Imagery-based strategies for memory for associations

Shrida S. Sahadevan

Department of Psychology and Neuroscience and Mental Health Institute, University of Alberta

Yvonne Y. Chen

Department of Psychology and Neuroscience and Mental Health Institute, University of Alberta; Currently: Department of Neurosurgery, Baylor College of Medicine

Jeremy B. Caplan

Department of Psychology and Neuroscience and Mental Health Institute, University of Alberta

Author Note

Corresponding author: Jeremy Caplan. jcaplan@ualberta.ca. Department of Psychology, Biological Sciences Building, University of Alberta, Edmonton, Alberta T6G 2E9, Canada, Tel:+1.780.492.5265, Fax:+1.780.492.1768.

The authors thank Tomi Ann A. Limcangco for assisting with earlier versions of the data analyses and Felicitas Kluger and Jeremy Thomas for helpful feedback on the manuscript. Supported by the Natural Sciences and Engineering Research Council of Canada and a studentship from the Undergraduate Research Initiative at the University of Alberta.

Abstract

Cued recall of word pairs is improved by asking participants to combine items in an interactive image. Meanwhile, interactive images facilitate serial-recall (Link Method), but even better when each item is imagined alongside a previously learned peg-word (Peg List Method). We asked if a peg system could support memory for pairs, hypothesizing it would outperform interactive imagery. Tested with cued recall, five study strategies were manipulated between-subjects, across two experiments: 1) Both words linked to one peg; 2) Each word linked to a different peg; 3) Peg list method but studying as a serial list; 4) Interactive imagery (within-pairs); 5) Link Method. Participants were able to apply peg-list strategies to pairs, as anticipated by mathematical modelling. Error-patterns spoke to mathematical models; peg lists exhibited distance-based confusability, characteristic of positional-coding models, and errors tended to preserve within-pair position, even for inter-item associative strategies, suggesting models of association should incorporate position. However, the peg list strategies came with a speed–accuracy tradeoff and did not challenge the superiority of the interactive imagery strategy. Without extensive practice with peg list strategies, interactive imagery remains superior for associations. Peg strategies may excel instead in tasks that primarily test serial order or with extensive training.

Keywords: Cued recall; association memory; visual imagery; interactive imagery; peg list method; mnemonic strategy

Imagery-based strategies for memory for associations

Introduction

Episodic memory for lists of word pairs (associations) has been extensively studied without instructing participants to apply any deliberate strategy. A lack of instruction does not imply that participants somehow apply no strategy; rather, their strategy is spontaneous, thus self-generated (Paivio & Yuille, 1969). The study strategy a participant selects could, in principle, influence overall accuracy, but also the logical organization of information in memory, for example, affecting the pattern of error responses. Paivio and colleagues identified two major classes of strategy, which they referred to as verbal and imagery mediation (Paivio, 1971). Here we focus on imagery mediation which, for verbal associations, refers to participants combining the two words into a single visual image. When the items are made to relate or interact with one another, this is specifically called interactive imagery (e.g., Bower, 1970; Bower & Winzenz, 1970; Hockley & Cristi, 1996a, 1996b; Lowry, 1974; McGee, 1980; Paivio, 1969, 1971), and is superior to rote repetition, but also, for example, sentence mediation (Bower & Winzenz, 1970). However, strategies based on interactive imagery directly between list items may not be superior for serial recall, compared to peg-list strategies (Roediger, 1980), as elaborated below. Also typically dependent on imagery, each image the memorizer forms combines a list-word with a peg from a pre-memorized peg list. Peg-based strategies, to our knowledge, have never been applied to lists of pairs. Our main goal was to test whether peg-based strategies have an association-memory advantage because participants can anchor new information onto a previously learned memory "scaffold," a hypothesis that emerges from several lines of research (summarized below). This hypothesis leads to the prediction that, as found by Roediger (1980) for serial recall, peg-based strategies should be more effective than interactive imagery, for association-memory as well, since interactive imagery links newly paired items and thus should have no benefit due to prior knowledge. Second, instructable strategies bear resemblances to mathematical models of memory, and enable us to test several specific hypotheses that derive from the modelling literature, detailed below. Before asking these theoretically motivated questions, we asked several pragmatic questions about application of the strategies of interest. Next, we elaborate each set of goals in turn.

Primary goal: generalization of peg-based strategies from research on serial recall

In research on serial recall, participants are asked to remember a list of items and reconstruct the list in correct order. In a pivotal study, Roediger (1980) found the method of loci (Yates, 1966) was superior for serial recall. In the method of loci, participants imagine placing visualizations of list items along a path through a familiar environment. However, nearly as effective in Roediger's experiment was the Peg List Method. To apply the Peg List Method (Bower, 1970; Bower & Reitman, 1972; Bugelski, 1968; Bugelski, Kidd, & Segmen, 1968; Bugelski, 1968; Carney & Levin, 2011; Delprato & Baker, 1974; J. E. Harris, 1980; L. J. Harris & Blaiser, 1997; Krinsky & Krinsky, 1994; Miller, Galanter, & Pribram, 1960; Paivio, 1971; Quinn & McConnell, 1996; Roediger, 1980; Wood, 1967), participants first master a "peg list," which includes one imageable word corresponding to each number over a range. For example, the peg list we use here (Figure 2a) is a set of words that rhyme with the numbers 1–10: 1–BUN, 2–SHOE, 3–TREE, 4–DOOR, 5–HIVE, 6–STICKS, 7–HEAVEN, 8–GATE, 9–VINE, 10–HEN (Lieberman, 2011). The participant can then learn a new list of words by forming an image between each new list item and the peg word (from the pre-learned peg list) corresponding to its numerical serial position. To recall the list in order, the participant runs through the peg list, using each peg item in turn as a cue to retrieve the item that had been linked to that peg word. The same peg list can be reused with many lists without losing its effectiveness (e.g., Bower & Reitman, 1972; Bugelski et al., 1968; Bugelski, 1968; Carney & Levin, 2011; Roediger, 1980; Ueno & Saito, 2013), suggesting it is not particularly susceptible to proactive interference and is quite effective relative to other known serial-recall strategies (Delin, 1969; Roediger, 1980).

Roediger (1980) also confirmed the efficacy of the Link Method, wherein participants combine all adjacent pairs of list items into interactive images. However, Roediger's Peg List participants substantially outperformed his Link Method participants in serial recall, particularly when recall was scored for strict order. This led us to hypothesize that the peg list method might likewise be a superior strategy for memory for lists of word pairs, which we tested here.

Theoretical motivation for the hypothesis that peg-based strategies will outperform those based on inter-item imagery. Numerous sources of evidence have suggested that knowledge can be acquired swiftly, even possibly bypassing the episodic memory system (or at least, medial temporal lobe and hippocampus), when new information can be anchored to previously mastered knowledge (O'Kane, Kensinger, & Corkin, 2004; Sharon, Moscovitch, & Gilboa, 2011; Skotko et al., 2004; Sommer, 2017; Tse et al., 2007, 2011). This provides a plausible theoretical account of Roediger's finding. It also provides theoretical motivation for our hypothesis that the peg list method should likewise outperform interactive imagery for lists of pairs. Namely, peg-based methods might be superior because the rhyming peg list is largely familiar to participants prior to the experiment. Although both link and peg methods rely on images combining two words together, only peg strategies involve binding list items each to a pre-memorized peg word, which could benefit from such anchoring effects.

Secondary questions motivated from mathematical modelling research

Due to their face-value resemblances to mathematical models, instructable, imagery-based strategies afford the opportunity to speak to current questions that have emerged in mathematical modelling research on both serial recall and association-memory. On the topic of serial recall, there is an ongoing debate about whether memory for a serial list is constructed from direct, inter-item associations ("associative chaining"), as first proposed by Ebbinghaus (1885/1913), or constructed only from associations between list items and a separate construct that provides order, so-called "positional-" or "order-coding" models, first proposed by Ladd and Woodworth (1911) and later by Conrad (1960). Thus, numerous positional/order-coding models, by design, exclude inter-item associations (e.g., Brown, Neath, & Chater, 2007; Burgess & Hitch, 1999; Farrell, 2012; Henson, 1998). On the other hand, many of the arguments against associative chaining have been challenged (e.g., Caplan, 2015; Farrell & Lewandowsky, 2003; Kahana, Mollison, & Addis, 2010; Lindsey & Logan, 2019; Solway, Murdock, & Kahana, 2012), and it is quite likely that a hybrid or mixture model is required to explain the full range of empirical benchmark findings (Caplan, 2015).

Moreover, Caplan (2005) showed how a positional coding model could be used to model memory for lists of pairs. The analytical and simulated implementations simply assumed that the model would use the cue word to retrieve the stored position code, then the position would be shifted by one position forward or backward, depending on the direction of the probe, and the shifted position used as a cue to attempt to retrieve the paired target item (see Howard, Jing, Rao, Provyn, & Datey, 2009 for a similar approach implemented within the Temporal Context Model). Thus, if it is the case that memory for a serial list does not rely on inter-item associations, the same might apply to lists of pairs. Without attempting to resolve the debate about uninstructed immediate serial recall, we simply observe that the link method resembles a very concrete implementation of associative chaining, whereas the peg list method resembles a very concrete and particular implementation of a positional-coding model.

We tested several hypotheses derived from research on mathematical models, as follows.

Feasibility of peg strategies for sets of pairs. Our first model-motivated question was whether Caplan's (2005) proposal that a list of associations could be well supported by a positional-coding model has any validity, or is complete fiction. We wondered whether participants would even be able to implement a peg list strategy with respect to a list of pairs, at any meaningful level of performance.

Contiguity effects. We were also interested in how memory for a list of pairs might differ depending on whether participants formed inter-item images or only item-peg images. We examined distance functions, both of cued recall and peg recall (Caplan, Glaholt, & McIntosh, 2006; Davis, Geller, Rizzuto, & Kahana, 2008). Contiguity effects have been presumed to emerge from item–position learning (e.g., Lee & Estes, 1977)— that is, when items are recalled in error, the errors tend to come from nearby positions. Henson, Norris, Page, and Baddeley (1996) called this the "locality constraint," and although associative chaining models can, in some implementations, meet this benchmark finding, the locality constraint has been argued to be diagnostic of variants of positional/order-coding models (e.g., Henson et al., 1996; Henson, 1998; Hurlstone, Hitch, & Baddeley, 2014). Our second model-motivated question was whether contiguity effects would emerge for peg-recall, a characteristic of current position/order-coding models, and likewise, whether contiguity effects would emerge for strategies based on inter-item imagery, which would test whether distance effects might be diagnostic of inter-item versus item–position associations.

Associative symmetry. Our third model-motivated question concerned the consistent finding of associative symmetry, that cued recall in the forward direction (for a studied pair AB, given A, recall B) and in the backward direction (given B, recall A) apparently test the same underlying variability in memory. That is, they are nearly equal overall (Asch & Ebenholtz, 1962), but more theoretically relevant, they are nearly perfectly correlated (Kahana, 2002). Caplan (2005) showed how this could be explained if memory for a pair was relatively isolated from the rest of the list, and how associative symmetry could appear to break down when substantial competition from other list items was introduced, such as when embedded within a serial list, or due to other pairs sharing an item (Caplan, Rehani, & Andrews, 2014; Kahana, 2002; Rehani & Caplan, 2011), as well as when associations gain internal order, such as compound words (Caplan, Boulton, &

Gagné, 2014). To test for associative symmetry, each pair was tested with cued recall twice, sometimes in the same direction (forward/forward or backward/backward) and sometimes in a different direction (forward/backward or backward/forward), enabling us to quantify the correlation between successive tests with Yule's Q (also known as the phi coefficient; Warrens, 2008). If associative symmetry holds, the value of Q in the "Different" condition would be nearly as high as for the "Same" condition. We hypothesized that the Link and Peg Serial List conditions would have elevated levels of interference from neighbouring list items, so their $Q_{\text{Different}}$ should be lower than in the Interactive Imagery and other two peg conditions. We also speculated that interactive images might incorporate order in such a way as to decouple forward and backward retrieval routes, raising the possibility that $Q_{\text{Different}}$ might be quite low in all conditions.

Within-pair position-cueing. Our fourth model-motivated question concerned the nature of the representation of associations. Experimental data have shown that despite associative symmetry, that is, the approximate equivalence of forward and backward cued recall, participants do nonetheless possess some memory for the relative positions of items within a pair (AB versus BA), but this knowledge is imperfect (Greene & Tussing, 2001; Kato & Caplan, 2017; Kounios, Smith, Yang, Bachman, & D'Esposito, 2001; Kounios, Bachman, Casasanto, Grossman, & Smith, 2003; Yang et al., 2013). Although both dominant models of association-memory, matrix models and convolution-based models, can produce associative symmetry (Kahana, 2002; Rizzuto & Kahana, 2001), they both make too extreme predictions about memory for within-pair order. Matrix models predict perfect order given that the pair is remembered and convolution models predict no memory for order. This has demanded that current models of association be expanded to incorporate some, but not perfect, within-pair order ability (Kato & Caplan, 2017; Rehani & Caplan, 2011). The analyses of within-list intrusions affords the opportunity to test whether the visual-imagery based inter-item association strategies (Interactive Imagery and Link Method) afford participants the ability to use within-pair position as a retrieval cue. If so, we should observe more intrusions to items sharing the same position as the target item than the other position, reminiscent of so-called "interpositions" between chunks (Henson, 1998). Such a result would extend the boundary conditions of the finding that within-pair memory for order is present, whereas the absence of such a result would suggest that convolution may be a sufficient model for inter-item associations learned via interactive imagery.

Initial considerations and questions regarding pragmatic aspects of peg-list strategies

Two ways to apply peg lists to pairs. Finding no precedents for the Peg-List Method applied to sets of pairs,[1] we devised two ways in which the pegs could be used (Figure 2b,c). In what we call the "One Peg/Pair" variant, the participant could study a pair, CAT–DOG, by forming an image of CAT with a peg (e.g., BUN) and a separate image between DOG and the same peg (BUN). In the "Two Pegs/Pair" variant, the participant could instead link each item to a different pair, so CAT would be linked to BUN and DOG would be linked to SHOE. We compared cued recall performance between participants using Interactive Imagery with participants using the Peg List Method.

Serial-list framing. Murdock and Franklin (1984) argued that participants adjusted how they organized their study processes depending on whether they expected a test of pairs versus serial-lists. We therefore wondered if it would make a large difference to cued-recall accuracy if participants were told that they were, in fact, studying a serial list rather than a set of pairs. To this end, we included two groups of participants who had the same experimental procedure, but were told they were studying a serial list (i.e., not a list of pairs). The Peg Serial List group used the peg list method but was told to conceive of the study set as comprising a serial list. The Link Method group formed interactive images, but according to standard Link Method instructions, were asked to form images

[1] although memory for pairs has been found to be successful with the method of loci, following three sessions of mnemonic training and tested with free recall; de Beni & Cornoldi, 1988.

between pairs as well as within pairs. If lists of pairs are encoded in distinct modes of operation than serial lists (Murdock & Franklin, 1984), we would predict a large disadvantage in cued recall for participants instructed they were studying serial lists. Alternatively, in case the peg-list strategies underperformed relative to interactive imagery, it could be that conceiving of the study set as a single serial list would restore the advantage of peg-based strategies as found by Roediger (1980), leading to the superiority of the Peg Serial List group over the Link Method group.

Mastery of the peg list and compliance. There were two technical problems we sought to solve before proceeding with the experiments. First, if a participant were to use the Peg-List Method, it would seem necessary to first master the peg list itself. In our reading of research on the Peg-List Method, we could not find examples of researchers implementing formal procedures to ensure mastery of the peg list. For example, Roediger (1980) asked participants to master the peg list at home, and upon returning, were asked to write out the peg list. Not all participants had fully mastered the peg list. We ensured all participants were trained to perfect memory for the peg list prior to its application, to examine that peg-list learning process itself, and to be able to test whether degree of mastery (i.e., number of trials to criterion) of the peg list plays a critical role in the effectiveness of peg-based strategies.

Second, as is typical, we planned to instruct participants to use the Peg-List Method versus other strategies. Instructing a participant to apply a strategy does not guarantee they will actually apply the strategy (Bellezza, 1981; Bugelski, 1968). We were concerned that our Peg-List group would include participants who were not using the Peg-List Method. Previous studies have either assumed that participants were compliant (e.g., Roediger, 1980) or provided peg cueing at study and/or at test (e.g., Quinn & McConnell, 1999), which we worried would alter the way participants apply the strategy. Bugelski (1968) asked participants to self-report their strategy usage at the end of the experiment (Bugelski, 1968), although this brings with it concerns about the validity of subjective report. For our goals here, we were particularly concerned with verifying compliance because we were testing a new application of the strategy. To get a concrete feel for the size of the problem compliance can produce, in a study of the Method of Loci applied to serial lists (Legge, Madan, Ng, & Caplan, 2012), which some people have suggested is an optimal strategy (Yates, 1966), *self-reported* compliance rates were not particularly high (in two strategy groups, 40% and 58% of participants, respectively, reported using the instructed strategy on more than half the lists). It was important to find out if participants were successful in applying the Peg-List Method to pairs as we had asked them to. We therefore included a "peg-recall" task: following study and cued recall of a set of pairs, we presented all the studied items, one at a time in a new random order, and asked the participant to respond with the corresponding peg they used with the item (cf. a related procedure, presenting peg numbers as cues for list items, used by Smith & Noble, 1965). We scored their responses based on what the correct peg should have been. If a participant did not use the Peg List Method, they should perform quite poorly on peg-recall, approaching the base rate of 1/5 or 1/10, depending on the peg strategy variant. High peg-recall accuracy can be taken as positive evidence of compliance with the strategy. In addition, in an attempt to ensure higher compliance rates, we included an incentive: participants were told that they would receive a monetary bonus based on their accuracy on the peg-recall task (but not for the other tasks in the experiment).

Next, we were concerned about potentially attributing inter-item association effects to the Peg-List Method. Although it seems, in principle, impossible to prevent participants from forming inter-item associations, or at least to verify that they are not doing so, we aimed to gain some insight into possible cross-contamination from inter-item associations by first, asking peg-list participants to avoid making inter-item associations, and then at the end of the testing session, asking participants to self-report the degree to which they felt they were successful. Such introspective results should always be taken with a pinch of salt, particularly because we were essentially asking participants to admit not following our instructions. Aiming to increase the level of honesty, we included a preamble, explaining that although we asked participants to avoid forming inter-item associations (if they were in the Peg List groups), as researchers, we were unsure if this was possible, and we would like to know if they felt they were successful. We then incorporated their responses into the data-analyses to test for the possible effects of inter-item imagery on Peg group behaviour.

Overview of the experimental design

The study was run as two experiments. The first experiment investigated the three Peg List groups. The second experiment included the Interactive Imagery and Link Method groups, but a second group of the Two Pegs/Pair strategy was included to check for cohort effects. As elaborated in the Results, little trace of cohort effects was found.

Our main measure of interest was cued-recall accuracy achieved by participants in each group, and patterns of errors were investigated with reference to questions derived from mathematical models of serial recall and memory for associations. For the Peg groups, we additionally interrogated the relationship between memory for the peg words and cued recall success, by probing with each list word individually and requesting the peg word it had been associated with.

Methods

Experiment 1: Three variants of the Peg List Method applied to sets of word-pairs

The first experiment included three groups, each using a variant of the peg-list method to learn sets of five pairs (a total of ten words).

Participants. A total of 327 participants (a sample of convenience, with a goal of about 100 participants per group to support individual-differences analyses) were recruited from the introductory psychology research participation pool at the University of Alberta in exchange for partial course credit plus a monetary bonus proportion to their performance on the peg-recall task (described below) at a rate of \$0.057 $\times$ number of correct peg responses summed over the 8 experimental lists plus one practice list, rounded to the nearest \$0.25, yielding a maximum of \$5.00. Participants were required to have learned English before the age of six, to have normal or corrected-to-normal vision, and provided written, informed consent. Participants were assigned, based on testing-order, to three groups, One Peg/Pair, Two Pegs/Pair and Peg-Serial List, described in detail in the Procedure. Participants were tested in groups of up to 15 participants assigned to the same condition, but each participant performed the study in an isolated testing chamber and received their own randomization of the stimuli. Nineteen particpants were excluded due to technical problems (logging errors or duplications of subject IDs), leaving 293 included participants (One Peg/Pair $N = 95$; Two Pegs/Pair: $N = 99$; Peg Serial List: $N = 99$). The procedures were approved by an ethics review board at the University of Alberta.

Materials. The peg list as cited in (Lieberman, 2011) was used, comprising ten high-imageability English nouns, each of which rhymes with its corresponding serial position numeral from 1 to 10: 1–bun, 2–shoe, 3–tree, 4–door, 5–hive, 6–sticks, 7–heaven, 8–gate, 9–vine, 10–hen.

Paired associate study sets were constructed with nouns of high imageability used in previous experiments (Caplan & Madan, 2016; Caplan, Madan, & Bedwell, 2015; Madan, Glaholt, & Caplan, 2010), where full details about word properties can be found. Briefly, the word pool contained 110 English nouns, 4–6 letters each, from the CELEX Lexical Database (Baayen, Piepenbrock, & Gulikers, 1995), with the single word "heaven" removed because it was already a peg word. Words were randomly assigned to participants, and to study sets and pairs without replacement.

Procedure. Each participant started with training task to learn the 10-word peg list, followed by the paired-associate task, where they were asked to use the pre-learned peg list to learn eight study sets comprised of five word pairs each. Each study set included a study phase followed by a mathematical distractor task, cued recall of the just-studied pairs, and a set of peg-recalls based on the just-studied pairs. Finally, participants filled out a questionnaire to obtain self-reported insight into their strategies.

Peg-list learning. Participants alternated study and test trials until they reached a criterion of three perfect consecutive recalls of the entire list (all participants reached the criterion). In a study trial, participants were instructed to study and learn the peg list, which was presented sequentially. For each peg, the corresponding peg number (numbers 1–10) was presented centrally for 2000 ms, followed by a blank inter-stimulus interval (ISI) of 150 ms, the peg word itself for 2000 ms and then a 150 ms before the next peg number. Each test trial was a sequential set of position probes; the numbers 1–10 were presented individually, in numerical order. Under each number probe was a blank underline, where participants were asked to type the corresponding peg word and submit it by pressing Enter. Backspacing was allowed while typing a given word, but once submitted, the peg-word response could no longer be edited (i.e., no backtracking). Participants were told they could type PASS if they did not know a word. Accuracy was scored automatically, online, with a strict correct-spelling criterion. To prevent participants from taking too much time during peg-list learning, the recall phase timed out and continued to the next study phase after 45 s, regardless of which positions had been probed.

Paired associate task.

Study phase. All groups were asked to learn eight study sets (plus an initial practice study set, identical to the experiment study sets, but not reported here), using the peg list they just learned, by forming images combining new words with the peg words, and to use the peg words in order, starting with peg 1 (bun). Study sets were presented one item at a time, centrally on the screen (Figure 1). Sequential presentation was chosen to better equate the association and serial-list paradigms. If we presented each pair simultaneously, we worried this would lead the serial-list groups to study the list as a set of pairs rather than a serial list. On the other hand, sequential presentation of pairs has produced very similar cued-recall data as simultaneous (e.g., Caplan, 2005; Caplan et al., 2006; Madan et al., 2010), so sequential presentation should not reduce the generalizability of the results. After a fixation cross (a plus sign) was presented for 500 ms, each item was presented for 4000 ms, with a 1000-ms ISI within pairs and a 3500-ms ISI following each pair, including the last. Following the last pair (and the 3500-ms post-pair blank interval) came a distractor task. The distractor was included as is typically done to attenuate possible sources of variance of non-interest due to study–test lag. Note that the timing during the study phase was identical for all groups. All participants were asked to avoid forming images directly between the target list items; we anticipated that this might be challenging, and included a self-report assessment of their success, which we describe below.

Strategy groups. The three strategy groups differed only in a few words in the instructions, as illustrated in Figure 2. *1) One Peg/Pair* participants were asked to link both items within a pair to the same peg (Figure 2b). Because there were five pairs in each study set, One Peg/Pair participants were only supposed to use pegs 1–5. *2) Two Pegs/Pair* participants were asked to link each item to a different (i.e., the next) peg word, regardless of whether the items were in the same pair or not. Thus, the first word should be linked with bun, then second word with shoe, etc. Two Pegs/Pair participants thus used all 10 pegs (Figure 2c). *3) Peg Serial List* participants were also instructed to link each item to a different (the next) peg word, but differed from the Two Pegs/Pair group in that the study sets were described to participants in this group as comprising a single, ordered list (Figure 2d). All participants were asked to avoid making images that combined two list words directly.

Distractor task. The distractor consisted of two sequentially presented trials of addition questions of the form "A + B + C = ", where A, B and C were numbers randomly selected from the range 2–8, inclusive. Participants were given 5000 ms to type each answer, followed by a 200-ms blank screen.

Cued recall. Following the two trials of the distractor task, participants were given cued recall tests of each of the studied pairs. In each cued-recall trial, the cue word (one of the two items of a pair) was displayed on the screen, with a response line either to the left or the right of the cue word. Participants had to respond by typing the corresponding target item. Note that these cued-recall probes contained directional information (as in Caplan et al., 2006; Kahana & Caplan, 2002; Madan et al., 2010; Rehani & Caplan, 2011). For a set of pairs, this directional cueing is not necessary; given B as a cue, the participant need not know whether the pair was AB or BA, only that A was paired with B. Probes were directional because we included the Peg Serial List group, and for those participants, a single-item probe is ambiguous; the participant needs to be told which associate (the subsequent or prior item) is desired. The cue remained on the screen until the participant submitted their response by pressing the Enter key, which was followed by a 250-ms blank screen before the next cue was displayed. Participants were instructed to type PASS if they could not think of the target item. Accuracy was scored according to a strict correct-spelling criterion.

All pairs were tested once, in random order, followed by a second complete set of randomly ordered cues. This successive testing procedure was included to examine associative symmetry (Kahana, 2002). The direction of test (whether the earlier— forward probes— or later— backward probes— item of a pair was the probe) was counterbalanced so that, over the whole eight sets (excluding the practice study set), there were an equal number of pairs tested Forward/Forward, Forward/Backward, Backward/Forward and Backward/Backward in test 1/test 2, respectively.

Peg recall. Following both sets of cued recall tests, participants were presented each studied word individually, in a new random order— disregarding pairings and cued-recall order— and asked to respond with the peg they had linked to each word. Prior to the set of peg probes, an instruction screen included a reminder that the participant would be rewarded with a bonus based on their peg-recall performance. Each cue word had a response line displayed underneath it; both were centered horizontally. The list of pegs words were displayed at the bottom of the screen, in a single, horizontal row, in peg-number order. The peg list was displayed to avoid participants simply failing to recall the peg words, and to reduce misspellings. The peg list was thus displayed in order, but without numerals. The word-probe remained on the screen until the participant submitted their typed response with the Enter key, after which a new screen displayed, for 500 ms, 'Correct: 10 points' or 'Incorrect: 0 points', based on a strict correct-spelling criterion. At the end of each set of peg probes, a screen displayed the running total of the participant's point-count, cumulative across all sets of peg-recall trials. Participants could type PASS if they did not know the peg word.

End-of-session questionnaire. At the end of the session, all participants were asked to complete a questionnaire, assessing strategy familiarity and difficulty as well as participant adherence to the instructions to only use the assigned strategy. Participants were asked whether they had prior knowledge of the Peg List Method technique, whether they found it easy to apply the strategy to a list of pairs or words. They were also asked if and how they used the peg list method. They were asked if they created an image between the items of one pair (excluding the pegs), as opposed to what they were instructed to do: create images between the item of each pair presented and the corresponding peg. This question was coded as Yes/No/Sometimes, and these responses are included in analyses reported below. Finally, they were asked, in free-form, to describe any other strategies they used to learn the items.

Experiment 2: Comparison of the Peg-List Method to Interactive Imagery strategies

In Experiment 2, we investigated two strategies that rely on images directly integrating pairs of list-items. To be able to contrast behaviour with these strategies to the data from Experiment 1, we included a single peg-group, Two Pegs/Pair (Figure 2c). This group enabled us to check for sampling biases between experimental cohorts. Because none were found (see Results), comparisons between experiments are better justified, although they should be interpreted with caution. The Two Pegs/Pair strategy was selected for this role because early inspections of self-reported inclusion of inter-item imagery were lower than for the One Peg/Pair participants in Experiment 1; thus, the Two Pegs/Pair strategy seemed to have less potential overlap with the two strategies investigated in Experiment 2, which explicitly rely on inter-item images.

The two new groups were an Interactive Imagery group, asked to form images in which the two paired items interact with one another (Figure 2e), and a Link Method group, who were asked to use interactive imagery, but were instructed that they were studying serial lists (Figure 2f), similar to the Peg Serial List group in Experiment 1.

The methods for the Two Pegs/Pair group were identical to those used for the Two Pegs/Pair group in Experiment 1, including the peg-list pre-training task. For the Interactive Imagery and Link Method groups, all methods were identical to those used in Experiment 1, except: the instructions (detailed next), the peg-list learning and peg-recall tasks were omitted, and the extra time freed up was used to include 10, rather than 8, study/test cycles of cued recall.

Participants. A total of 250 participants (a convenience sample, aiming to match the sensitivity of Experiment 1, taking into account the greater data-yield per participant in Experiment 2) participated for course credit. After excluding two participants due to technical problems, the final sample size was 248 (Two Pegs/Pair: $N = 75$; Interactive Imagery: $N = 75$; Link Method:[2] $N = 94$).

Strategy groups. The instructions differed across the three groups as follows. 1) *Two Pegs/Pair* participants were instructed identically to the corresponding group in Experiment 1. 2) *Interactive Imagery* participants were asked to form an image combining the two items of each pair. 3) *Link Method* participants were told they were studying single, 10-item lists, and asked to form an image between each item and the subsequent

[2]The larger sample size in this group is because we ran a separate cohort of Link Method participants intended for a different study (unpublished).

item. The first image is formed between the first two items of the serial list, the second image is formed between the second and third item in the list and so on.

At the end of the session, the Two Pegs/Pair participants were given the same questions as in Experiment 1, and participants who were instructed to use either the interactive imagery or link method were asked on a scale from 1 to 7 how easy it was to create an image of the pair or item, justification of their rating and what other strategies they used to learn the items.

Data analyses

Analyses were conducted in MATLAB (The Mathworks, Inc.) and JASP (JASP Team, 2016). Violations of non-sphericity are addressed using the Greenhouse-Geisser correction, and post-hoc multiple comparisons are Bonferroni-corrected to control Type *I* error rates. Bayesian versions of analyses, run in JASP (JASP Team, 2016), assuming uniform priors, are used both to check whether null findings are under-powered or more consistent with a null effect, and to check whether significant findings might be too small to be considered meaningful (but magnitudes of effects are also considered in these cases). The Bayes Factor quantifies the ratio of evidence for the effect versus the evidence for the null, given the data. By convention, a Bayes factor above 3 (3:1 ratio) will be viewed as "some" support for the effect, and below 1/3 (1:3 ratio) as "some" support for the null; "strong" evidence is inferred when BF > 10 or < 0.1, respectively (Kass & Raftery, 1995). For t tests and correlations, this Bayes Factor is denoted BF_{10}, and for effects reported from ANOVAs, we report the Bayes Factor for inclusion of the effect (which is derived from Bayes Factors for all possible models), denoted $\mathrm{BF}_{\mathrm{inclusion}}$.

To test for contiguity effects in error responses, we analyzed distance functions both in cued-recall and peg-recall data. A distance function measures the proportion of error responses as a function of lag, where lag is defined as the output position minus the input position, illustrated in Figure 5c. After excluding omissions and extra-list intrusions (responses other than peg words, for peg recall), the value at each lag was the ratio of the number of errors made at the lag to the number of trials on which that lag was available as a possible response. The distance function is thus the proportion of error responses to a lag, given that that lag was available. Incorporating availability is essential for lag functions like this, because edge effects would otherwise produce an artifactual contiguity effect (Howard & Kahana, 1999).

To quantify contiguity effects, we conducted a linear regression[3] to the positive lags of a given distance function. We did the same for the negative lags. Any lags with denominators of zero (i.e., no opportunity for the participant to produce an error at that lag) were considered missing data and left out of the regression. Because a regression is undefined with fewer than two points, participants with fewer than two lags available for any of these analyses were excluded from the respective analysis. The result was a positive slope and a negative slope for each included participant for each distance function.

Results

First we report data on peg-list learning, collapsing across both experiments, because peg-list learning was the first activity in all groups. We then report analyses establishing that List Number can be safely collapsed across in analyses of the cued-recall data, again for both experiment. Then we report the findings of the remaining tasks, first for Experiment 1 and then Experiment 2.

Peg-list learning (both experiments)

First we examine learning of the peg list itself. The numerical rhyming peg list was expected to be relatively easy to learn, due to the rhyming scheme, and because many of the number–peg pairings are part of popular nursery rhymes, but we could find no

[3] Although distance functions tend to be approximately exponential, we sought a very simple way of quantifying the size of the contiguity effect, and with few degrees of freedom and requiring fewer data points given that some participants had missing data at some lags.

published reports of acquisition of the peg list itself. A histogram of trials to criterion (TTC), requiring three successive perfect serial recalls of the peg list (Figure 3a), shows that nearly one third of participants either could recall the peg list perfectly after a single study trial. Most of the remaining required only a second study trial to achieve perfect recall, and a smaller number required more trials or never mastered the peg list perfectly. Collapsing together cells with low counts (6–10 trials to criterion), the three peg groups in Experiment 1 did not differ significantly in TTC. TTC also did not differ significantly across experiments, between the two groups of Two Pegs/Pair participants, $\chi^2(3) = 1.00,\ p = 0.80$, suggesting no cohort effects between experiments. Due to the range of trials to criterion within groups, we later break down analyses of interest by trials to criterion to determine if level of mastery (i.e., overlearning) of the peg list influences its effectiveness.

For participants who mastered the peg list over several trials, Figure 3b–d plots serial position curves. As is typical of serial recall (Murdock, 1974), one can see a dominant primacy effect (although the last few positions are confounded because the entire recall phase timed out after 45 seconds). Thus, despite the peg list being special, in that each peg word is a high-frequency, imageable word that rhymes with the corresponding peg number, when participants do need to learn it, they apparently learn the peg list very much like any episodic serial list of words. Two exceptions to this are that there seems to be a slight advantage for the odd-numbered pegs, and a relative disadvantage for the sixth peg word, STICKS.

Effects of List Number on cued recall

List Number and Pair Position were intended to be factors of no interest. In Experiment 2, participants studied ten, compared to eight lists in Experiment 1. Before proceeding with analyses of cued-recall data (Figure 4), we first checked whether or not List Number needed to be carried throughout the data analyses, or could be safely collapsed across. Therefore, in this section, we report analyses of cued-recall accuracy incorporating List Number and both experiments. First, because the Interactive Imagery and Link Method groups had two more lists than the other groups, we started by analyzing the effect of List Number of cued recall accuracy for just those two groups, using all the data. We conducted a mixed, repeated-measures ANOVA with design Group[2]×List Number[10]. The main effect of List, $F(8.0, 1343) = 0.61$, $MSE = 0.036$, $p = 0.77$, $\eta_p^2 = 0.004$, $\mathrm{BF}_{\mathrm{inclusion}} < 0.0001$ and interaction, Group×List Number, $F(8.0, 1343) = 1.20$, $MSE = 0.036$, $p = 0.29$, $\eta_p^2 = 0.007$, $\mathrm{BF}_{\mathrm{inclusion}} < 0.0001$, were both non-significant. Because all groups had at least 8 experimental lists, we next conducted a mixed, repeated-measures ANOVA on Group[6]×List Number[8]. The interaction, Group×List Number, was again non-significant, $F(29.5, 3137) = 1.05$, $MSE = 0.041$, $p = 0.39$ $\eta_p^2 = 0.010$, $\mathrm{BF}_{\mathrm{inclusion}} < 0.0001$. In light of these findings, we collapse across List Number for both experiments, and use all all available data, including from all ten lists of Experiment 2.

Results: Experiment 1

Turning to the main task of interest, we report cued recall data, distance functions, tests of associative symmetry and response times. We then turn to the responses to the post-session questionnaire about usage of inter-item imagery, and break down subsequent analyses based on this self-reported characteristic, to test if inter-item associations might explain any of the effects we might like to attribute to the peg-list method. We then report the peg-recall data, including a distance-function analysis of those responses. Finally, we test whether the speed of mastery of the peg list predicts accuracy in both cued recall and peg recall.

Cued recall accuracy. To check for effects of Pair Position, expected to be a factor of no interest, a mixed, repeated-measures ANOVA with design Strategy[3] × Pair Position[5] produced a significant main effect of Strategy, $F(2, 290) = 18.13$,

$MSE = 0.46$, $p < 0.001$, $\eta_p^2 = 0.11$, $\text{BF}_{\text{inclusion}} > 1000$. Bonferroni-corrected post-hoc t tests were all significant, with One Peg/Pair > Two Pegs/Pair > Peg Serial List. The main effect of Pair Position was also significant, $F(3.6, 1039) = 20.15$, $MSE = 0.018$, $p < 0.001$, $\eta_p^2 = 0.065$, $\text{BF}_{\text{inclusion}} > 1000$. Bonferroni-corrected post-hoc t-tests found position 1 significantly greater than all subsequent positions but no other significant differences. This confirms an overall decrease in accuracy from the first two pairs to the rest of the list. The interaction was not significant ($p = 0.21$, $\text{BF}_{\text{inclusion}} = 0.030$).

Next, we collapsed across List Number and Pair Position. A mixed, repeated-measures ANOVA with design Strategy[3] × Test[1/2] × Direction[Forward/Backward] produced a significant main effect of Strategy, $F(2, 365) = 17.36$, $MSE = 0.38$, $p < 0.001$, $\eta_p^2 = 0.087$, $\text{BF}_{\text{inclusion}} > 1000$, reiterating the effect in the previous ANOVAs. Post-hoc pairwise t tests revealed Interactive Imagery > all but Link Method and One Peg/Pair (trend: $p = 0.061$); Link Method > all peg groups except One Peg/Pair; and One Peg/Pair > Two Pegs/Pair > Peg Serial List. The main effect of Test was also significant, $F(1, 365) = 97.50$, $MSE = 0.004$, $p < 0.001$, $\eta_p^2 = 0.21$, $\text{BF}_{\text{inclusion}} > 1000$. The main effect of Direction was significant, $F(1, 365) = 11.76$, $MSE = 0.010$, $p < 0.001$, $\eta_p^2 = 0.031$, $\text{BF}_{\text{inclusion}} = 77.40$, with greater accuracy on backward than forward probes. The interaction, Strategy × Direction was also significant, $F(2, 365) = 3.31$, $MSE = 0.010$, $p = 0.038$, $\eta_p^2 = 0.018$, $\text{BF}_{\text{inclusion}} = 4.31$. All other effects were clearly non-significant ($p > 0.1$).

Cued recall response times. Response times during correct cued-recall trials (quantified as the onset of typing following display of the probe word) could either show congruent effects to accuracy (higher accuracy associated with faster = lower response times) or could reveal speed–accuracy tradeoffs that might indicate ways in which participants could achieve greater memory success by working longer to retrieve associates. An ANOVA on correct-response time with design Strategy[3] × Test[2] × Direction[2] revealed significant main effects of Strategy, $F(2, 245) = 19.08$, $MSE = 12.2 \times 10^6$ ms^2,

$p < 0.001$, $\eta_p^2 = 0.14$, $\mathrm{BF_{inclusion}} > 1000$; and Test, $F(1, 245) = 118.09$, $MSE = 8.7 \times 10^5$ ms^2, $p < 0.001$, $\eta_p^2 = 0.33$, $\mathrm{BF_{inclusion}} > 1000$; Bonferroni-corrected post-hoc t tests on Strategy revealed One Peg/Pair < Two Pegs/Pair = Peg Serial List, where < denotes faster response times (lower value) and = denotes non-significant difference. Test 2 responses were faster than Test 1.

Cued recall distance functions. When participants made intrusions of other items within the same list, distance functions (Figure 5) show that these errors were primarily from neighbouring serial positions (contiguity effect). The zig-zag pattern indicates that errors were more likely to the made to items from the target-item's position within the pair; thus, even lags (relative to the target's serial position) were far more frequent than odd lags. Summing over even and odd lags separately, this was confirmed in a mixed, repeated-measures ANOVA with design Strategy[3] × Lag [Even/Odd] × Direction[Forward/Backward]. The main effect of Lag was significant, $F(1, 144) = 66.38$, $MSE = 0.43$, $p < 0.001$, $\eta_p^2 = 0.32$, $\mathrm{BF_{inclusion}} > 1000$, with greater proportions of errors to even than to odd lags. All other effects were non-significant ($p > 0.4$, $\mathrm{BF_{inclusion}} < 0.2$). These findings show that all groups used knowledge of within-pair position during cued recall.

Associative Symmetry. We use Yule's Q (Figure 6) to compute the correlation between successive cued-recall tests (Kahana, 2002). Yule's Q functions quite like Pearson correlation, but is appropriate for dichotomous data (i.e., 2×2 contingency tables). For statistical tests, Q values are log-odds transformed, since these values are theoretically Gaussian-distributed, and the p value for a Q is equivalent to the significance of the χ^2 test of the same 2×2 contingency table. The correlation of interest, $Q_{\mathrm{Different}}$, computes the correlation between accuracy in test 1 and test 2 when the direction changes between tests (forward/backward and backward/forward). This is typically nearly as high as the correlation for Q_{Same}, test/re-test in the same direction (forward/forward and backward/backward). This provides an estimate of the realistic upper bound one could obtain for $Q_{\text{Different}}$. Finally, due to Simpson's Paradox (Hintzman, 1980), the correlation between successive tests, even for different pairs within the same list, is expected to be somewhat positive, due to common across-list variability. We therefore also compute, for comparison, Q_{Control}, which is a bootstrap that measures the correlation between successive tests of different pairs within a list (Caplan, 2005). This provides an estimate of the realistic lower bound one could obtain for $Q_{\text{Different}}$. Participants with zero or perfect recall on either test 1 or test 2 were excluded from these analyses, because Q is undefined in such cases.

A repeated-measures ANOVA on log-odds-transformed Q values (to satisfy the assumption of normality), with design Strategy[5] $\times$ Correlation Type[Q_{Same}, $Q_{\text{Different}}$, Q_{Control}] produced a significant main effect of Correlation Type, $F(1.9, 456) = 643.11$, $MSE = 1.34$, $p < 0.001$, $\eta_p^2 = 0.73$, $\text{BF}_{\text{inclusion}} > 1000$. Bonferroni-corrected post-hoc pairwise t tests found all comparisons significant, with $Q_{\text{Same}} > Q_{\text{Different}} > Q_{\text{Control}}$. The main effect of Strategy was not significant, $F(2, 239) = 1.13$, $MSE = 1.34$, $p = 0.32$, $\eta_p^2 = 0.009$, $\text{BF}_{\text{inclusion}} = 0.052$. The interaction, Strategy$\times$Correlation Type was also non-significant, $F(2.5, 456) = 1.89$, $MSE = 1.34$, $p = 0.11$, $\eta_p^2 = 0.016$, with $\text{BF}_{\text{inclusion}} = 0.034$, strongly favouring the null and suggesting that the interaction was negligible in magnitude. Thus, associative symmetry holds, and is not significantly modulated by strategy instruction.

Effects of self-reported inter-item imagery. In an attempt to keep the peg list method similar to positional-coding models and distinct from associative chaining or inter-item association models, we instructed participants in all groups to avoid forming images directly between list items. That said, we were unsure whether or not this would be possible, and if possible, whether participants could do this successfully. To get a handle on how much "contamination" by inter-item images there might be, at the end of the testing session, we asked participants to confess how successful they felt they were in avoiding inter-item images (see Methods). Note that four participants did not respond to the inter-item image question, and are excluded from any analyses involving inter-item images. Responses differed across the group (Table 1), $\chi^2(4) = 16.47$, $p = 0.0024$. This was due to the One Peg/Pair group reporting more inter-item images than the other groups; when this group was excluded, the χ^2 test was non-significant (but both pairwise comparisons with One Peg/Pair were significant, $p < 0.005$).

We wondered if peg participants who confessed also to forming inter-item images performed better than those who reported sticking strictly to item–peg images. A mixed, repeated-measures ANOVA on cued recall accuracy, with design Inter-Item Images[No, Sometimes, Yes]×Peg Strategy[3]×Test[2]×Direction[2] produced no support for this hypothesis; the main effect and all interactions involving the Inter-Item Images factor were non-significant ($F < 1.5$, $p > 0.2$, $\eta_p^2 < 0.02$, $\mathrm{BF}_{\mathrm{inclusion}} < 0.05$).

Next, we wondered if the inclusion of inter-item images affected the distance functions. We conducted four ANOVAs with design Strategy[3]×Inter-Item Images[No, Sometimes, Yes], on the positive and negative slope of the linear regressions conducted on positive and negative lags, respectively. All effects involving the factor Inter-Item Images were non-significant ($p > 0.05$, $\mathrm{BF}_{\mathrm{inclusion}} < 0.3$) except as follows. For Backward cued recall, negative lags: the main effect of Inter-Item Imagery was significant, $F(2, 151) = 3.86$, $MSE = 0.085$, $p = 0.023$, $\eta_p^2 = 0.049$, $\mathrm{BF}_{\mathrm{inclusion}} = 0.45$, where no Bonferroni-corrected post-hoc tests were significant, but marginal effects suggested "No" participants having a smaller positive (or even negative) slope than both "Sometimes" and "Yes" participants ($0.05 < p < 0.1$). For Backward cued recall, positive lags, the interaction Strategy × Inter-Item Imagery was also significant, $F(2, 147) = 4.56$, $MSE = 0.028$, $p = 0.002$, $\eta_p^2 = 0.11$, $\mathrm{BF}_{\mathrm{inclusion}} = 11.71$. Follow-up tests suggested non-monotonic patterns, with the One Peg/Pair group, "Sometimes" participants exhibiting the most steeply negative slope. Thus, a clear pattern does not emerge from this set of analyses.

Peg Recall. After cued recall of each list, participants in the Peg groups were asked to produce the peg attached to each list-word, tested in a new random order. This enables us to independently assess the success with which a participant actually formed word–peg associations (as requested). We first report the peg recall data on their own, and then check whether peg recall directly determines cued recall accuracy. Peg-accuracy was very high, with little trace of any effect of strategy (Figure 7a). Serial-position curves for peg recall show the familiar bow-shape, with a dominant primacy effect and a gentle recency effect.

First, focusing on Two Pegs/Pair and Peg Serial List, including all 10 peg positions, a mixed, repeated-measures ANOVA on Peg Strategy[2]×Serial Position[10] revealed only a significant main effect of Serial Position, $F(7.6, 1495) = 22.68$, $MSE = 0.017$, $p < 0.001$, $\eta_p^2 = 0.097$, $\mathrm{BF}_{\mathrm{inclusion}} > 1000$. Bonferroni-corrected post-hoc t-tests revealed serial-position 1 was superior to all other positions except 2; 2 superior to all subsequent positions; and 3 and 4 both superior to positions 6–9; ($p < 0.05$; and all other comparisons non-significant. This broadly confirms the U-shaped function.

Adding the Self-confessed Inter-Item Images[3] as a between-subjects factor, Inter-Item Images was a significant main effect, $F(1, 189) = 8.18$, $MSE = 0.33$, $p < 0.001$, $\eta_p^2 = 0.080$ $\mathrm{BF}_{\mathrm{inclusion}} > 1000$, as was the interaction, Inter-Item Images × Serial Position, $F(15.5, 2002) = 3.55$, $MSE = 0.016$, $p < 0.001$, $\eta_p^2 = 0.027$, $\mathrm{BF}_{\mathrm{inclusion}} > 1000$. Bonferroni-corrected post-hoc t tests revealed only significantly greater peg-recall accuracy for the “Yes” participants than for the “No” participants, suggesting that the presence of inter-item associations does not compromise item–peg learning.

Next incorporating the One Peg/Pair group, analyzing only peg positions 1–5, a mixed, repeated-measures ANOVA on Peg Strategy[3]×Serial Position[5] revealed only a significant main effect of Serial Position, $F(3.7, 1076) = 33.52$, $MSE = 0.012$, $p < 0.001$, $\eta_p^2 = 0.103$, $\mathrm{BF}_{\mathrm{inclusion}} > 1000$. Bonferroni-corrected post-hoc pair-wise t tests confirmed a decreasing trend from serial position 1 through 5, all significant ($p < 0.05$) except positions 3 and 4 were not significantly different. Adding the self-confessed Inter-Item Images factor, the main effect of this factor was not significant, nor was its interaction with Serial Position.

Distance functions were similar for all strategies (Figure 7b), with all strategies showing large contiguity effects. A one-way ANOVA on the slope of the negative distances, with Strategy[3] as the factor, revealed a significant main effect, $F(2, 252) = 11.40$, $MSE = 0.010$, $p < 0.001$, $\eta_p^2 = 0.083$, $\mathrm{BF}_{\mathrm{inclusion}} = 911$, with Bonferroni-corrected t test finding the One Peg/Pair strategy produced the steepest (most positive) slope, but no difference between the Two Pegs/Pair and Peg Serial List groups. The same analysis for the slope of the positive distances was not significant ($p = 0.73$, $\mathrm{BF}_{\mathrm{inclusion}} = 0.056$).

Prior mastery of the peg list. As we suggested in the Introduction, the peg list method might be effective because the peg list is typically pre-memorized, even days or weeks prior to application to learning new lists. We wondered if prior knowledge of the peg list played a role in how well participants could apply the peg list strategies, such as the hypothesized anchoring advantage. We compared participants who required only three trials to criterion to participants who required more trials to master the peg list. Although we did not test knowledge of the peg list before presenting it for study, participants who needed only three trials to reach our criterion would have made zero errors, even after only one exposure to the list. This subset, then, comprises the subset of participants who had the most foreknowledge of the peg list. Trials to criterion is also influenced by speed of acquisition of memory of the peg list. Thus, any effects of trials to criterion could be due to effects of prior knowledge or effects of speed of learning or some combination; this ambiguity cannot be resolved with the present data. Each participant was assigned a category of 3, 4 or > 4 trials to criterion. A two-way ANOVA on peg recall accuracy, collapsed across peg positions, with design Peg Strategy[3] $\times$ TTC[3] produced a significant main effect of TTC, $F(2, 284) = 9.34$, $MSE = 0.038$, $p < 0.001$, $\eta_p^2 = 0.0622$, $\mathrm{BF}_{\mathrm{inclusion}} = 216$. Fewer trials to criterion were associated with higher peg-recall accuracy (TTC $3 > 4 >$ more than 4), but Bonferroni-corrected post-hoc t tests confirmed the significance only of $3 >$ more than 4 and $4 >$ more than 4 TTC (Figure 8a). No other effects were significant ($p > 0.5$, $\mathrm{BF}_{\mathrm{inclusion}} < 0.03$). Thus (presumed) prior knowledge of the peg list did have a positive influence on participants' success learning and remembering item–peg associations.

An ANOVA on cued recall accuracy with design Peg Strategy[3] × TTC[3] also produced a significant main effect of TTC, $F(2, 284) = 5.61$, $MSE = 0.088$, $p = 0.004$, $\eta_p^2 = 0.038$, $\mathrm{BF}_{\mathrm{inclusion}} = 5.41$ (Figure 8b) but the interaction with Peg Strategy was not significant($p = 0.65$, $\mathrm{BF}_{\mathrm{inclusion}} = 0.18$). All differences were in the expected order, but Bonferroni-corrected post-hoc t-tests confirmed only a significant difference between 3 and more than 4 trials to criterion ($p = 0.003$). Thus, prior knowledge of the peg list did positively influence cued recall for the participants asked to use peg-based strategies. Surprisingly, even the lowest-TTC participants, using peg strategies, still do not quite reach the accuracy level of the entire, unselected pool of participants instructed to use Interactive Imagery in Experiment 2 (compare Figure 8b with Figure 4).

Effect of item–peg memory on cued recall. If participants are applying the peg strategy as instructed and can accurately recall which peg was associated with both items in the pair (in the peg-recall phase), they should have performed accurately on cued recall of the pair. We broke down cued-recall accuracy based on the number of pegs correctly recalled for the pair (0, 1 or 2). A mixed, repeated-measures ANOVA on cued-recall accuracy (Figure 9), with design Peg Strategy[3]×Peg Accuracy[0,1,2] produced significant main effects of Strategy, $F(2, 159) = 7.72$, $MSE = 0.19$, $p < 0.001$, $\eta_p^2 = 0.089$, $\mathrm{BF}_{\mathrm{inclusion}} = 30.16$, and Peg Accuracy, $F(1.73, 274) = 121.64$, $MSE = 0.042$, $p < 0.001$, $\eta_p^2 = 0.43$, $\mathrm{BF}_{\mathrm{inclusion}} > 1000$. Bonferroni-corrected post-hoc t tests found all pairwise comparisons significant, with $2 > 1 > 0$ correct, as expected. The interaction was not significant ($p = 0.65$, $\mathrm{BF}_{\mathrm{inclusion}} = 0.10$). A remarkable feature of the relationship is that even when both pegs were known, accuracy was not at ceiling; accuracy for these highly selected trials also did not quite rise to the accuracy of the (nonselected) Interactive Imagery group of Experiment 2 (cf. Figure 4). Thus, even when participants had complete knowledge of the pegs associated with both items, they were not consistent in using that information to retrieve the correct response word.

Interim Summary of Experiment 1. Participants were able to apply peg-based strategies to memorize a list of pairs and perform well on cued-recall tests. Greater mastery of the peg list did result in greater success in cued recall. Self-reported application of peg strategies was not process-pure, but included direct inter-item associations, particularly for the One Peg/Pair strategy group, but we found no evidence that such spontaneous incorporation of inter-item images facilitated or impaired cued-recall. Correct memory for item–peg associations was also associated with better cued-recall, providing some confirmation that participants were succeeding with the peg strategy as instructed. The peg list exhibited a key characteristic of positional-coding models, contiguity, and participants appeared to use knowledge of within-pair position as part of their retrieval cue.

Results of Experiment 2

Having established that peg-based strategies are possible for participants to apply with success to lists of pairs, we next asked how these strategies compare with inter-item imagery and its serial-recall analogue, the link method. We included the Two Pegs/Pair strategy because it was explicitly described as a strategy for pairs, not serial lists, and it was more process-pure than the One Peg/Pair strategy (less self-report of inter-item imagery).

Cued recall accuracy. To check for effects of Pair Position, expected to be a factor of no interest, a mixed, repeated-measures ANOVA on cued recall accuracy (Figure 4) with design Strategy[3] × Pair Position[5] produced a significant main effect of Strategy, $F(2, 241) = 19.80$, $MSE = 0.30$, $p < 0.001$, $\eta_p^2 = 0.14$, $\mathrm{BF}_{\mathrm{inclusion}} > 1000$. Bonferroni-corrected post-hoc t tests revealed Two Pegs/Pair produced lower accuracy than both Interactive Imagery and Link Method ($p < 0.001$), but the latter two did not differ significantly ($t = 2.02$, $p = 0.14$) The main effect of Pair Position was also significant, $F(3.5, 847) = 30.34$, $MSE = 0.015$, $p < 0.001$, $\eta_p^2 = 0.11$, $\mathrm{BF}_{\mathrm{inclusion}} > 1000$. Bonferroni-corrected post-hoc t-tests found position 1 significantly greater than all subsequent positions, position 2 significantly greater than positions 4 and 5 but no other significant differences. This confirms an overall decrease in accuracy from the first two pairs to the rest of the list. The interaction was also significant, $F(7, 847) = 2.32$, $MSE = 0.015$, $p = 0.024$, but the Bayesian ANOVA resulted in an inconclusive $\text{BF}_{\text{inclusion}} = 1.02$ for the interaction. Also considering the small $\eta_p^2 = 0.024$, we did not pursue it further.

Next, we collapsed across List Number and Pair Position. A mixed, repeated-measures ANOVA with design Strategy[3] × Test[1/2] × Direction[Forward/Backward] produced a significant main effect of Strategy, $F(2, 241) = 19.83$, $MSE = 0.24$, $p < 0.001$, $\eta_p^2 = 0.14$, $\text{BF}_{\text{inclusion}} > 1000$, reiterating the effect in the previous ANOVAs. The main effect of Test was also significant, $F(1, 241) = 43.77$, $MSE = 0.003$, $p < 0.001$, $\eta_p^2 = 0.15$, $\text{BF}_{\text{inclusion}} > 1000$, with greater accuracy on Test 2 than Test 1. The three-way interaction, Strategy × Test × Direction, $F(2, 241) = 3.26$, $MSE = 0.004$, $p = 0.04$, but with small effect size, $\eta_p^2 = 0.026$ The Bayesian ANOVA, in fact, provided strong support for the a null three-way interaction, $\text{BF}_{\text{inclusion}} = 0.0001$, so we consider this effect negligible. All other effects were clearly non-significant ($p > 0.1$).

Cued recall response times. An ANOVA on correct-response time with design Strategy[3] × Test[2] × Direction[2] revealed significant main effects of Strategy, $F(2, 245) = 19.08$, $MSE = 1.23 \times 10^7$ ms^2, $p < 0.001$, $\eta_p^2 = 0.14$, $\text{BF}_{\text{inclusion}} > 1000$; and Test, $F(1, 245) = 118.09$, $MSE = 8.7 \times 10^5$ ms^2, $p < 0.001$, $\eta_p^2 = 0.33$, $\text{BF}_{\text{inclusion}} > 1000$; Bonferroni-corrected post-hoc t tests on Strategy revealed Two Pegs/Pair was significantly slower than both Interactive Imagery and Peg Serial List ($p < 0.001$), but the latter were not significantly different. Test 2 responses were faster than Test 1.

Cued recall distance functions. As in Experiment 1, the zig-zag pattern in (Figure 5) indicates that errors were more likely to the made to items from the target-item's position within the pair. Summing over even and odd lags separately, this was confirmed in a mixed, repeated-measures ANOVA with design Strategy[3] × Lag

[Even/Odd] × Direction[Forward/Backward]. The main effect of Lag was significant, $F(1, 124) = 54.79$, $MSE = 0.44$, $p < 0.001$, $\eta_p^2 = 0.31$, $\text{BF}_{\text{inclusion}} > 1000$, with greater proportions of errors to even than to odd lags. The interaction, Strategy×Lag, was significant, $F(2, 124) = 3.38$, $MSE = 0.44$, $p = 0.037$, $\eta_p^2 = 0.052$ but $\text{BF}_{\text{inclusion}} = 1.83$, in the inconclusive zone, suggesting the magnitude of the interaction is quite small, as also indicated by the effect size. All other effects were non-significant ($p > 0.4$, $\text{BF}_{\text{inclusion}} < 0.1$; $\text{BF}_{\text{inclusion}} = 0.35$ for the main effect of Strategy). Intriguingly, these findings show that all groups used knowledge of within-pair position during cued recall, including both inter-item imagery groups.

Associative Symmetry. A repeated-measures ANOVA on log-odds-transformed Q values (Figure 6) with design Strategy[3] × Correlation Type[Q_{Same}, $Q_{\text{Different}}$, Q_{Control}] produced a significant main effect of Correlation Type, $F(1.9, 391) = 714.64$, $MSE = 1.22$, $p < 0.001$, $\eta_p^2 = 0.77$, $\text{BF}_{\text{inclusion}} > 1000$. Bonferroni-corrected post-hoc pairwise t tests found all comparisons significant, with $Q_{\text{Same}} > Q_{\text{Different}} > Q_{\text{Control}}$. The main effect of Strategy and interaction, Strategy×Correlation Type were also non-significant ($p > 0.1$, $\text{BF}_{\text{inclusion}} < 0.1$). Thus, associative symmetry holds, and is not significantly modulated by strategy instruction.

Results: Comparisons spanning both experiments

We finish with some follow-up analyses that consider the data from both experiments together. Because the experiments were run sequentially, these findings should be interpreted with caution, as sampling differences could produce spurious differences or render what otherwise would have been significant differences non-significant. The null differences between the Two Pegs/Pair group between the two experiments argues against such a sampling bias, but this should not be considered definitive.

Interactive Imagery versus One Peg/Pair. We wondered if the Interactive Imagery strategy was not only better than the Two Pegs/Pair instruction (both collected in Experiment 2), but might be better than even the highest-accuracy form of the Peg strategy, the One Peg/Pair group of Experiment 1. We thus compared these groups cautiously, acknowledging that inter-experiment comparisons can be subject to sampling bias. The Interactive Imagery group outperformed the One Peg/Pair group, according to independent-samples t tests on cued-recall accuracy on Test 1 in both the Forward (M (SE) $= 0.80$ (0.022) and 0.65 (0.027), respectively; $t(168) = 4.14$, $p < 0.001$, $BF_{10} = 359$) and Backward (M (SE) $= 0.80$ (0.022) and 0.69 (0.26), respectively; $t(168) = 4.14$, $p < 0.001$, $BF_{10} = 359$) direction.

Speed–accuracy tradeoffs. Because the Peg strategies produced lower accuracy but also slower response times than the inter-item imagery strategies, we wondered if the peg strategies simply require more time to execute, and the less-motivated participants may simply not have wanted to put in the full effort required to succeed with those strategies. If this were the case, we would expect a positive correlation across participants, between mean (correct) response time and accuracy. These correlations were non-significant for each of the six experimental groups, uncorrected ($|r| < 0.17$, $p > 0.1$, BF< 0.3). Contrary to the hypothesis, the peg strategies took longer to carry out, despite underperforming the inter-item imagery strategies.

General Discussion

We review our results in light of our main goals, to test whether participants were able to apply peg-list strategies to lists of pairs, and to test whether the superiority of peg-list over interactive imagery might generalize from serial recall (Roediger, 1980) to cued recall of pairs, as is expected based on theories about anchoring new information to prior knowledge. We then discuss our findings with respect to mathematical models of memory for associations and order.

Support for feasibility of peg-based strategies but with less success than inter-item imagery

Participants were able to apply two variants of peg-list strategies to perform well on cued recall of a set of word-pairs. However, in contrast to serial recall (Roediger, 1980), for which the peg list method clearly exceeded the link method, none of our peg-list variants surpassed the level of accuracy produced by participants using simple interactive imagery (note that the comparison between cued recall of Interactive Imagery and the One Peg/Pair groups was across experiments). This is despite a) peg list participants being rewarded for their peg-recall, which if anything, may have increased their motivation, b) peg-list participants taking longer to respond, and c) peg-list participants requiring pre-training on the peg list, itself. Participants given a serial-list framing (Peg Serial List and Link Method groups) did not reverse the superiority of inter-item imagery over peg list strategies, ruling out the study phase as an explanation for why our findings appeared opposite to Roediger's. Interactive imagery apparently remains champion when it comes to cued recall of word pairs, with the added benefit of requiring only a very simple verbal instruction and no training phase.

Thus, the idea, raised in the introduction, that peg-list strategies may benefit from anchoring new information onto old knowledge is not supported. For inexperienced memorizers, interactive imagery requires less cognitive infrastructure and produces high levels of accuracy. Even when both pegs of a pair were fully known, cued-recall accuracy was not perfect, nor even up to the level of the (non-selected) Interactive Imagery participants. It could be the case that peg–item associations are asymmetric, and item→peg accuracy does not determine peg→item accuracy (Caplan, 2005). Alternatively, if peg–item associations are symmetric, this would imply that participants are not consistently able or willing to complete the full inference from probe word → probe peg word → target peg word → target word. It is, of course, possible that with substantial additional training, peg-based strategies could surpass interactive imagery— an interesting possibility that could be tested in future studies. Our findings showed that prior knowledge of the peg list, itself, improved successful application of the peg strategies, so it is plausible that further training on the peg list, as well as further practice applying the strategy (Foer, 2011) could produce substantial further improvements. The superiority of inter-item imagery for memory for verbal associations might be general, but our findings leave open the possibility that this result is limited to participants with a small amount of practice with the strategies (tens of minutes). For naïve participants, our results suggest that inter-item imagery requires less training and less practice than item–peg imagery-based strategies to support very high levels of accuracy in cued recall.

Reconsidering the findings of Roediger (1980), when serial recall was scored with a lenient criterion, i.e., disregarding order of report, the Link and Peg groups produced equivalent accuracy. Only when scored strictly (an item was correct only if recalled in the correct position) did the Peg group outperform the Link group. Thus, the particular benefit of peg-list based strategies may be when precise serial order is required. Although this does not explain why our peg groups did not reach the level of the Interactive Imagery group, the lack of demand for order may explain why peg strategies were not optimal; their advantages did not offset their costs. An additional study could test whether peg-list strategies confer an advantage to association-memory when order of items within pairs is required, as with order-recognition (Greene & Tussing, 2001; Kato & Caplan, 2017; Kounios et al., 2001, 2003; Yang et al., 2013). Alternatively, given that our lists were half the length (5 pairs = 10 words, compared to Roediger's 20-word serial lists), it is possible that a superiority of peg-list over interactive imagery strategies would emerge for cued recall of longer lists of pairs.

Strategies are not process-pure. In particular, the Peg-strategy participants, especially the One Peg/Pair participants, reported also using inter-item imagery. If anything, the superiority of One Peg/Pair over the other peg groups suggests that a combination strategy is a boon to cued recall, but the presence of inter-item images did not compromise either item–peg learning or accuracy in resulting cued recall. The self-report responses at least suggest that modellers of cued-recall data may need to seriously consider mixture models.

Implications for mathematical models

Positional coding for cued recall of pairs. Although initially proposed as a logical possibility in mathematical modelling work (Caplan, 2005; Caplan et al., 2006; Howard et al., 2009), the current findings provide empirical support for the idea that participants can, indeed, learn lists of word pairs using a strategy resembling positional-coding models, the peg list method (both experiments). This proof of principle suggests that modellers may need to keep positional-coding accounts of association-memory among the set of models considered as accounts of experimental data from typical experiments where participants are not instructed to use any particular strategy.

Contiguity effects. Distance functions are thought to be diagnostic of model variants (Henson et al., 1996; Henson, 1998; Hurlstone et al., 2014; Shiffrin & Cook, 1978). Distance effects were prominent in peg recall, showing that the peg list is not special in this regard, and echoes contiguity effects (Henson's "locality constraint") that positional and order-coding models produce (Henson, 1998; Hurlstone et al., 2014; Kahana, 2012; Neath & Brown, 2007). However, in the cued-recall data presented here, distance effects were clearly present, regardless of strategy. This suggests either that contiguity effects are not a diagnostic of models, or the five strategies examined here are equivalent in that regard.

Associative symmetry. Caplan (2005) proposed that the high correlation between forward and backward recall of associations is due to associations being relatively isolated from the remaining studied items. When embedded within a serial list, substantial levels of competition from nearby items are introduced. If different sources of competition predominate for forward than for backward cued recall, that should decouple forward and backward cued recall, reducing the value of $Q_{\text{Different}}$. Our hypothesis, then, was that the serial-list groups (Peg SL and Link Method) should have reduced $Q_{\text{Different}}$ than corresponding participants who were told to conceptualize the list as a set of pairs. In fact, $Q_{\text{Different}}$ did not differ significantly by strategy whatsoever. This either challenges Caplan's so-called Isolation Principle account of associative symmetry, or else implies that something makes these effects too small relative to the statistical sensitivity of the current data. Given the strong presence of contiguity effects in the current data compared to prior findings (cf. Figure 6 in Caplan et al., 2006), there may, indeed, be more competition from nearby list items that decorrelates forward from backward cued recall directions. However, this was not accompanied by particularly low $Q_{\text{Different}}$ values, as would have been predicted by the Isolation Principle account.

Within-pair order. The zig-zag characteristic, that indicates the use of within-pair position as a retrieval cue, was found in all groups, including the two inter-item imagery groups. This indicates that inter-item associations, at least when mediated by visual imagery, must include some representation of relative position, converging with findings on memory for within-pair order (Greene & Tussing, 2001; Kato & Caplan, 2017; Kounios et al., 2001, 2003; Rehani & Caplan, 2011; Yang et al., 2013).

Different modes of operation between lists and pairs. Murdock and Franklin (1984) argued that lists of pairs and serial lists are mutually exclusive modes of operation. Within their chaining model, inter-item associations were either present or absent between pairs, which they found to be consistent with their empirical findings. One may thus have expected that conceptualising the study set as a serial list versus a set of pairs would severely hurt cued-recall accuracy. However, although the Link Method group was nominally worse than the Interactive Imagery group, this difference did not reach significance. Deviating from Murdock and Franklin, including inter-pair images was not a huge liability for a test of association-memory. For the peg strategies, although formally, participants could have been doing quite the same thing, the disadvantage of the Peg Serial List group was significant, relative to the One Peg/Pair and Two Pegs/Pair groups. Thus, preparing to be tested between-pair (i.e., not distinguishing “pairs”) may have produced some confusion in applying the peg strategy to cued recall. Alternatively, awareness of the constraints of the task (probes should only be expected within-pair) may have been useful to participants, for example, making educated guesses (see Vicente & Wang, 1998). In any case, the inferiority of peg-based strategies was not ameliorated when participants were instructed to study the set as a serial list, ruling out one potential explanation of the divergence of our results from those of Roediger (1980).

The rhyming peg list, itself

Finally, to our knowledge, ours is the first report of acquisition of a peg list, itself. Serial learning data revealed that about one third of participants essentially knew the peg list prior to the experiment, making zero errors in recall of the pegs in order. However, there were also about a third of participants who took numerous trials, or failed to master the peg list perfectly. This subject variability did, indeed, translate into accuracy when applying peg-based strategies to cued recall (Figure 8) and it would not be surprising if the same applied to serial-recall performance in a future study. Thus, mastery, and arguably, time of mastery (e.g., just prior to application versus days or weeks before) is an important factor to take into account when investigating peg-based memory strategies. Apart from prior knowledge, serial-position effects during learning showed the classic primacy-dominant characteristic that has been found across a very broad range of serial-recall tasks, both episodic and semantic, adding further support to the generality of distinctiveness-based accounts of serial-order memory (Brown et al., 2007; Kelley, Neath, & Surprenant, 2013; Neath, 2010; Neath & Saint-Aubin, 2011; Overstreet, Healy, & Neath, 2017).

In sum, participants are able to apply peg list strategies to perform well in cued recall, although not to the level of untrained participants applying interactive imagery. The rhyming peg list produces contiguity effects that are central to the operation of position/order-coding models, suggesting that peg-list strategies may be useful to guide and develop mathematically implemented positional-coding models in the future. Strategies based on peg lists and those based on inter-item imagery, although formally quite different, and echoing different mathematical models, produce qualitative features that are more similar (distance effects, associative symmetry and evidence of cueing with within-pair position) than different.

Conclusion

In sum, peg-list strategies can, indeed, be applied to cued recall of lists of pairs, but unlike serial recall, at least with relatively short lists and minimal experience with the strategies, interactive imagery is superior, which is at odds with the hypothesis that peg-based methods achieve their superiority in serial recall due to the benefits of anchoring list-items to prior knowledge. Due to their formal similarity to mathematical models, the kinds of instructable, imagery-based strategies studied here lead to insights about the diagnostic value of empirical findings within ongoing mathematical modelling research, including findings that suggest contiguity effects and within-pair position-cuing are general, and unlikely to be clearly diagnostic of chaining versus positional coding models. In turn, the mathematical-modelling perspective led to additional insights into the functioning of those subjectively applied strategies, such as that associations mediated by interactive imagery incorporate some amount of within-pair positional information that can be used as a part of the retrieval cue. Finally, regarding the rhyming peg peg list, itself, this list exhibits the standard serial-position curve one typically finds with serial recall, but is virtually mastered prior to participation by about one third of our participants. Our novel report of peg-list exhibiting standard contiguity-based errors suggests that it resembles current positional-coding models and comparisons of peg-list versus link method strategies may continue to inform the development of positional-coding, associative chaining and hybrid or mixture models of list memory.

References

Asch, S. E., & Ebenholtz, S. M. (1962). The principle of associative symmetry. *Proceedings of the American Philosophical Society*, *106*(2), 135-163.

Baayen, R., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX Lexical Database (Release 2) [CD-ROM].* Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania [Distributor].

Bellezza, F. S. (1981). Mnemonic devices: classification, characteristics, and criteria. *Review of Educational Research*, *51*(2), 247-275.

Bower, G. H. (1970). Imagery as a relational organizer in associative learning. *Journal of Verbal Learning and Verbal Behavior*, *9*, 529-533.

Bower, G. H., & Reitman, J. S. (1972). Mnemonic elaboration in multilist learning. *Journal of Verbal Learning and Verbal Behavior*, *11*, 478-485.

Bower, G. H., & Winzenz, D. (1970). Comparison of associative learning strategies. *Psychonomic Science*, *20*(2), 119-120.

Brown, G. D. A., Neath, I., & Chater, N. (2007). A temporal ratio model of memory. *Psychological Review*, *114*(3), 539-576.

Bugelski, B. R. (1968). Images as mediators in one-trial paired-associate learning. II: self-timing in successive lists. *Journal of Experimental Psychology*, *77*(2), 328-334.

Bugelski, B. R., Kidd, E., & Segmen, J. (1968). Image as a mediator in one-trial paired-associate learning. *Journal of Experimental Psychology*, *76*(1), 69-73.

Burgess, N., & Hitch, G. J. (1999). Memory for serial order: A network model of the phonological loop and its timing. *Psychological Review*, *106*(3), 551-581.

Caplan, J. B. (2005). Associative isolation: unifying associative and order paradigms. *Journal of Mathematical Psychology*, *49*(5), 383-402.

Caplan, J. B. (2015). Order-memory and association-memory. *Canadian Journal of Experimental Psychology*, *69*(3), 221-232.

Caplan, J. B., Boulton, K. L., & Gagné, C. L. (2014). Associative asymmetry of compound words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(4), 1163-1171.

Caplan, J. B., Glaholt, M., & McIntosh, A. R. (2006). Linking associative and list memory: pairs versus triples. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(6), 1244-1265.

Caplan, J. B., & Madan, C. R. (2016). Word-imageability enhances association-memory by increasing hippocampal engagement. *Journal of Cognitive Neuroscience*, *28*(10), 1522-1538.

Caplan, J. B., Madan, C. R., & Bedwell, D. J. (2015). Item-properties may influence item–item associations in serial recall. *Psychonomic Bulletin & Review*, *22*(2), 483-489.

Caplan, J. B., Rehani, M., & Andrews, J. C. (2014). Associations compete directly in memory. *Quarterly Journal of Experimental Psychology*, *67*(5), 955-978.

Carney, R. N., & Levin, J. R. (2011). Delayed mnemonic benefits for a combined pegword–keyword strategy, time after time, rhyme after rhyme. *Applied Cognitive Psychology*, *25*, 204-211.

Conrad, R. (1960). Serial order intrusions in immediate memory. *British Journal of Psychology*, *51*(1), 45-48.

Davis, O. C., Geller, A. S., Rizzuto, D. S., & Kahana, M. J. (2008). Temporal associative processes revealed by intrusions in paired-associate recall. *Psychonomic Bulletin & Review*, *15*(1), 64-69.

de Beni, R., & Cornoldi, C. (1988). Does the repeated use of loci create interference? *Perceptual and Motor Skills*, *67*, 415-418.

Delin, P. S. (1969). The learning to criterion of a serial list with and without mnemonic instructions. *Psychonomic Science*, *16*(4), 169-170.

Delprato, D. J., & Baker, E. J. (1974). Concreteness of peg words in two mnemonic systems. *Journal of Experimental Psychology*, *102*(3), 520-522.

Ebbinghaus, H. (1885/1913). *Memory: A contribution to experimental psychology.* New York: Teachers College, Columbia University.

Farrell, S. (2012). Temporal clustering and sequencing in short-term memory and episodic memory. *Psychological Review*, *119*(2), 223-271.

Farrell, S., & Lewandowsky, S. (2003). Dissimilar items benefit from phonological similarity in serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(5), 838-849.

Foer, J. (2011). *Moonwalking with Einstein: the art and science of remembering everything.* New York, NY, USA: Penguin Press.

Greene, R. L., & Tussing, A. A. (2001). Similarity and associative recognition. *Journal of Memory and Language*, *45*, 573-584.

Harris, J. E. (1980). Memory aids people use: two interview studies. *Memory & Cognition*, *8*(1), 31-38.

Harris, L. J., & Blaiser, M. J. (1997). Effects of a mnemonic peg system on the recall of daily tasks. *Perceptual and Motor Skills*, *84*, 721-722.

Henson, R. N. A. (1998). Short-term memory for serial order: the Start-End Model. *Cognitive Psychology*, *36*(2), 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*(1), 80-115.

Hintzman, D. L. (1980). Simpson's paradox and the analysis of memory retrieval. *Psychological Review*, *87*(4), 398-410.

Hockley, W. E., & Cristi, C. (1996a). Tests of encoding tradeoffs between item and associative information. *Memory & Cognition*, *24*, 202-216.

Hockley, W. E., & Cristi, C. (1996b). Tests of the separate retrieval of item and associative information using a frequency-judgment task. *Memory & Cognition*, *24*(6), 796-811.

Howard, M. W., Jing, B., Rao, V. A., Provyn, J. P., & Datey, A. V. (2009). Bridging the gap: transitive associations between items presented in similar temporal contexts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(2), 391-407.

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*(4), 923-941.

Hurlstone, M. J., Hitch, G. J., & Baddeley, A. D. (2014). Memory for serial order across domains: an overview of the literature and directions for future research. *Psychological Bulletin*, *140*(2), 339-373.

JASP Team. (2016). *JASP (Version 0.8.1.1)[computer software].* Retrieved from `https://jasp-stats.org [jasp-stats.org]`

Kahana, M. J. (2002). Associative symmetry and memory theory. *Memory & Cognition*, *30*(6), 823-840.

Kahana, M. J. (2012). *Foundations of human memory.* New York: Oxford University Press.

Kahana, M. J., & Caplan, J. B. (2002). Associative asymmetry in probed recall of serial lists. *Memory & Cognition*, *30*(6), 841-849.

Kahana, M. J., Mollison, M. V., & Addis, K. M. (2010). Positional cues in serial learning: the spin-list technique. *Memory & Cognition*, *38*(1), 92-101.

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Society*, *90*(430), 773-795.

Kato, K., & Caplan, J. B. (2017). Order of items within associations. *Journal of Memory and Language*, *97*, 81-102.

Kelley, M. R., Neath, I., & Surprenant, A. M. (2013). Three more semantic serial position functions and a SIMPLE explanation. *Memory & Cognition*, *41*, 600-610.

Kounios, J., Bachman, P., Casasanto, D., Grossman, M., & Smith, W., Roderick W. Yang. (2003). Novel concepts mediate word retrieval from human episodic associative memory: evidence from event-related potentials. *Neuroscience Letters*, *345*, 157-160.

Kounios, J., Smith, R. W., Yang, W., Bachman, P., & D'Esposito, M. (2001). Cognitive association formation in human memory revealed by spatiotemporal brain imaging. *Neuron*, *29*, 297-306.

Krinsky, R., & Krinsky, S. G. (1994). The peg-word mnemonic facilitates immediate but not long-term memory in fifth-grade children. *Contemporary Educational Psychology*, *19*, 217-229.

Ladd, G. T., & Woodworth, R. S. (Eds.). (1911). *Elements of physiological psychology: A treatise of the activities and nature of the mind from the physical and experimental point of view.* New York, NY: Charles Scribner's Sons.

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.

Legge, E. L. G., Madan, C. R., Ng, E. T., & Caplan, J. B. (2012). Building a memory palace in minutes: equivalent memory performance using virtual versus conventional environments with the Method of Loci. *Acta Psychologica*, *141*(3), 380-390.

Lieberman, D. A. (2011). *Human learning and memory.* Cambridge University Press.

Lindsey, D. R. B., & Logan, G. D. (2019). Item-to-item associations in typing: evidence from spin list sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(3), 397-416.

Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review*, *1*(4), 476-490.

Lowry, D. H. (1974). The effects of mnemonic learning strategies on transfer, interference, and 48-hour retention. *Journal of Experimental Psychology*, *103*(1), 16-20.

Madan, C. R., Glaholt, M. G., & Caplan, J. B. (2010). The influence of item properties on association-memory. *Journal of Memory and Language*, *63*(1), 46-63.

McGee, R. (1980). Imagery and recognition memory: the effects of relational organization. *Memory & Cognition*, *8*(5), 394-399.

Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of behavior*. New York: Holt.

Murdock, B. B. (1974). *Human memory: theory and data*. New York: John Wiley & Sons.

Murdock, B. B., & Franklin, P. E. (1984). Associative and serial-order information: different modes of operation? *Memory & Cognition*, *12*(3), 243-249.

Neath, I. (2010). Evidence for similar principles in episodic and semantic memory: the presidential serial position function. *Memory & Cognition*, *38*(5), 659-666.

Neath, I., & Brown, G. D. A. (2007). Making distinctiveness models of memory distinct. In J. S. Nairne (Ed.), *The foundations of remembering: essays in honor of Henry L. Roediger III* (p. 125-140). Psychology Press.

Neath, I., & Saint-Aubin, J. (2011). Further evidence that similar principles govern recall from episodic and semantic memory: the Canadian prime ministerial serial position function. *Canadian Journal of Experimental Psychology*, *65*(2), 77-83.

O'Kane, G., Kensinger, E. A., & Corkin, S. (2004). Evidence for semantic learning in profound amnesia: an investigation with patient H.M. *Hippocampus*, *14*, 417-425.

Overstreet, M. F., Healy, A. F., & Neath, I. (2017). Further differentiating item and order information in semantic memory: students' recall of words from the "CU Fight Song", Harry Potter book titles, and Scooby Doo theme song. *Memory*, *25*(1), 69-83.

Paivio, A. (1969). Mental imagery in associative learning and memory. *Psychological Review*, *76*(3), 241-263.

Paivio, A. (1971). *Imagery and verbal processes*. New York: Holt, Rinehart and Winston, Inc.

Paivio, A., & Yuille, J. C. (1969). Changes in associative strategies and paired-associate learning over trials as a function of word imagery and type of learning set. *Journal of Experimental Psychology*, *79*(3), 458-463.

Quinn, J. G., & McConnell, J. (1996). Irrelevant pictures in visual working memory. *Quarterly Journal of Experimental Psychology*, *49A*(1), 200-215.

Quinn, J. G., & McConnell, J. (1999). Manipulation of interference in the passive visual store. *European Journal of Cognitive Psychology*, *11* (3), 373-389.

Rehani, M., & Caplan, J. B. (2011). Interference and the representation of order within associations. *Quarterly Journal of Experimental Psychology*, *64* (7), 1409-1429.

Rizzuto, D. S., & Kahana, M. J. (2001). An autoassociative neural network model of paired-associate learning. *Neural Computation*, *13*, 2075-2092.

Roediger, H. L. I. (1980). The effectiveness of four mnemonics on ordering recall. *Journal of Experimental Psychology: Human Learning and Memory*, *6* (5), 558-567.

Sharon, T., Moscovitch, M., & Gilboa, A. (2011). Rapid neocortical acquisition of long-term arbitrary associations independent of the hippocampus. *Proceedings of the National Academy of Sciences, USA*, *108* (3), 1146-1151.

Shiffrin, R. M., & Cook, J. R. (1978). Short-term forgetting of item and order information. *Journal of Verbal Learning and Verbal Behavior*, *17*, 189-218.

Skotko, B. G., Kensinger, E. A., Locascio, J. J., Einstein, G., Rubin, D. C., Tupler, L. A., ... Corkin, S. (2004). Puzzling thoughts for H.M.: can new semantic information be anchored to old semantic memories? *Neuropsychology*, *18* (4), 756-769.

Smith, R. K., & Noble, C. E. (1965). Effects of a mnemonic technique applied to verbal learning and memory. *Perceptual and Motor Skills*, *21*, 123-134.

Solway, A., Murdock, B. B., & Kahana, M. J. (2012). Positional and temporal clustering in serial order memory. *Memory & Cognition*, *40* (2), 177-190.

Sommer, T. (2017). The emergence of knowledge and how it supports the memory for novel related information. *Cerebral Cortex*, *27* (3), 1906-1921.

Tse, D., Langston, R. F., Kakeyama, M., Bethus, I., Spooner, P. A., Wood, E. R., ... Morris, R. G. M. (2007). Schemas and memory consolidation. *Science*, *316*, 76-82.

Tse, D., Takeuchi, T., Kakeyama, M., Kajii, Y., Okuno, H., Tohyama, C., ... Morris, R. G. M. (2011). Schema-dependent gene activation and memory encoding in neocortex. *Science*, *333* (6044), 891-895.

Ueno, T., & Saito, S. (2013). The role of visual representations within working memory for paired-associate and serial order of spoken words. *Quarterly Journal of Experimental Psychology*, *66*(9), 1858-1872.

Vicente, K. J., & Wang, J. H. (1998). An ecological theory of expertise effects in memory recall. *Psychological Review*, *105*(1), 33-57.

Warrens, M. J. (2008). On association coefficients for 2 × 2 tables and properties that do not depend on the marginal distributions. *Psychometrika*, *73*(4), 777-789.

Wood, G. (1967). Mnemonic systems in recall. *Journal of Educational Psychology*, *58*(6, Part 2), 1-27.

Yang, J., Zhao, P., Zhu, Z., Mecklinger, A., Fang, Z., & Han, L. (2013). Memory asymmetry of forward and backward associations in recognition tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(1), 253-269.

Yates, F. A. (1966). *The art of memory.* London: Routledge & Kegan Paul.

Group	No	Sometimes	Yes	Abstention
One Peg/Pair	24	27	42	2
Two Pegs/Pair (Exp. 1)	49	13	35	2
Peg Serial List	47	16	35	0
Two Pegs/Pair (Exp. 2)	41	12	22	1

Table 1

Rates of confessions by peg-strategy participants of forming direct associations between list items.

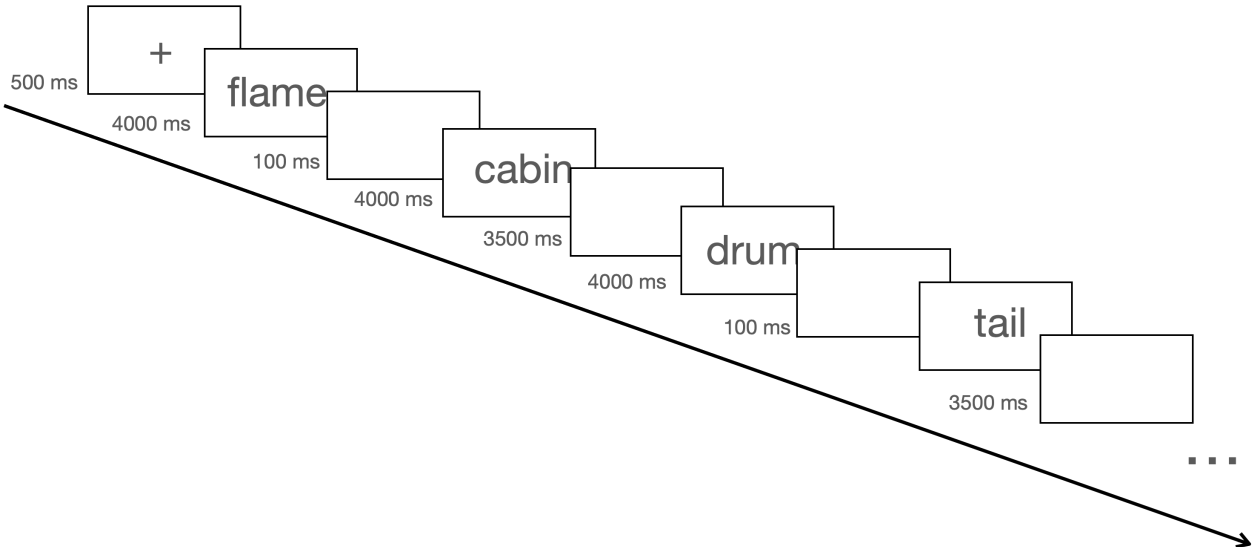


Figure 1. Illustration of the timing during the study phase, displayed here for the first two pairs of an example list.

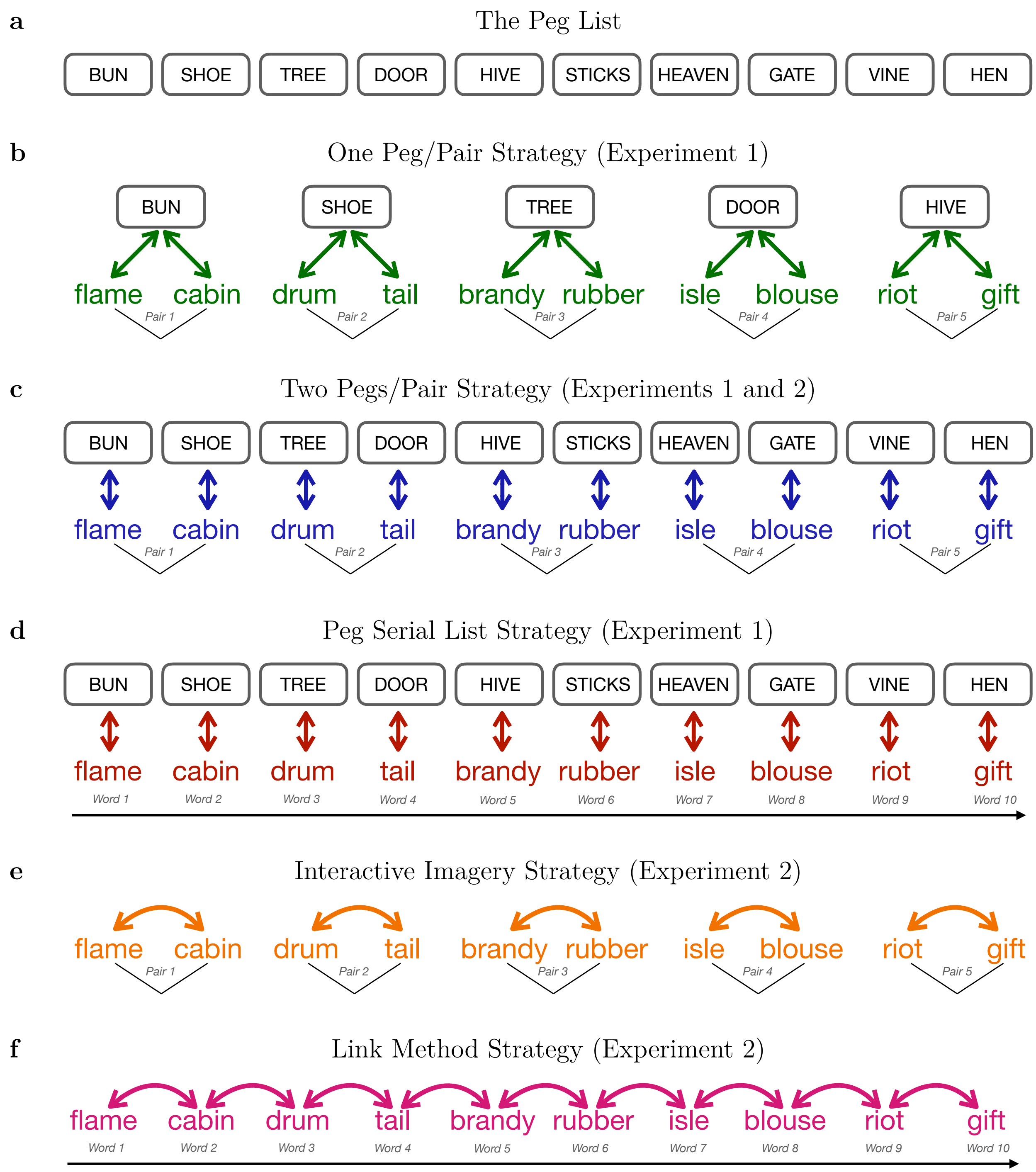


Figure 2. a) Rhyming Peg list. Schematic representations of the strategies: b) One Peg/Pair: each word of a pair imagined with a single peg word; c) Two Pegs/Pair: words within a pair linked to different pegs; d) Peg Serial List: like Two Pegs/Pair, but framed as a serial list; e) Interactive Imagery: images formed between the paired words; and f) Link Method: framed as a serial list, images formed between all nearest-neighbour words. Double-arrow lines denote interactive images.

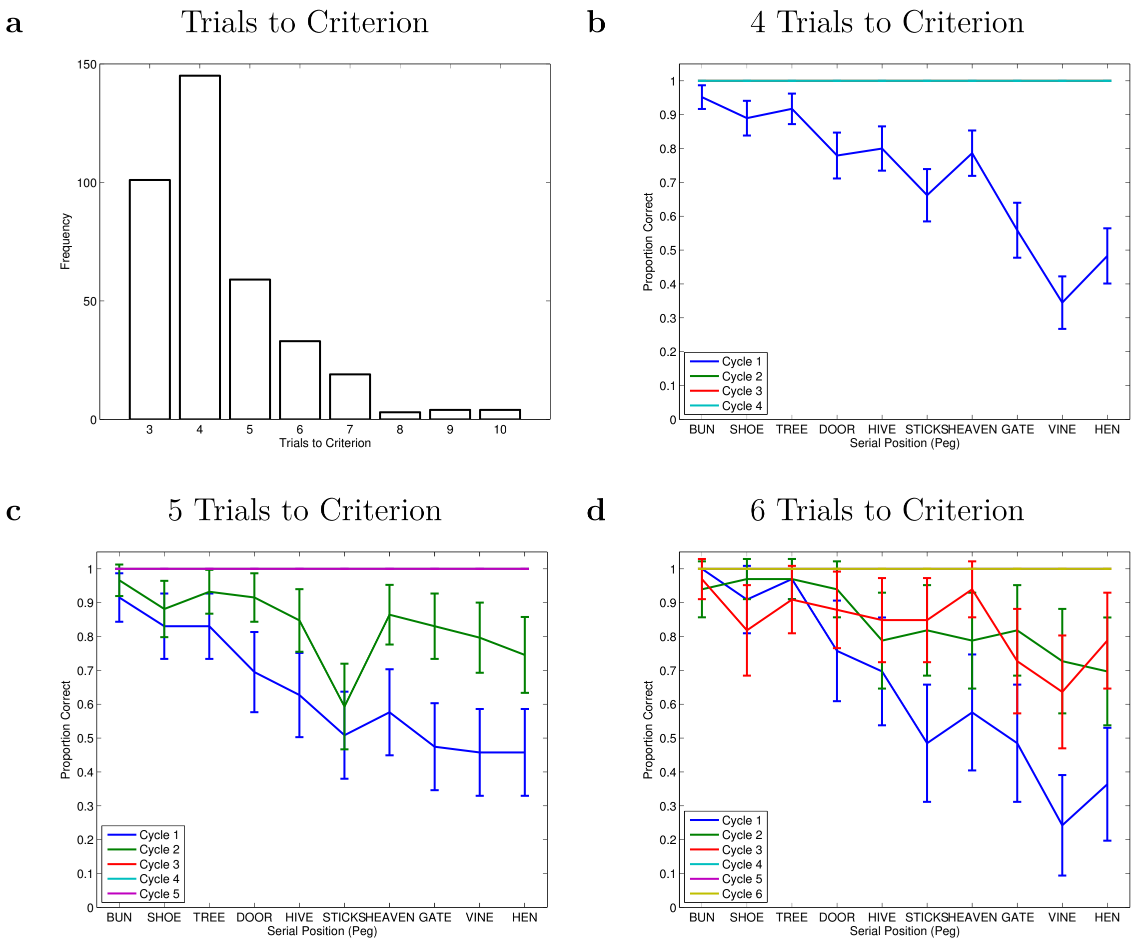


Figure 3. Peg-list learning. (a) Distribution of trials to criterion, collapsed across all groups. Serial position curves (strict serial-position scoring) are plotted for each trial, separated by participants who mastered the peg list in the following number of trials: (b) 4, (c) 5, or (d) 6. Note that because of the criterion (three perfect consecutive recalls), the last two cycles are obscured by the third-last. Error bars plot standard error of the mean, corrected for subject variability (Loftus and Masson, 1994).

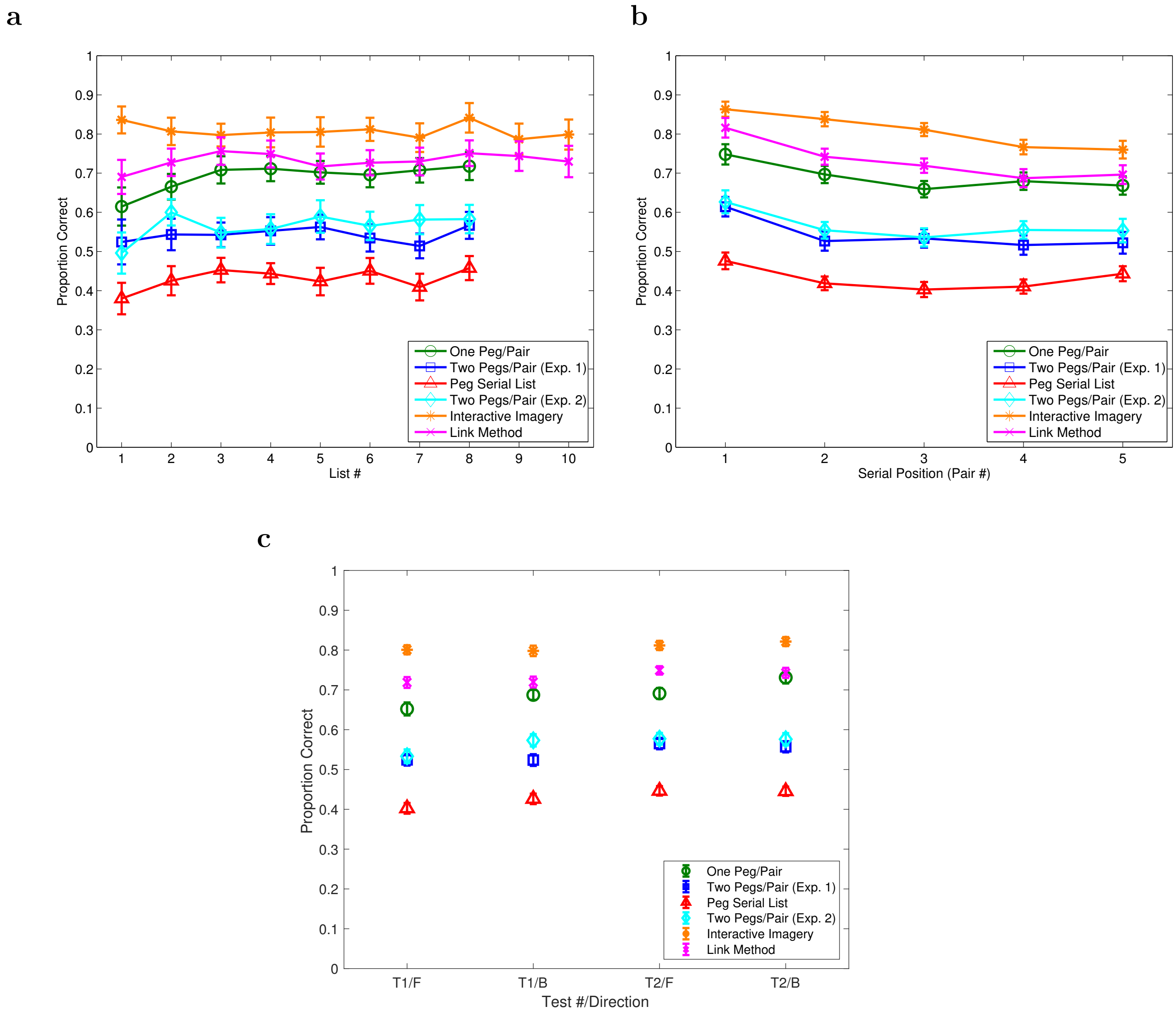
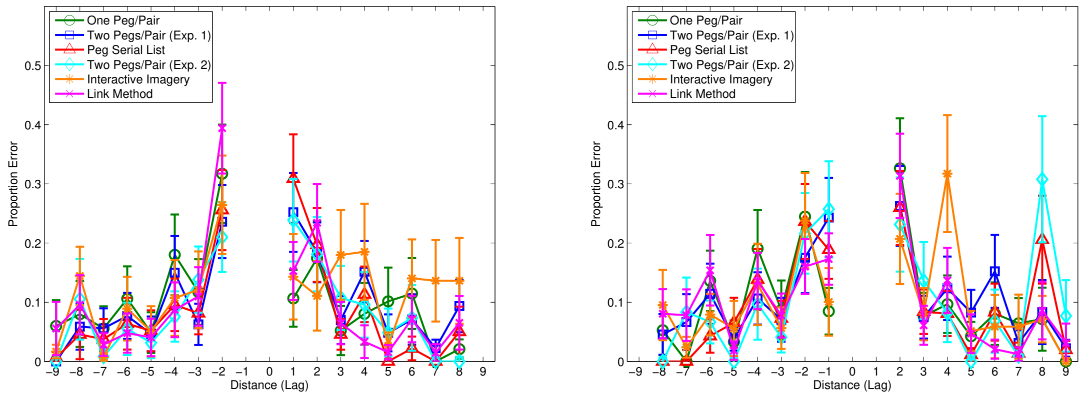


Figure 4. Cued recall accuracy, as functions of (a) list number, (b) serial position of pair, (c) Test Number×Direction, while collapsing over the remaining factors. Error bars plot standard error of the mean, corrected for subject variability (Loftus and Masson, 1994).

a Forward Cued Recall **b** Backward Cued Recall



c *Studied List:* **flame cabin drum tail brandy rubber isle blouse riot gift**

Cue Word	Available Lags	Response	Lag
tail	–3, –2, +1, +2, +3, +4, +5, +6	**cabin**	–2
brandy	–4, –3, –2, –1, +2, +3, +4, +5	**gift**	+5

Figure 5. Distance functions for within-list intrusions during cued recall in the (a) forward and (b) backward directions. Zero lag denotes the target position (i.e., correct responses, excluded from these distance functions); positive lags refer to intrusions of items from later in the list, relative to target (item-wise) serial position and negative lags refer to intrusions of items from earlier in the list, also relative to target serial position. Even-numbered lags are errors to other pairs, that nonetheless preserve the correct within-pair position of the target. Error bars are 95% confidence intervals based on standard error of the mean, controlling for subject variability (Loftus and Masson, 1994). (c) Illustration of how distance functions were computed. Given the studied list at the top of the panel, two hypothetical cued-recall trials are illustrated below. Assuming the Cue Word was presented to the participant, Available Lags shows the lags that are possible (i.e., excluding beyond the start and end of the list, and excluding the probe itself and the correct response, since only within-list errors enter into this analysis). The denominator at each Available Lag is incremented regardless of the Response. Given a particular Response, the Lag column shows the single lag where the numerator of the distance function value is incremented.

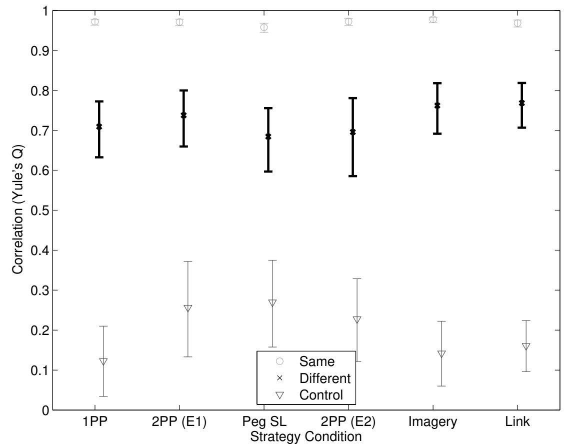


Figure 6. Correlations (Yule's Q) testing for associative symmetry (see Methods). Error bars plot 95% confidence intervals based on the log-odds transform.

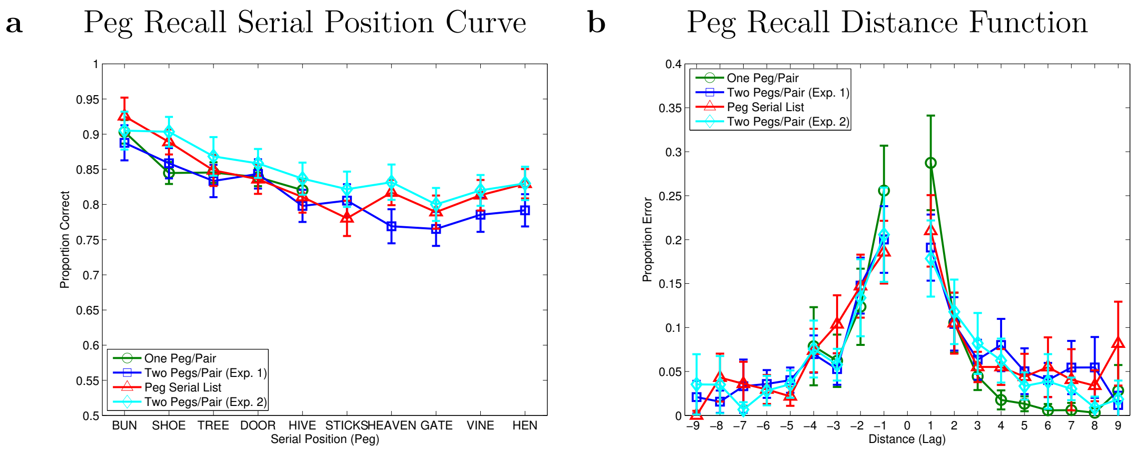


Figure 7. (a) Peg recall accuracy as a function of serial position. (b) Distance functions for peg-recall intrusions. Error bars are 95% confidence intervals based on standard error of the mean, controlling for subject variability (Loftus and Masson, 1994).

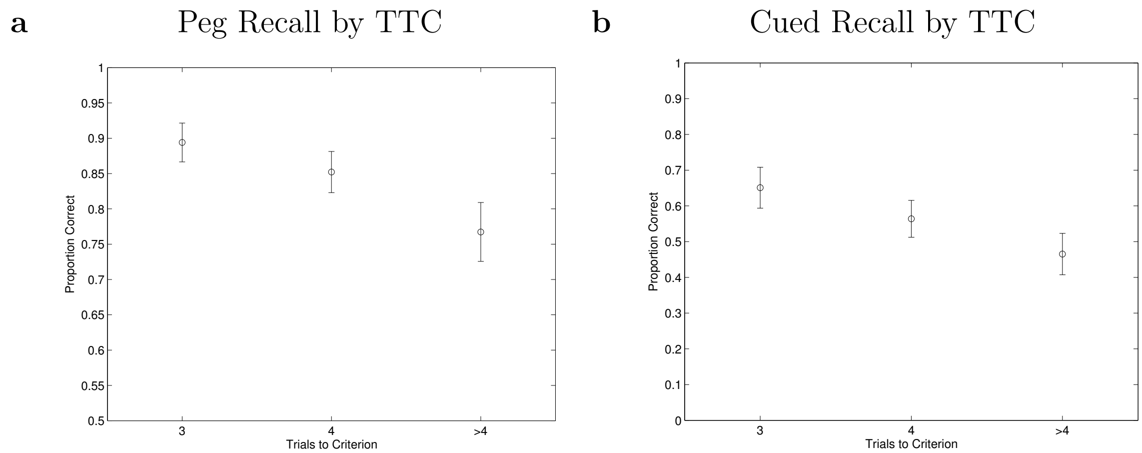


Figure 8. Effect of trials to criterion on (a) peg recall and (b) cued recall, collapsing across all Peg List strategy participants. Error bars are 95% confidence intervals based on standard error of the mean.

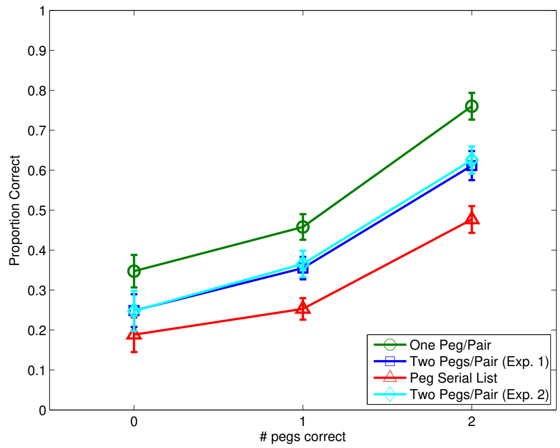


Figure 9. Cued-recall accuracy as a function of peg accuracy (0, 1 or 2 pegs correctly recalled for the pair during the subsequent peg-recall phase). Error bars are 95% confidence intervals based on standard error of the mean, corrected for between-subject variability (Loftus and Masson, 1994).