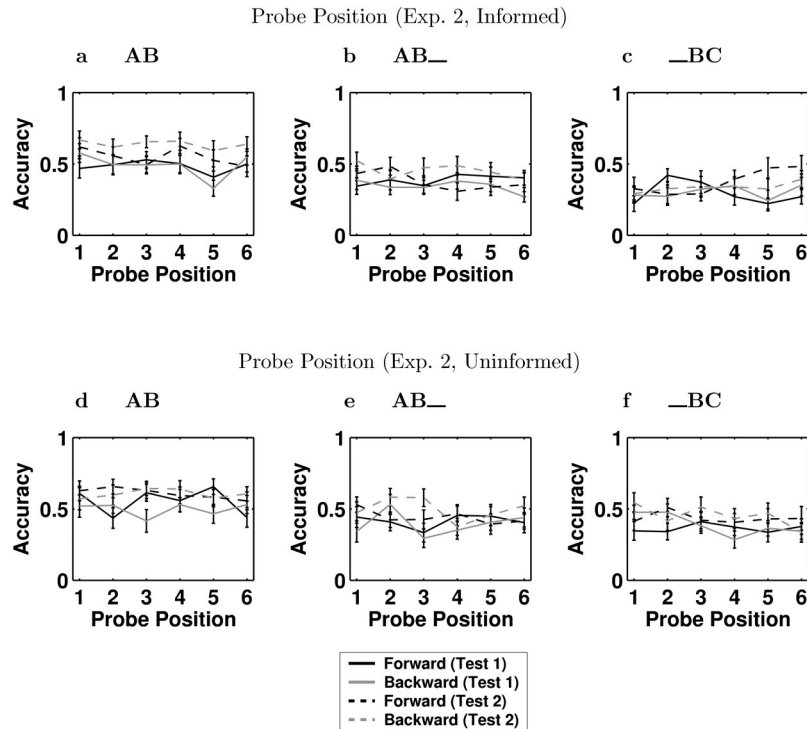


Figure 9. Experiment 2, probe position curves: Accuracy as a function of probe position, probe type (columns), Test 1 versus Test 2 (solid vs. dashed lines, respectively), and forward versus backward probe direction (black vs. grey lines, respectively). Panels a–c are for the informed group; Panels d–f are for the uninformed group. Error bars are *SEM* across participants, corrected for between-subjects variability (Loftus & Masson, 1994).

significant effect of probe direction for the informed group's pairs and for the uninformed group's  $\_BC$  triples, both with a backward-probe advantage.

*Probe position analysis.* We performed an ANOVA on Probe Position ( $6 \times$ ) Test Number ( $2 \times$ ) Pair/Triple Type ( $3 \times$ ) Probe Direction ( $2 \times$ ) Group ( $2$ ). The only significant effects in this analysis replicated the serial position bin analysis on both accuracy and response time measures.

### Summary

A main effect of serial position bin was found (sloped curves in Figure 8); accuracy decreased with increasing serial position bin. Thus, serial position bin is responsible for some variance, which would have contributed to the control correlation in the correlational analyses. However, serial position bin did not interact with probe direction or with pair/triple type in accuracy, so it does not pose a challenge to the correlation dissociation between pairs and triples. Probe position showed neither main effects nor interactions with any factors, so we need not worry about collapsing across this factor.

As in Experiment 1, accuracy overall was highest on pairs, intermediate for  $AB\_\_$  triples, and lowest for  $\_BC$  triples; response time data reinforced this relation. In Experiment 1, the total list presentation time differed for pair lists versus triple lists. In Experiment 2, there was no such confound because all lists contained both pairs and triples; nonetheless, two differences remain. First, there were fewer pairs per list than triples (two versus four),

the reverse of Experiment 1. Second, if one assumes that most of the rehearsal for a given pair or triple occurs following its presentation, then the effective amount of rehearsal time per word is still greater for pairs than for triples ( $4 \text{ s}/2 \text{ words} = 2 \text{ s}/\text{word}$  for pairs;  $4 \text{ s}/3 \text{ words} = 1.33 \text{ s}/\text{word}$  for triples). Thus, difficulty was not necessarily equated between pairs and triples, but it does not explain the difference between  $AB\_\_$  and  $\_BC$  triples, which may be an indirect result of a primacy effect across the entire study set, or a phenomenon akin to the scalloping of chunked lists (e.g., Brannon, 1997; Hitch, Burgess, Towse, & Culpin, 1996; Ng & Maybery, 2002), suggesting a parallel between our pairs and triples with subjective or induced groupings. Also, as in Experiment 1, performance was superior on Test 2 compared with Test 1 (dashed vs. solid lines in Figures 8 and 9), suggesting the use of output encoding.

The probe direction factor showed no main effects, supporting the overall finding of symmetric mean accuracy and response time. However, probe direction participated in two interactions. The three-way interaction in accuracy with test number and pair/triple type is somewhat problematic with regard to the correlation dissociations and should be borne in mind as a caveat, although it still does not directly speak to the question of the correlation between forward and backward probes, which is, mathematically, an independent measure. Interactions with probe direction represent evidence of occasional deviations from symmetric mean accuracy and response times in pairs and triples, in some cases with a forward advantage and in other cases with a backward advantage. Forward-

probe advantages were found in cued recall of triples and lists (Kahana & Caplan, 2002), but as argued above, those experiments involved several repetitions of the to-be-learned material. Thus, a strong forward bias may require multiple exposures to materialize (see the General Discussion for elaboration of this point).

Finally, there was no main effect of group (Figure 8, Panels a–c vs. Panels d–f and Figure 9, Panels a–c vs. Panels d–f), suggesting that overall task difficulty was matched. Group did interact with pair/triple type and probe direction, suggesting that the two groups used different study or cued recall strategies. This was further evident in the correlational analyses.

### Correlations on Successive Tests

We next examined the correlation between the two successive tests as a function of probe direction on Test 1 and Test 2. Figure 4b shows Yule's  $Q$  as a function of probe type and direction. Collapsing across groups, all  $Q_{\text{Diff}}$  values were significantly different from  $Q_{\text{Control}}$ : AB (pairs):  $z = 10.2$ ; AB:  $z = 9.3$ ; BC:  $z = 10.6$ ,  $p < .0001$ . All  $Q_{\text{Same}}$  values were significantly different from  $Q_{\text{Diff}}$ : AB (pairs):  $z = 4.5$ ; AB:  $z = 5.7$ ; BC:  $z = 5.3$ ,  $p < .0001$ .  $Q_{\text{Same}} - Q_{\text{Diff}}$  was significantly larger for AB\_\_ triples than for pairs ( $z = 3.2$ ,  $p < .005$ ), but not for \_\_BC triples ( $z = 1.1$ ,  $p > .1$ ).  $Q_{\text{Same}} - Q_{\text{Diff}}$  was significantly greater for AB\_\_ than \_\_BC triples ( $z = 2.5$ ,  $p < .05$ ). Thus, cued recall of triples taps directionally sensitive information to a greater degree than cued recall of pairs.  $Q_{\text{Control}}$  was greater for pairs than for triples, but these differences were not significant. This finding is followed up in the Simulation section. Response-time correlations did not show significant differences between pairs and triples and are not reported.

*Informed versus uninformed groups.* When the correlation analyses are broken down by group (Figures 4c and 4d), it becomes clear that the distinction between pairs and triples is not present for the informed participants; their  $Q_{\text{Same}} - Q_{\text{Diff}}$  values did not differ between pairs and both AB ( $z = 1.4$ ,  $p > .1$ ) and BC ( $z = -1.0$ ,  $p > .1$ ) probes. Consistent with this,  $Q_{\text{Same}} - Q_{\text{Diff}}$  was significantly different from zero ( $p < .01$ ) but had more similar effect sizes across the three conditions, AB (pairs):  $z = 3.7$ ; AB:  $z = 2.9$ ; BC:  $z = 3.0$ .

In contrast, for uninformed participants, this difference was significant for both AB probes ( $z = 2.8$ ,  $p < .005$ ) and BC probes ( $z = 2.1$ ,  $p < .05$ ), and  $Q_{\text{Diff}}$  was significantly different from  $Q_{\text{Same}}$  for AB ( $z = 4.8$ ,  $p < .0001$ ) and BC ( $z = 4.3$ ,  $p < .001$ ) triples and though significant, was weaker for pairs ( $z = 2.5$ ,  $p < .05$ ).

### Summary

When collapsed across groups,  $Q_{\text{Diff}}$  was significantly lower than  $Q_{\text{Same}}$  for pairs, AB\_\_ triples, and \_\_BC triples (Figure 4b). This, along with a similar finding in Experiment 1, shows that the finding of perfect forward–backward correlations in paired associates learning is not strictly perfect, but only nearly so. This slightly weakens the theory of associative symmetry but is quite consistent with the isolation principle, in which paired associates learning and serial list learning lie upon a continuum rather than being totally distinct paradigms. However, this reduced correlation for opposite probe directions was only significantly larger for AB\_\_ triples compared with pairs. When the analysis is broken down by participant group, the pattern becomes clearer. The informed group (Figure 4c) showed no dissociation in correlations

for pairs versus triples. On the other hand, the uninformed group (Figure 4d), with just over half the statistical power as in Experiment 1, showed a reliable difference between pairs and both AB\_\_ and \_\_BC triples.

What does the lack of dissociation for informed participants mean? First, it should be noted that a lack of dissociation is fully consistent with a unified model and thus does not challenge the isolation principle. However, the null finding points to an interesting boundary condition on the differentiation between pairs and triples. A more detailed inspection of the pattern of correlations for the informed group suggests that, rather than treating triples more like pairs, this group of participants treated pairs more like triples, to the point that associative symmetry broke down to a similar degree for pairs and for triples. In the framework of the isolation principle, this would reflect a decrease in the degree of isolation for pairs as well as for triples, which is consistent with the overall greater rate of within-list intrusions for informed compared with uninformed participants, despite a lack of differences in mean accuracy or response time between groups. As is demonstrated in the Simulation section, in conditions in which correlation differences between pairs and triples are found (Experiment 1 and uninformed participants in Experiment 2), the isolation principle stands as a plausible, parsimonious account of those differences. Furthermore, these differences are predicted precisely in settings in which pairs and triples are treated equally (i.e., definitely for uninformed participants but not necessarily for informed participants, who may have been shifting strategies more).

### Intrusions

Participants made a total of 2,235 intrusions, or  $50 \pm 35\%$  of all probes. Each participant made at least one intrusion.

*Triple-lure intrusions.* Of the 2,235 intrusions, 482 were the unprobed item of a triple. These are plotted as dashed error bars in Figure 6d (the first of these two points plots the rate of intrusions to A for BC probes, and the second plots the rate of intrusions to C for AB probes) and represent responses to  $16 \pm 16\%$  of all probed triples. All but 1 participant committed one or more such triple-lure intrusion.

*Prior-list intrusions.* Intrusions came from prior studied lists 259 times, representing  $1.4 \pm 1.8\%$  of responses to all probes. Forty-two of the 50 participants committed one or more prior-list intrusion. Table 1 shows the rates of each possible transition (pair–pair, pair–triple, triple–pair, triple–triple). Despite the modest amount of data available for this analysis, the confidence intervals exclude moderate-to-high positive and negative correlations, and the value of  $Q$  is in fact near zero. It is interesting that, as in Experiment 1, prior-list intrusions came predominantly from triples, regardless of whether the probe was from a pair or a triple. However, more triples than pairs were presented. If participants were drawing prior-list intrusions at random, the marginal probabilities (overall probability of recalling pairs and overall probability of recalling triples) would be different (namely, there are more triple-words available to intrude). However, Yule's  $Q$  controls for this in a manner analogous to that of Pearson's  $r$ .

*Within-list intrusions.* Intrusions came from other pairs/triples of the same list 1,021 times, representing  $5.7 \pm 4.9\%$  of all probes. All but 1 participant committed one or more within-list intrusion. By examining further these within-list intrusions, we can better pinpoint how within-list interference comes into play.

*Contiguity and asymmetry in within-list intrusions.* We asked whether within-list intrusions came from nearby list lags (contiguity) and from later versus earlier list lags (asymmetry). Figure 6 shows the proportion of intrusions as a function of lag, for within-list intrusions, for probes of pairs (Panel c) and triples (Panel d). These plots include only participants with enough data (a minimum of four intrusions overall and data on at least two lag values for negative and positive lags, respectively). Intrusions to both pairs and triples decrease with increasing list lag, as in Experiment 1. We tested for contiguity and asymmetry effects as was done in Experiment 1.

For pairs, the average slope was  $5.0 \pm 2.4\%$  per lag (here a lag represents a pair or triple) and was significantly nonzero,  $T(33) = 2.10$ ,  $p < .05$ . For triples, the average slope was  $1.7 \pm 0.5\%$  per lag and was also significantly nonzero,  $T(39) = 3.42$ ,  $p < .005$ . Interestingly, these slopes are close to the slopes estimated in Experiment 1.

We tested for asymmetry by comparing the value of the average regression line at lag = +1 with that at lag = -1. This test was significant for pairs,  $T(32) = -2.18$ ,  $p < .05$ , showing a greater number of intrusions to negative lags than to positive lags (a difference of  $9.3 \pm 4.3\%$ ), but not for triples,  $T(39) = 0.56$ . Thus, contiguity effects were found, as well as a barely significant advantage for intrusions to backward lags for pairs, but no asymmetry effects were found for triples.

Triple-lure intrusions were common (Figure 6d) and were more frequent than intrusions to items from adjacent triples, for AB probes compared to items from the subsequent pair/triple,  $T(91) = 7.4$ ,  $p < .0001$ , as well as for BC probes compared to items from the prior pair/triple,  $T(64) = 4.6$ ,  $p < .0001$ . Within-triple asymmetry was not found. Intrusions to the C item from AB probes were not more frequent than intrusions to the A item from the BC probes,  $T(96) = 0.40$ ,  $p > .5$ . Also, lure-item intrusion rates did not differ by cue direction, for either AB,  $T(89) = 1.1$ , or BC,  $T(93) = 0.67$ , probes.

Thus, asymmetry effects were largely absent for triples, but a backward intrusion advantage was found for pairs. Contiguity effects were found for both pairs and triples but were far more pronounced in comparisons between within-triple to extra-triple intrusions.

*Pair-to-pair and triple-to-triple intrusions.* We asked whether within-list intrusions to pairs tended to come from pairs and intrusions to triples tended to come from triples. Table 1 shows the rates of each possible transition (pair-pair, pair-triple, triple-pair, triple-triple). The value of  $Q$  is negative. Though the confidence intervals are large, they exclude zero and all positive values.

*Summary.* As in Experiment 1, a large number of intrusions came from the triple-lure item (neither the probe nor the target; Figure 6d, dashed error bars), providing further support for the isolation principle. The within-list intrusions again showed effects of lag but with an overall dominance of intrusions to the triple-lure item, supporting the predictions of the isolation principle in a mixed-list design.

### Simulation

Across the two experiments, the data analyses found probed recall of pairs and triples to be largely similar. Where they differ, we argued that these differences could be explained in a parsimo-

nious model (e.g., Caplan, 2004, 2005) that employs the same study and test processes for pairs and triples. The chief differences were:

1. The correlation between forward and backward probe performance was lower for triples than for pairs (pure lists design and the uninformed group of the mixed-list design), suggesting that forward and backward probes of triples tap partially different studied information.

2. Both pairs and triples show some evidence of contiguity effects in within-list intrusions, and triples show a very high rate of intrusions of the triple-lure item. This could be explained in a dual-framework approach, but it would be more concisely explained by a single-framework model, in which isolation produces not only the correlation dissociation but also relatively high rates of lure-item intrusion.

3. The "control" correlation is not exactly comparable for pairs and triples. Although this could be used to argue for separate frameworks, it is possible that the same isolation pattern could produce lower control correlations for triples due to greater overall within-list interference.

The analytic models derived by Caplan (2005) to explain these types of effects required several assumptions for tractability, including large list length and high dimensionality of item vector representations. Here we implement a simple model and simulate it to test whether the isolation principle could account for our empirical differences when the model is subject to more realistic constraints. We fit the data from the Experiment 2 uninformed group with a positional coding model simulation. We focus on this data set because it poses the greatest apparent challenge to the unified paired associates-serial list framework. The model had to use the same processes and parameter values to fit data on pairs and triples, apart from the grouping structure of the list. The isolation principle can be implemented in both chaining models and positional coding models (Caplan, 2005). Because modeling pairs within a positional coding model is more novel, and because we have already foreshadowed the present findings in a chaining model (Caplan, 2004) that was simulated prior to data collection, we focus on the positional coding model. We are certainly not ruling out chaining model accounts of our findings; indeed, Caplan (2004) predicted the correlation dissociation as well as some other aspects of the behavior pattern seen in the present paradigm. Furthermore, chaining model simulation fits to a subset of the present data performed with a level of success similar to that of the positional coding model reported here.

Rather than attempt to model the complex and potentially rich, detailed strategies participants likely use to learn pairs and triples of nouns, such as imagery and verbal mediators (Paivio, 1971), we focus on a more abstract level: the formal structure of the learned information, following in the tradition of the modeling work we have cited. More detailed models would thus generalize from the type of model we implement here. As a strength model of positional coding, it has no explicit vector representation for items but learns by assigning a positional code and a strength to each list item. Note that in this approach, pairs and triples are not unitized directly; instead, they acquire properties similar to unitization by virtue of their encoded positions being relatively near one another within pairs and within triples compared with positional codes between pairs and triples. At test, it probes with an item, retrieves its positional code, then updates the positional code in the desired

direction and probes with this new position to retrieve the item associated with the target position. At test, probing with position retrieves not only the item with the closest position, but also items that had been stored at nearby positions.

The model learns lists designed to parallel those learned by the participants. Each list contained two pairs and four triples. Retrieval is based on a winner-take-all selection criterion (here, the item with the strongest retrieval strength is the item recalled, given that it exceeds some minimum strength threshold) among all word pool items and incorporates output encoding (further learning of the retrieved associations during the cued recall phase). The model incorporates list-to-list variability by selecting the mean encoding strength from a Gaussian distribution once per list. Critically, the model has a single parameter,  $I$ , that controls the isolation between pairs and triples. In this model,  $I$  controls how similar the stored positions of items are depending on whether the items are within the same pair or triple, or cross between two pairs/triples. The effect is that the positions with which the two items of a pair are stored are highly similar to each other and much less similar to the positions of other items in the list. Similarly, the positions with which the three items of a triple are stored are mutually similar and highly dissimilar to the positions of other list items. Detailed simulation methods are reported in Appendix A, and simulation fitting methods are given in Appendix B.

### Simulation Results

*Fits.* The final parameter values for the best fitting model (lowest root-mean-squared deviation [*RMSD*] over all eight models; all genetic algorithm fits ended up with *RMSD* under 0.01) are listed in Table 2. The best fitting model fits the data well (*RMSD* = 0.008). Of great relevance to our hypothesis that the isolation principle could account for the present data, the model required a substantial degree of isolation (parameter  $I$ ) in order to fit the data. However, the value for  $I$  is not at ceiling. More concretely, we can compute the ratio of positional similarity between adjacent items that are within a pair/triple to the similarity between adjacent items that cross from one pair/triple to another, based on this value of  $I$ . From Equations 6 and 8 we can calculate this ratio:

$$\frac{\text{Sim}(\textit{between})}{\text{Sim}(\textit{within})} = e^{-I[(2-I)\tau]} = 0.197. \quad (5)$$

where *Sim(between)* and *Sim(within)* denote similarity between adjacent items that cross pair/triple boundaries or stay within

Table 2  
Parameter Values for the Model

Parameter	Value	Range
$N_{\text{pool}}$	<b>500</b>	—
$\mu_s$	<b>1</b>	—
$I$	0.88	0.001–0.999
$\sigma_s$	0.73	0.01–2.00
$\sigma_{\mu_s}$	0.56	0–2
$\theta$	0.09	0–1
$C_c$	0.32	0–2
$\tau$	0.95	0.1–4.0

*Note.* Values for the best-fitting models are reported. Bold parameters were fixed; others were free. See text for explanations of the parameters.

pair/triple boundaries, respectively, and  $\tau$  is a parameter that controls the overall steepness with which the similarity of positional codes falls off with lag. This means that retrieval of neighboring items within pairs or triples purely as a result of positional similarity will be about five times as strong as retrieval of neighboring items in other pairs or triples. This suggests that the isolation principle need not be implausibly exaggerated to account for dissociations between pairs and triples.

*Mean accuracy.* Data and best fitting model behavior for accuracy are shown in Figure 10. Model data are plotted for the best fitting parameter sets, rerun on 10,000 virtual lists. The model was fit explicitly to these values, and fit reasonably well, capturing the overall lower levels of performance on triples compared with pairs. The only way the model could produce this dissociation was through differential interference effects due to relative isolation, spanned by a single parameter. Pairs and triples had to be equally isolated from the rest of the list (using the same value for parameter  $I$ ), and items within pairs and triples had to be equally strongly isolated and have equally separated positional codes. This suggests that the accuracy advantage in pairs over triples can be accounted for as a simple consequence of the isolation principle, without the need to invoke any different study or recall processes. The model also fits symmetry in accuracy for pairs and triples.

### Correlations

Data and best fitting model behavior are shown in Figure 11. Even though the model was fit only to the contingency tables for Same correlations and Different–Control pair correlations, it also correctly produced the dissociation in the Different correlations, giving lower correlations for both AB\_\_ and \_\_BC triples than for pairs. The only way this pattern could have arisen was by the added interference produced by the third items of triples, which were not isolated from probe and target items.

The model also produced the dissociation in the control correlation that was hinted at in the data, with lower control correlation for triples than for pairs. It was necessary to include a single parameter,  $\sigma_{\mu_s}$ , to introduce list-to-list variability and thus produce nonzero control correlations, but no additional assumptions were necessary to produce the dissociation in control correlation between pairs and triples. To understand why, consider the effect of list-to-list variability given a competitive retrieval process. List-to-list variability increases or decreases the item-position strengths throughout a particular list. For pairs, this should introduce correlations between forward and backward probes. But for triples, the interference level for the triple-lure item (item-position strength) is also correlated with the probe and target encoding levels. Because the effect of the triple-lure opposes recall success, this tends to introduce a corresponding negative list-to-list correlation. The net result is that in the presence of list-to-list variability, the control correlation is lower for triples than for pairs. This leads to a testable prediction, namely, experimental manipulations that reduce list-to-list variability in encoding should also reduce the pair versus triple difference in control correlations.

Finally, the model produced lower correlations than the data for Different and Control conditions, even though the fitting procedure did not select directly for this effect. This suggests that this pair–triple dissociation arises from internal constraints of the model that are a simple consequence of isolation.

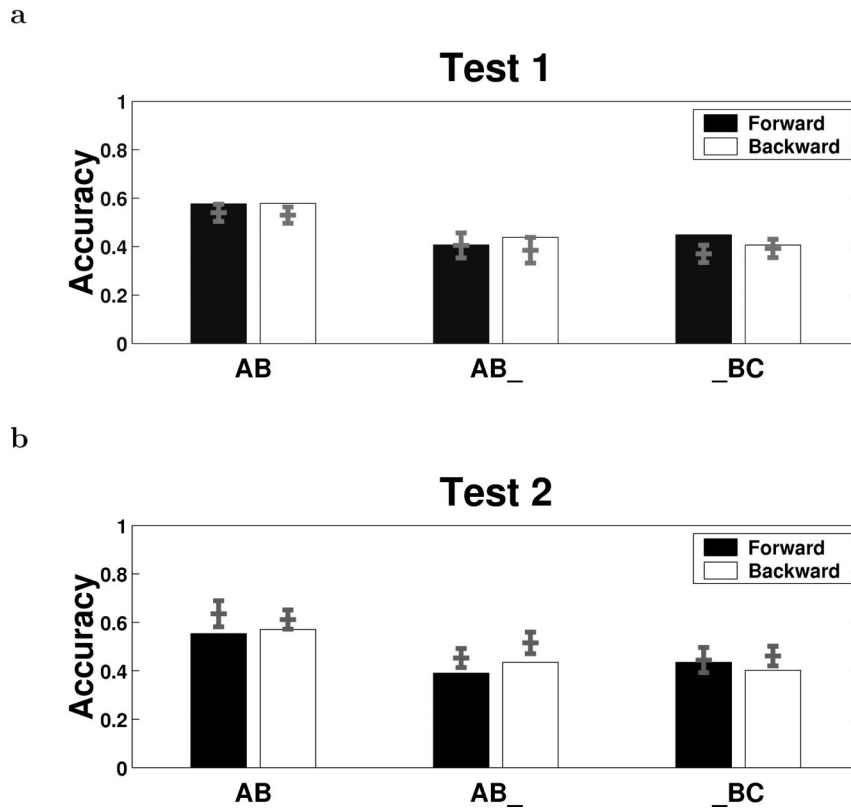


Figure 10. Simulation fits to mean accuracy. Data are plotted with error bars, simulation means (i.e., from model fits) are plotted as bars. Panels plot the results for the best-fitting parameter set, Tests 1 (a) and 2 (b). Error bars denote *SEM*.

### Within-List Intrusions

The simulated intrusions produced by the best fitting model are shown in Figure 12 (compare with Figure 6). Despite the fact that the model was not fit to the intrusion data at all, it captures the effects of contiguity and the comparably high level of lure-item intrusions from triple probes. This suggests that the differences in accuracy and correlation as well as the within-list intrusion pattern can be parsimoniously accounted for by additional interference due to the third, unisolated item present in triples but not in pairs.

### Parameter Values

We now discuss the values of the best fitting parameter set (Table 2). The model required a high level of isolation ( $I$ ) but it did not require perfect isolation. It fit to a modest level of encoding variability ( $\sigma_{\mu,s}$ ) and a substantial level of list-to-list variability ( $\sigma_{\mu,l}$ ). The response threshold ( $\theta$ ) was quite low but non-zero. Output encoding ( $C_e$ ) fit to a small, but non-zero level. Finally,  $\tau$ , which dictates the width of the positional similarity function, fit to approximately one position.

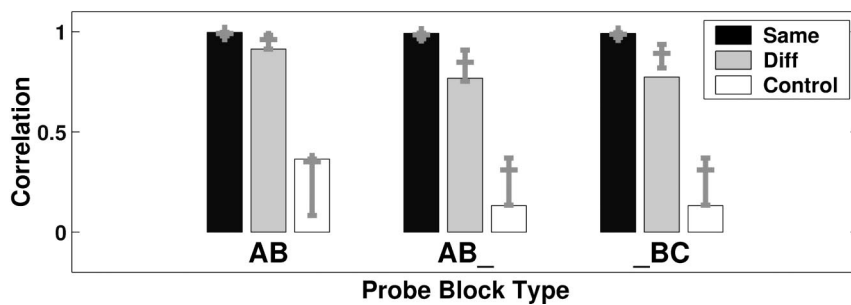


Figure 11. Simulation fits to correlation (Yule's  $Q$ ). Data are plotted with error bars; simulation  $Q$  values (from model fits) are plotted as bars. The panels plot the results for the best-fitting parameter set.



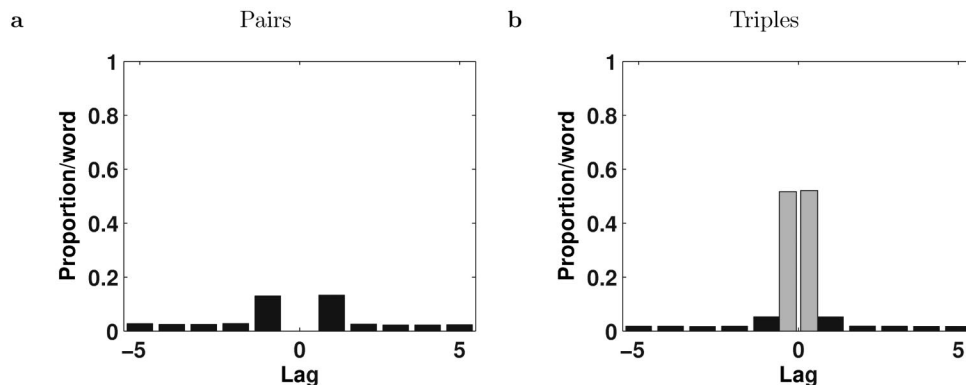


Figure 12. Simulation fits to within-list intrusions. Lag denotes position (in units of numbers of pairs or triples) of the intruded item relative to the probed item. Gray bars denote the rates of intrusions to the triple-lure items (at lag =  $-1/3$ : intrusions to the A item on probes of the BC subpair; at lag =  $+1/3$ : intrusions to the C item on probes of the AB subpair).

### Discussion

The model presented here is a specific implementation and limited in scope. It was designed only to account for cued recall behavior relevant to the present data, and would need to be modified to produce behavior in free recall, serial recall, or associative recognition paradigms or to handle multitrial presentation study phases. Nonetheless, the reasons it fit the data qualitatively and quantitatively are more general. The simulation fit is a practical test of the isolation principle. The simulation showed that it is reasonable to treat associative and serial list memory within a single model, with only the pattern of isolation differing between paradigms, as was introduced by Caplan (2004, 2005).

One advantage of this approach is that the unified model is more constrained, in that it must account for more patterns of behavior than would be required if one were to model memory for pairs and lists separately. This gives the modeler an increased ability to falsify specific assumptions and implementation choices. It could also lead experimentalists to design memory paradigms that dwell in the intermediate range between associations and lists as allowed by the mathematics of the isolation principle (moderate values of  $I$ ). The model fit also lends support to the notion that paired associates learning may be reasonably well modeled using positional coding models, as suggested by Caplan (2005). However, the chaining model account is by no means ruled out. Indeed, we suggest that paired associates learning and serial list learning may rely on multiple underlying processes. Thus, it is a strength of the isolation principle that it may be implemented in both major model classes; we also stress that it also could plausibly be implemented in a hybrid positional/chaining model.

### General Discussion

The tasks presented here enabled a closely matched comparison of memory for pairs and memory for short lists (triples). Whereas previous work suggested dissociations between associative and list memory (Kahana & Caplan, 2002), our findings represent the first direct evidence of such differences in two within-experiment comparisons. The chief finding is that forward and backward cued recall probes are nearly perfectly correlated for pairs but less

correlated for lists. This is the first time that the forward–backward probe correlations have been reported for triples. Also for the first time, we report the correlation in response time between forward and backward probes of both pairs and triples, supporting associative symmetry in pairs and suggesting a weakening of associative symmetry in triples. This extends the finding from the measure of accuracy to latency and helps to rule out speed–accuracy tradeoff as a confound. Other novel findings are the contiguity effect for within-list intrusions to both pairs and triples and the high rate of triple-lure item intrusions rising above these contiguity effects. This suggested that the third item is the principal culprit dissociating recall performance of pairs from triples. These dissociations were found in a pure-lists design (Experiment 1), and in a mixed-lists design (Experiment 2), with the exception that the dissociation in forward–backward correlation was present only for participants who were not informed whether each upcoming block of words were part of a pair or a triple (the uninformed group). The correlation dissociations, when present, were small in magnitude, as expected given that triples are very short lists. Finally, a positional coding model simulation showed that these dissociations can be accounted for parsimoniously using the same model processes and parameter values for pairs and triples. The simulation reproduced dissociations in correlations and within-list intrusion patterns without fitting to them directly, suggesting that these dissociations are consequences of the internal constraints of the model.

### Associative Symmetry

Probe direction had no main effect on accuracy or response time. This extends the classic findings of equal forward and backward probe accuracy to measures of response time and to pairs in the context of an experiment in which participants knew they would be studying both pairs and triples (lists). Interestingly, triples for the most part also showed symmetric accuracy and response times. In the only other study testing for asymmetries in probed recall of lists, Kahana and Caplan (2002) found forward-probe advantages in all probes of triples and 19-word lists. Thus, the findings of asymmetric accuracy and response time are by no means universal, but rather may relate to particular experimental

settings. For triples, asymmetries can arise from various types of associative interference or target ambiguity (Caplan, 2004; Kahana and Caplan, 2002). For pairs, breakdowns in symmetric means have been linked to differential treatment of the A versus B items (e.g., Horowitz et al. 1964, 1966; Kahana, 2000, 2002; Lockhart, 1969; Wollen, Fox, and Lowry, 1970). It is possible that, to some extent, the presence of triples induced participants to study or rehearse the A and B items of pairs differently.

Furthermore, the forward-probe advantages were found when participants had several exposures to a triple or list (Kahana & Caplan, 2002). Here we presented triples only a single time. It is possible that the strong directionality materializes only with repeated practice, as suggested for pairs by Waugh (1970). In particular, on repetitions, participants might recognize the repeated first ("A") item and then play a "serial anticipation" game for the remaining two items. This would be expected to progressively enhance what Horowitz and colleagues termed "item availability" (Horowitz et al., 1964, 1966). That is, participants would have more experience (covertly) producing the second ("B") items than first items, and the third ("C") items than the second items. Reinspection of Figure 1 from Kahana and Caplan (2002) supports this notion: The single-item probes can be ranked in accuracy (and in response latency) according to the serial position of the target item, independent of the probe item. Thus:  $B? > A? \cong ?C > ?B$ .

In addition, Kahana and Caplan (2002) presented triples sequentially, but with additional spatial cueing information. Thus, the A item was presented on the left side of the screen, the B item in the middle, and the C item to the right. This spatial cueing information was also present at test. Thus, a test probe of type A-? would have displayed the A word on the left portion of the screen, a series of question marks in the middle, and a blank box to the right. A probe of type B-? would have display the B word in the middle of the screen, question marks to the right, and a blank box to the left. Thus, participants had explicit positional information they could have used at test. Furthermore, this method might have encouraged participants to use explicit spatial codes, and the asymmetries observed in that dataset might have related to asymmetries in the use of the spatial location information. In the present experiments, no such spatial cueing was used, either at study or at test. Only the direction of probe was indicated by placement of the question marks either to the right (forward direction) or to the left (backward direction) of the probe item.

Kahana (2000, 2002) pointed out that mean performance does not provide a strong test of associative symmetry; instead, one must measure the correlations between forward and backward probes. The high correlations between forward and backward probes of pairs in the present study replicate this stronger test of associative symmetry (Kahana, 2002; Rizzuto & Kahana, 2000, 2001) in a setting in which participants were not only studying pairs, but also triples. The forward-backward correlation for triples, in contrast, was lower. The high correlation for pairs and the dissociation between pairs and triples breaks down somewhat in the informed group of Experiment 2. These participants had advance warning as to whether the upcoming chunk of words would be a pair or a triple. Although it is difficult to know why this is the case, we suggest that participants in this group failed to achieve isolation of pairs and triples as high as that of the uninformed group and the participants of Experiment 1. Perhaps they were attempting to study pairs and triples differentially but failed, and

ended up with fuzzier boundaries between pairs and triples. Given that the participants were informed only as to whether they were faced with a pair or triple at study, but were not similarly cued at test, this condition may have ultimately reduced their ability to keep pairs and triples separate from one another within the study sets.

### *Within-List Intrusions*

Participants made only a small number of intrusions to probes of pairs, replicating an old finding by Woodworth (1915) and extending it to probes of triples (within-list, extra-triple intrusions). Intrusions to pairs and triples showed evidence of contiguity effects, consistent with preliminary findings showing contiguity in intrusions to cued recall of pairs (Davis, Rizzuto, Geller, & Kahana, 2006), but little evidence of asymmetry effects. Far more common were intrusions to the lure item of triples; these were also relatively symmetric. This supports a key prediction of the isolation principle, that the dominant source of interference comes from the triple-lure item because it is not isolated from the tested subpair. An alternate account of triple-lure intrusions is that they result from failed attempts to recall the competing association, along the lines of recall-to-reject processes in associative recognition of pairs (Rotello & Heit, 2000), hence the strong lure item is sometimes selected when positional information within the triple is not sufficiently diagnostic. Furthermore, contiguity effects are consistent with certain classes of models, namely, those that at least partially rely on positional information or contextual information that is correlated in time, as well as chaining models that include more than nearest neighbor associations. Contiguity effects are inconsistent with models that rely on item-to-item associations that exclude remote associations or, more concretely, associations that cross from one pair or triple to another. The contiguity effect also rules out positional coding models that store position only within pairs and triples but don't represent overall position within the study session.

### *Triples as Short Lists*

According to the isolation principle, even very short lists containing three items should show effects of interference that differentiate them from probes of pairs. This prediction, which we confirmed, does not necessarily follow from approaches that would treat associative and serial list memory separately. There are many models that have been used to explain data on paired associates learning; for example, the theory of disturbed associative memory (TODAM; Murdock, 1982), the composite holographic associative recall/recognition model (CHARM; Eich, 1982), and the matrix model (Humphreys, Bain, & Pike, 1989). Likewise, there are numerous models of serial list memory, including oscillator-based associative recall (OSCAR; Brown et al., 2000), Scale-Invariant Memory Perception and Learning (SIMPLE; Brown, Neath, & Chater, 2006; Hulme, Surprenant, Bireta, Stuart, & Neath, 2004), the phonological loop model of Burgess and Hitch (1999), TODAM (e.g., Lewandowsky & Murdock, 1989), Serial Order in a Box (SOB; Farrell & Lewandowsky, 2002) the start-end model (Henson, 1998), the attractor network model of Jones and Polk (2002), the perturbation model (Lee & Estes, 1977), and the feature model (Nairne, 1990). One could

account for simultaneously acquired associative and serial list memory using a pair of such models, one for associations and one for lists. In such dichotomous models, one must still ask when a cluster of words turns from a holistic unit (as in pairs), which should have near-perfectly correlated forward and backward probe performance (Kahana, 2002; Rizzuto and Kahana, 2000, 2001), to a list, which requires a lower forward–backward correlation (Caplan, 2005). The most obvious assumption would be that triples are a hybrid between pairs and lists, but this hybridization would have to be further specified, resulting in a much more complex model. The isolation principle allows us instead to modify existing models of list learning or paired associates learning without substantially complicating them. The result is a model that can intrinsically shift from one paradigm to the other, including intermediate paradigms, and the relative position of memory for triples within this continuum is determined by the overall framework. The isolation principle framework provided a highly constrained means of explaining the present probed recall data on pairs and triples using the same processes and parameter values.

### *Rehearsal and Isolation*

The long interpair and intertriple intervals are unfilled and therefore permit participants to rehearse considerably. It is thus possible that this rehearsal process is what produced the dissociations in the first place. In particular, as proposed previously (Caplan, 2005), forward and backward association strengths within pairs may become highly correlated due to participants rehearsing A–B–A–B–A–B . . . ; in this type of situation, the participant rehearses the pair in the backward direction (B–A) nearly as many times as in the forward direction (A–B). Rehearsal of the entire triple (A–B–C–A–B–C . . .) would clearly differ. Fortunately, this account leads to one particular hypothesis: As a pair (or list) is studied over more presentations, the forward–backward correlation should increase. However, no effect of the number of presentations was found, either for probes of pairs (Kahana, 2000; Rizzuto & Kahana, 2001) or for probes of long lists (Caplan, 2005).

### *Chunking and the Control of Isolation*

The continuum notion of the isolation principle could be extended to the phenomenon of *chunking*, in which participants group together long lists into smaller subunits (Bower, 1969; Brannon, 1997; Martin & Noreen, 1974). Bower (1969) in fact suggested that chunks act to isolate subgroups of words from interference. Although participants appear to create their own idiosyncratic chunks, grouped presentation tends to induce participants to draw consistent, experimentally controlled chunk boundaries. This could lend itself to follow-up experiments involving probed recall of serial lists with grouped presentation, as suggested by Caplan (2005). The existing literature on chunking already shows some findings consistent with the isolation principle notion presented here. In particular, order errors are more common within chunks than between chunks, and intrusions that cross chunk boundaries are more rare than equivalent relative-position intrusions in ungrouped lists (e.g., Lee & Estes, 1981; Ng & Maybery, 2002). These phenomena may be related to the present finding of

extremely high rates of intrusions to the “lure” item of triples in cued recall (as compared with serial recall; see Figures 6b and 6d) compared with intrusions of other items in study set.

A more general question is what controls *I*, the level of isolation? It is unlikely that the grouped timing passively produces the level of isolation required to account for our findings, given that temporal isolation per se does not influence serial recall accuracy or error patterns (Lewandowsky & Brown, 2005; Ng & Maybery, 2002). However, the points made in the present manuscript are neutral with regard to the precise mechanism that controls the level of isolation. It could be under participants’ control; as, for example, Hockley and Cristi (1996) show, participants’ studied information differs as a function of whether they expect an item test or an order test. Alternatively, differences in *I* may be induced more passively via properties of the study session (e.g., temporal presentation schedule), as in the chunking studies cited above. Rather than having a priori expectations of under which particular circumstances *I* will be large or small, we suggest that *I* is a free parameter that could be fit to experimental data acquired under a variety of experimental conditions and even potentially used to explain individual variability. However, this account is more constrained than a dual-model approach in that the model is only allowed to adjust a single parameter rather than entirely switching regimes between study of associations and study of lists to invoke distinct models with independent parameters. Furthermore, *I* simply modulates an existing mechanism, rather than requiring different encoding mechanisms between the two paradigms. In a realistic positional coding model, the positional codes are variable anyway, and *I* simply acts to modulate this positional variability.

### *Double-Function Paired Associates Lists*

In double-function paired associates learning experiments, participants learn lists in which each item serves as the A item in one pair and the B item in another (Horowitz et al., 1966; Primoff, 1938; Slamecka, 1976; Stark, 1968). For example, a participant might study the list {A–B, D–E, B–C, C–D}. The isolation principle predicts that probed recall of such lists should result in performance quite similar to cued recall of lists. As proposed by Caplan (2005), forward and backward probes are susceptible to interference from different pairs; if pairs are encoded in separate, independent operations, this reduces the correlation between forward and backward probes relative to probes of standard, single-function pairs. If double-function pairs failed to exhibit lower forward–backward correlations than single-function pairs, the generality of the isolation principle would be questioned.

### *Associative Chaining Versus Positional Coding*

The model applied here was a positional coding one. This class of model has recently been shown to account for findings that the other class of model, associative chaining (e.g., Ebbinghaus, 1885/1913; Lewandowsky & Murdock, 1989), cannot explain. Simple associative chaining models have been challenged by empirical findings (e.g., Baddeley, 1968; Henson, Norris, Page, & Baddeley, 1996; Wickelgren, 1966), which is why we focus on a positional coding model in this article. However, it is not clear that more complex chaining models (e.g., with remote associations) could not overcome such challenges. Furthermore, in circumstances in

which position has little diagnostic value, like circular lists in which positional information is disrupted (Addis & Kahana, 2006), chaining strategies may dominate participants' strategies. Most likely participants use a variety of cues to learn and retrieve serial lists: chaining, positional, or ordinal (e.g., Giurintano, 1973; Maisto & Ward, 1976; Woodward & Murdock, 1968; Young, 1968). A complete model of serial list learning would most likely be a hybrid of chaining and positional coding. Caplan (2005) showed how the isolation principle may be built into chaining models as well as positional coding models, with a similar effect on the correlation between forward and backward probes. Caplan (2004) demonstrated how such a chaining model might be applied to data on probed recall of pairs and triples such as the data sets presented here. Thus, our conclusions are not necessarily limited to models that rely only on position or order coding but may also generalize to models that include associative chaining.

### *Theoretical Significance and Falsifiability*

We now ask how the present data findings and simulation fit provide new information relevant for modeling. First, our experiments might have failed to dissociate pairs from triples, suggesting that prior indications of dissociations were purely due to between-experiment differences. This would have, by default, supported the more parsimonious unified position. Second, the data obtained in the experiments could have been impossible to reconcile with a unified model, challenging the parsimonious, unified account. Third, if the simulation had successfully fit the data, it might have done so by requiring very extreme parameter values, thus questioning the model's plausibility. Instead, for example, the required level of isolation was not extreme at all, but instead was suggestive of a continuously modifiable encoding strategy.

A model is only useful insofar as it can be falsified. Although the isolation principle is more of a heuristic strategy for tracking interference within a broad range of models, it leads to strong constraints within those models. Thus, with future empirical work, it could be challenged as an explanation of dissociations between associative and serial list memory. In particular, for the positional coding model presented here, it was necessary that the strength of association from an item to a position be nearly perfectly correlated with the corresponding association strength from the position to the item. Experimental evidence contradicting this high correlation would make it impossible to model the high correlation in pairs, thus rendering the isolation principle unable to model memory for pairs and lists simultaneously. An analogous constraint applies to associative chaining models (e.g., Caplan, 2004, 2005): that the strength of the forward associations stored between a pair of items must be nearly perfectly correlated with the corresponding backward association. As Rizzuto and Kahana (2000, 2001) showed, this property is necessary to model the high correlation in forward and backward probes of paired associates.

As discussed in the previous section, a failure to dissociate the forward-backward probed recall correlation for double-function from single-function paired associates lists would cast doubt on the isolation principle as an account of dissociations between associative and serial list memory.

Finally, the isolation principle notion of treating memory for associations and lists the same may apply only for certain experimental regimes. Thus, experiments that seek boundary conditions on the applicability of the isolation principle to various experimental parameters might find that it fails to account for specific paradigms.

### *Conclusion*

We found direct support for a distinction in forward and backward probe correlations between associations and lists by applying a successive testing paradigm to probed recall of pairs and, for the first time, triples. This demonstrated the robustness of the apparent breakdown in holistic coding when participants move from study of associations to study of lists. The model simulation shows that this dissociation, in addition to other dissociations (contiguity of within-list intrusions and correlation due to list-to-list variability) can be accounted for by parsimonious models that treat association and list learning identically, if one allows pairs and lists to be relatively isolated from other studied items. This suggests that it is still possible to account for behavioral data on memory for associations in the same theoretical framework as memory for serial lists. Doing so will produce more highly constrained models and thus lead to greater insights about the structure of memory in both paradigms, as well as in intermediate paradigms.

### *References*

- Addis, K. M., & Kahana, M. J. (2006). *Positional uncertainty in serial learning*. Manuscript in preparation.
- Asch, S. E., & Ebenholtz, S. M. (1962). The principle of associative symmetry. *Proceedings of the American Philosophical Society*, 106(2), 135-163.
- Baddeley, A. D. (1968). How does acoustic similarity influence short-term memory? *Quarterly Journal of Experimental Psychology*, 20(3), 249-264.
- Bishop, Y. M. M., Fienberg, S. E., & Holland, P. W. (1975). *Discrete multivariate analysis: Theory and practice*. Cambridge, MA: MIT Press.
- Bower, G. H. (1969). Chunks as interference units in free recall. *Journal of Verbal Learning and Verbal Behavior*, 8, 610-613.
- Brannon, E. M. (1997). *Chunking in memory: Toward an operational definition of a chunk*. Unpublished master's thesis, Columbia University, New York.
- Brown, G. D. A., Neath, I., & Chater, N. (2002). *A ratio model of scale-invariant memory and identification*. Manuscript submitted for publication.
- Brown, G. D. A., Preece, T., & Hulme, C. (2000). Oscillatory-based memory for serial order. *Psychological Review*, 107, 127-181.
- 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.
- Caplan, J. B. (2004). Unifying models of paired associates and serial learning: Insights from simulating a chaining model. *NeuroComputing*, 58-60, 739-743.
- Caplan, J. B. (2005). Associative isolation: Unifying associative and order paradigms. *Journal of Mathematical Psychology*, 49(5), 383-402.
- Conrad, R. (1965). Order error in immediate recall of sequences. *Journal of Verbal Learning and Verbal Behavior*, 4, 161-169.

- Davis, O. C., Rizzuto, D. R., & Geller, A. S. & Kahana, M. J. (2006). *Temporal associative processes revealed by intrusions in paired-associates recall*. Manuscript submitted for publication.
- Detterman, D. K. (1977). A comparison of item and position probes in short-term memory. *American Journal of Psychology*, 90(1), 45–53.
- Ebbinghaus, H. (1885/1913). *Memory: A contribution to experimental psychology*. New York: Teachers College, Columbia University.
- Eich, J. M. (1982). A composite holographic associative recall model. *Psychological Review*, 89, 627–661.
- Farrell, S., & Lewandowsky, S. (2002). An endogenous distributed model of ordering in serial recall. *Psychonomic Bulletin & Review*, 9(1), 59–79.
- Friendly, M., Franklin, P. E., Hoffman, D., & Rubin, D. C. (1982). The Toronto Word Pool: Norms for imagery, concreteness, orthographic variables, and grammatical usage for 1,080 words. *Behavior Research Methods and Instrumentation*, 14, 375–399.
- Giurintano, S. L. (1973). Serial learning process: Test of chaining, position, and dual-process hypotheses. *Journal of Experimental Psychology*, 97(2), 154–157.
- Good, P. (2000). *Permutation tests: A practical guide to resampling methods for testing hypotheses*. New York: Springer.
- Hayman, C. A. G., & Tulving, E. (1989). Contingent dissociation between recognition and fragment completion: The method of triangulation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 228–240.
- Henson, R. N. A. (1998). Short-term memory for serial order: The start-end model. *Cognitive Psychology*, 36, 73–137.
- Henson, R. N. A., Norris, D. G., Page, M. P. A., & Baddeley, A. D. (1996). Unchained memory: Error patterns rule out chaining models of immediate serial recall. *Quarterly Journal of Experimental Psychology*, 49A, 80–115.
- Hitch, G. J., Burgess, N., Towse, J. N., & Culpin, V. (1996). Temporal grouping effects in immediate recall: A working memory analysis. *Quarterly Journal of Experimental Psychology*, 49A, 116–139.
- Hockley, W. E., & Cristi, C. (1996). Tests of encoding tradeoffs between item and associative information. *Memory & Cognition*, 24, 202–216.
- Horowitz, L. M., Brown, Z. M., & Weissbluth, S. (1964). Availability and the direction of associations. *Journal of Experimental Psychology*, 68(6), 541–549.
- Horowitz, L. M., Norman, S. A., & Day, R. S. (1966). Availability and associative symmetry. *Psychological Review*, 73, 1–15.
- Howard, M. W., & Kahana, M. J. (1999). Contextual variability and serial position effects in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25, 923–941.
- Hulme, C., Surprenant, A. M., Bireta, T. J., Stuart, G., & Neath, I. (2004). Abolishing the word-length effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30, 98–106.
- Humphreys, M. S., Bain, J. D., & Pike, R. (1989). Different ways to cue a coherent memory system: A theory for episodic, semantic, and procedural tasks. *Psychological Review*, 96, 208–233.
- Jones, M., & Polk, T. A. (2002). An attractor network model of serial recall. *Cognitive Systems Research*, 3, 45–55.
- Köhler, W. (1947). *Gestalt psychology*. New York: Liveright.
- Kahana, M. J. (1996). Associative retrieval processes in free recall. *Memory & Cognition*, 24, 103–109.
- Kahana, M. J. (2000). Contingency analyses of memory. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 59–72). New York: Oxford University Press.
- Kahana, M. J. (2002). Associative symmetry and memory theory. *Memory & Cognition*, 30, 823–840.
- Kahana, M. J., & Caplan, J. B. (2002). Associative asymmetry in probed recall of serial lists. *Memory & Cognition*, 30, 841–849.
- Kahana, M. J., Howard, M. W., Zaromb, F., & Wingfield, A. (2002). Age dissociates recency and lag recency effects in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28, 530–540.
- Lee, C. L., & Estes, W. K. (1977). Order and position in primary memory for letter strings. *Journal of Verbal Learning and Verbal Behavior*, 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: Learning, Memory, and Cognition*, 7, 149–169.
- Lewandowsky, S., & Brown, G. D. A. (2005). Serial recall and presentation schedule: A micro-analysis of local distinctiveness. *Memory*, 13(3/4), 283–292.
- Lewandowsky, S., & Murdock, B. B. (1989). Memory for serial order. *Psychological Review*, 96, 25–57.
- Lockhart, R. S. (1969). Retrieval asymmetry in the recall of adjectives and nouns. *Journal of Experimental Psychology*, 79(1), 12–17.
- Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review*, 1(4), 476–490.
- Maisto, A. A., & Ward, L. C. (1976). A test of the associative chaining hypothesis by serial-recall and serial-anticipation procedures. *Journal of General Psychology*, 94, 75–83.
- Martin, E., & Noreen, D. L. (1974). Serial learning: Identification of subjective subsequences. *Cognitive Psychology*, 6, 421–435.
- Mitchell, M. (1996). *An introduction to genetic algorithms*. Cambridge, MA: MIT Press.
- Murdock, B. B. (1962). Direction of recall in short-term memory. *Journal of Verbal Learning and Verbal Behavior*, 1, 119–124.
- Murdock, B. B. (1965). Associative symmetry and dichotic presentation. *Journal of Verbal Learning and Verbal Behavior*, 4, 222–226.
- Murdock, B. B. (1966). Forward and backward associations in paired associates. *Journal of Experimental Psychology*, 71(5), 732–737.
- Murdock, B. B. (1968). Serial order effects in short-term memory. *Journal of Experimental Psychology Monographs*, 76(4, Pt. 2), 1–15.
- Murdock, B. B. (1982). A theory for the storage and retrieval of item and associative information. *Psychological Review*, 89, 609–626.
- Murdock, B. B., & Franklin, P. E. (1984). Associative and serial-order information: Different modes of operation? *Memory & Cognition*, 12, 243–249.
- Nairne, J. S. (1990). A feature model of immediate memory. *Memory & Cognition*, 18, 251–269.
- Ng, H. L. H., & Maybery, M. T. (2002). Grouping in short-term verbal memory: Is position coded temporally? *Quarterly Journal of Experimental Psychology*, 55A, 391–424.
- Paivio, A. (1971). *Imagery and verbal processes*. New York: Holt, Rinehart and Winston.
- Posnansky, C. J. (1972). Probing for the functional stimuli in serial learning. *Journal of Experimental Psychology*, 96, 184–193.
- Primoff, E. (1938). Backward and forward associations as an organizing act in serial and in paired-associate learning. *Journal of Psychology*, 5, 375–395.
- Rizzuto, D. S., & Kahana, M. J. (2000). Associative symmetry vs. independent associations. *NeuroComputing*, 32–33, 973–978.
- Rizzuto, D. S., & Kahana, M. J. (2001). An autoassociative neural network model of paired-associate learning. *Neural Computation*, 13, 2075–2092.
- Rotello, C. M., & Heit, E. (2000). Associative recognition: A case of recall-to-reject processing. *Memory & Cognition*, 28, 907–922.
- Slamecka, N. J. (1976). An analysis of double-function lists. *Memory & Cognition*, 4, 581–585.
- Stark, K. (1968). Transfer from serial to paired-associate learning: A reappraisal. *Journal of Verbal Learning and Verbal Behavior*, 7, 20–30.
- Waugh, N. C. (1970). Associative symmetry and recall latencies: A distinction between learning and performance. *Acta Psychologica*, 33, 326–337.

- Wickelgren, W. A. (1966). Associative intrusions in short-term recall. *Journal of Experimental Psychology*, 72(6), 853–858.
- Wollen, K. A., Fox, R. A., & Lowry, D. H. (1970). Variations in asymmetry as a function of degree of forward learning. *Journal of Experimental Psychology*, 86(3), 416–419.
- Woodward, A. E. (1970). Continuity between serial memory and serial learning. *Journal of Experimental Psychology*, 85(1), 90–94.
- Woodward, A. E., & Murdock, B. B. (1968). Positional and sequential probes in serial learning. *Canadian Journal of Psychology*, 22(2), 131–138.
- Woodworth, R. S. (1915). A revision of imageless thought. *Psychological Review*, 22(1), 1–27.
- Young, R. K. (1968). Serial learning. In T. R. Dixon & D. L. Horton (Eds.), *Verbal behavior and general behavior theory* (pp. 122–148). Englewood Cliffs, NJ: Prentice-Hall.

## Appendix A

### Positional Coding Model Simulation

The positional coding model is a strength model with no explicit vector representation of items or positional codes. In a study trial, the model stores a list of  $L$  items by assigning each item  $\mathbf{f}_i$  a positional code,  $\mathbf{t}_i = (p_i, a_i)$  where  $p_i$  denotes a (scalar) position and  $a_i$  denotes a (scalar) strength. The values of  $p_i$  are determined according to the isolation parameter,  $I$ , where

$$I = 1 - \frac{s_w}{s_b}, \quad s_w + s_b = 1$$

$$p_{i+1} - p_i = \begin{cases} s_w & \text{within pair/triple} \\ s_b & \text{between pair/triple} \end{cases} \quad (\text{A1})$$

The strength values are drawn from a Gaussian distribution  $N(\mu_{\text{list}}, \sigma_s)$  where  $\mu_{\text{list}}$  is drawn once per list from a Gaussian distribution  $N(\mu_s, \sigma_{\mu_s})$  to implement list-to-list variability. Prior to study, all word pool items are initialized to zero strength.

At test, the model retrieves the positional code  $\mathbf{t}_x$  that was associated with the probe item  $\mathbf{f}_x$ . Then, the model adds or subtracts the value  $s_w$  from the retrieved position,  $p_x$ , for forward and

backward probes, respectively. Note that we are assuming that the model participant has direct access to the stored pattern of positional codes. Then, the model probes all items with the target positional code, obtaining a strength,  $\xi_i$  for each item in the pool:

$$\xi_i = a_x a_i \text{Sim}(p_x \pm s_w, p_i) \quad (\text{A2})$$

where positional code similarity is computed as

$$\text{Sim}(p_i, p_j) = e^{-|p_i - p_j|/\tau} \quad (\text{A3})$$

and  $\tau$ , a free parameter, sets the decay rate of similarity across nearby list positions. Note that this similarity function is commutative:  $\text{Sim}(p_i, p_j) = \text{Sim}(p_j, p_i)$ . Finally, a winner-take-all retrieval heuristic is applied to retrieve the word pool item,  $r$ , with the greatest retrieval strength, and is recalled if  $\xi_r > \theta$ , where  $\theta$  is a response threshold. If the model recalls an item, output encoding proceeds by incrementing the recalled item's strength  $a_r$  by an amount drawn from a Gaussian distribution  $N(C_e \mu_s, C_e \sigma_s)$ . Note that in this particular implementation of the positional coding model, the actual positions,  $p_i$ , are fixed from each study trial.

## Appendix B

### Simulation Fitting Methods

Simulation behavior (accuracy) was fit to the empirical values obtained from the participants. The values that were fit were all cells of the contingency tables for Test 1/Test 2 in the Same condition for pairs, AB\_\_ triples, and \_\_BC triples (4 quadrants  $\times$  3 types) and the contingency tables in the Control and Different conditions for pairs only (4 quadrants  $\times$  2 conditions) for a total of 20 data points. However, because of probability normalization, each contingency table has only 3 degrees of freedom; thus, the total degrees of freedom were 15, or 2.5 times the number of free parameters (6) in each model. Fitness was determined by root-mean-squared deviation.

Fits were performed in two stages using a genetic algorithm (Mitchell, 1996) with mutation and recombination. In the first stage, the parameter space was divided into 8 partitions and the genetic algorithm was run in each partition for eight generations in

order to allow the genetic algorithm for each partition to stabilize, with a generation size of 400 and a 50% survival rate. These partitions helped to ensure that the parameter space was evenly explored. The partitions were determined by dividing three parameters ( $I$ ,  $\tau$ , and  $\sigma_s$ ). Each parameter set was run on 1,200 virtual lists to estimate performance. In the second stage, each of the eight genetic algorithm partitions was run in the complete parameter space. Each parameter set was run on 15,000 virtual lists to estimate performance for 22 generations of size 800 with a 50% survival rate. The searched ranges for each parameter are listed in Table 2.

Received December 28, 2005

Revision received March 11, 2006

Accepted March 14, 2006 ■