

Semantic relatedness proactively benefits learning, memory, and interdependence across episodes

Reviewed Preprint

v2 • November 14, 2024

Revised by authors

Reviewed Preprint

v1 • May 2, 2024

Kelly A Bennion , Jade Phong, Mytien Le, Kunhua Cheng, Christopher N Wahlheim, James W Antony 

Department of Psychology and Child Development, California Polytechnic State University, San Luis Obispo, United States • Department of Psychology, University of North Carolina at Greensboro, Greensboro, United States

 https://en.wikipedia.org/wiki/Open_access
 Copyright information

eLife Assessment

This **important** work advances our understanding of how memories interact to facilitate or interfere with one another, also informing our understanding of how humans build knowledge. The study provides **compelling** evidence that semantic relatedness proactively benefits memory using clean experimental design, rigorous statistics, large N samples, and well-characterized stimuli. The study also demonstrates the boundaries of these proactive benefits, showing that when studied items have weaker semantic relationships, proactive interference may be observed. This research will be of interest to memory researchers as well as cognitive psychologists, neuroscientists, and educators more broadly.

<https://doi.org/10.7554/eLife.95480.2.sa4>

Abstract

Over the past century of memory research, the interplay between initial and later-learned information in determining long-term memory retention has been of central interest. A likely candidate for determining whether initial and later memories interfere with or strengthen each other is semantic relatedness. Relatedness has been shown to *retroactively* boost initial memory and increase the interdependence between earlier and more recent experiences in memory. Here, we investigated the converse relationship of how relatedness *proactively* affects later memory for paired associates. In five experiments ($N = 1000$ total), we varied the relatedness between initial and later cues, initial and later targets, or both. Across experiments and conditions, relatedness profoundly benefited later-learned memories – in some conditions, low relatedness reliably produced proactive interference (versus a control condition) while high relatedness produced proactive facilitation within the same experiment. Additionally, relatedness also accelerated learning and increased interdependence between initial and later-learned pairs. In sum, we demonstrate the robust effects of relatedness in scaffolding memory for recently learned information and creating strong integrative links with prior experiences.

When one begins learning a set of associations in a new domain, the associations do not exist in a vacuum but rather affect each other in various ways. Consider the example that a novice bartender learns that a “martini” is made with *dry vermouth* and later that a “Manhattan” is made with *sweet vermouth*. In this case, previously learning about the martini could affect memory for the Manhattan by diminishing it [proactive interference (PI)] or improving it [proactive facilitation (PF)].

Here, we ask whether this outcome can be predicted by the *semantic relatedness* between the learning events. In so doing, we re-visit a classic proposal by Osgood (1949) [that](#) has been elaborated upon in the intervening decades (Bugelski & Cadwallader, 1956 [; Eich, 1982 \[; Mensink & Raaijmakers, 1988 \\[; Antony et al., 2022\\]\\(#\\)\]\(#\)](#)) that memory changes critically depend on the relatedness between events. A recent paper of ours (Antony et al., 2022) showed that later learning affected memory for prior events (i.e., retroactively) based on relatedness. The current paper can be considered a twin companion of that prior effort, in that we are keeping the same subject population, stimuli, experimental procedure, and timing, and we vary only the order of list learning to focus on proactive, rather than retroactive, effects. Additionally, we mirror that paper’s formatting and figures to ease comparisons among the studies, and we directly compare some data between the studies.

There are theoretical reasons to support directly opposing outcomes for how semantic relatedness with previously learned stimuli proactively affects new episodic memories (Brunmair & Richter, 2019). That is, relatedness could interfere with the new memory, promoting PI, or it could cause one to recall the prior memory and link it with the new one, which may make the new memory more resistant to forgetting and promote PF. In favor of the PI account, the presence of multiple targets for one cue could benefit from more mental separation among the targets (Underwood, 1969). In fact, increasing relatedness in some paradigms can increase interference (Bower et al., 1994 [; McGeoch & McDonald, 1931 \[; McGeoch & McGeoch, 1937\]\(#\)](#)) and/or intrusions (Dallett, 1962 [; 1964 \[; Osgood, 1946 \\[; Postman, 1961 \\\[; Underwood, 1951\\\]\\\(#\\\)\\]\\(#\\)\]\(#\) \), and intrusions have been shown to increase specifically with similarity in proactive paradigms \(Young, 1955 \). Conversely, in favor of the PF account, we consider two main, related theories. The first is the importance of “reminders” in memory, which involve reinstating representations from an earlier study phase during later learning \(Hintzman, 2011 \). This idea centers study-phase retrieval, which involves being able to mentally recall prior information and is usually applied to exact repetitions of the same material \(Benjamin & Tullis, 2010 \[; Hintzman et al., 1975 \\[; Siegel & Kahana, 2014 \\\[; Thios & D’Agostino, 1976 \\\\[; Zou et al., 2023\\\\]\\\\(#\\\\)\\\]\\\(#\\\)\\]\\(#\\) \\). However, reminders can occur upon the presentation of related \\(but not identical\\) material and can result in better memory for both prior and new information when memory for the linked events becomes more interdependent \\(Hintzman, 2011 \\[; Hintzman et al., 1975 \\\[; McKinley et al., 2019 \\\\[; McKinley & Benjamin, 2020 \\\\\[; Schlichting & Preston, 2017 \\\\\\[; Tullis et al., 2014 \\\\\\\[; Wahlheim & Zacks, 2019\\\\\\\]\\\\\\\(#\\\\\\\)\\\\\\]\\\\\\(#\\\\\\)\\\\\]\\\\\(#\\\\\) \\\\\). The second is the memory-for-change framework, which builds upon these ideas and argues that humans often retrieve prior experiences during new learning, either spontaneously by noticing changes from what was learned previously or by instruction \\\\\(Jacoby et al., 2015 \\\\\[; Jacoby & Wahlheim, 2013\\\\\]\\\\\(#\\\\\) \\\\\). The key advance of this framework is that recollecting changes is necessary for PF, whereas PI occurs without recollection. This framework has been applied to paradigms including stimulus changes, including common paired associate paradigms \\\\\(e.g., A-B, A-D\\\\\) that we cover extensively later. Because humans may be more likely to notice and recall prior information when it is more related to new information, these two accounts would predict that semantic relatedness instead promotes successful reminders, which would create PF and interdependence among the traces.\\\\]\\\\(#\\\\)\\\]\\\(#\\\)\\]\\(#\\)\]\(#\)](#)

One popular method for testing these accounts is to use paired associated learning paradigms in which original associations (A-B) change on a later presentation (A-D; Barnes & Underwood, 1959 [; Briggs, 1954](#)). Here, we refer to this A-B, A-D paradigm as Δ Target learning. Under conditions of no semantic relatedness between the B and D targets, Δ Target learning typically

causes PI in the form of impaired A-D memory relative to a control condition, likely because B and D compete at retrieval (Anderson & Neely, 1996 [↗](#); Bower et al., 1994 [↗](#); Caplan et al., 2014 [↗](#); Castro et al., 2002 [↗](#); Jung, 1962 [↗](#), 1967 [↗](#); Postman, 1962 [↗](#); Postman & Underwood, 1973 [↗](#); Twedt & Underwood, 1959 [↗](#); Young, 1955 [↗](#)). However, as the semantic relationship between B and D increases, PI decreases and, in some cases, PF occurs (Bruce, 1933 [↗](#); Dallett, 1964 [↗](#); Martin & Dean, 1964 [↗](#); Mehler & Miller, 1964 [↗](#); Metcalfe et al., 1993 [↗](#); Morgan & Underwood, 1950 [↗](#); Osgood, 1946 [↗](#); Russell & Storms, 1955 [↗](#)) [though see (Mills, 1973 [↗](#)) for one case in which PI increases]. An important early result using Δ Target learning showed PF only when participants performed a mental linking strategy, in which they thought back to and integrated the prior and new pairs (Martin & Dean, 1964 [↗](#)). Collectively, these findings indicate that prior learning can improve or impair memory for more recent information based on how people notice and consider relationships to prior information during new learning.

Over the last decade, several studies have rigorously addressed this issue by asking participants during new learning if they notice changes in new stimuli relative to existing memories, and if they can recall the existing memories. On later memory tests of the updated associations, participants are also asked to indicate if they recollect earlier changes. These studies consistently showed that noticing and recollecting changes are associated with enhanced memory for new associations (for a review, see Wahlheim et al., 2021 [↗](#)). Largely, these proactive benefits parallel those in the retroactive direction – that is, increased similarity between B and D also reduces retroactive interference (RI) or causes retroactive facilitation (RF) (Antony et al., 2022 [↗](#); Barnes & Underwood, 1959 [↗](#); Dallett, 1962 [↗](#); Kanungo, 1967 [↗](#); Osgood, 1946 [↗](#); Young, 1955 [↗](#)); we review the similarities and differences between proactive and retroactive operations more extensively below. In sum, proactive benefits of relatedness suggest that higher semantic relatedness may result in A-B and A-D becoming integrated (Wahlheim, 2014 [↗](#)), perhaps by promoting the extent to which changed associations trigger reminders of existing memories with both shared and unique features.

An association can be modified in ways other than changing the target. Indeed, in Osgood's proposal and in our prior paper (Antony et al., 2022 [↗](#)), we investigated how numerous associational changes affected original memories. Under A-B, C-B learning conditions, which we call Δ Cue learning, C-B memory relative to memory for unrelated stimuli (e.g., C-D) tends to be similar (recall condition of Postman et al., 1969 [↗](#); Traxler, 1973 [↗](#)) or slightly worse (Amundson et al., 2003 [↗](#); Postman et al., 1969 [↗](#); Twedt & Underwood, 1959b [↗](#); see Amundson et al. 2003 [↗](#) for a similar finding in rodents). Nevertheless, new learning of C-B associations can occur more quickly with Δ Cue learning (Mandler & Heinemann, 1956 [↗](#)), and it appears at least possible that under certain conditions, C-B could also show proactive facilitation in subsequent memory, perhaps by making B responses more accessible (Estes, 1979 [↗](#); Martin, 1965 [↗](#)). When instead investigating retroactive effects of Δ Cue learning, studies have found little-to-no RI when A and C are unrelated (Houston, 1966 [↗](#); Keppel & Underwood, 1962 [↗](#); Twedt & Underwood, 1959 [↗](#)) and even RF when A and C are semantically related (Antony et al., 2022 [↗](#); Bugelski & Cadwallader, 1956 [↗](#); Hamilton, 1943 [↗](#)).

Finally, when learning associations with completely new cues and targets (typically called A-B, C-D learning, but which we call Δ Both learning), the C-D memory is generally regarded as distinct from A-B. In the retroactive direction, when A is related to C and B related to D, RI has been observed when testing memory after a few minutes (Baddeley & Dale, 1966 [↗](#); Bugelski & Cadwallader, 1956 [↗](#); McGeoch & McGeoch, 1937 [↗](#); Saltz & Hamilton, 1967 [↗](#)). However, when both relationships are very strong (e.g., learning beer-late followed by keg-tardy) and a longer delay is used, we previously showed reliable RF (Antony et al., 2022 [↗](#)). These latter results also accord with the idea that, with high similarity between both items of the pair, participants may think back and recollect prior associations, which aids their long-term maintenance. However, no studies to our knowledge have used Δ Both learning with varying relationships between earlier and recent word pairs and tests of proactive effects of memory, as we do here.

To visualize this complicated collection of effects, Osgood proposed a three-dimensional surface by which old and new associations may interact to influence memory (Eich, 1982 [↗](#); Mensink & Raaijmakers, 1988 [↗](#)) (**Fig 1A** [↗](#)). Briefly, one can summarize Osgood's argument by carefully considering three continuous relationships between cues and targets. With a constant cue, one can draw a Δ Target line from an unrelated to an identical target (i.e., from A-B, A-D to A-B, A-B). Somewhere along this line, PI should shift to PF. Next, with a constant target starting from the A-B, A-D point (i.e., from A-B, A-D to A-B, C-D), PI should disappear as the pair becomes a wholly new association. Finally, with a constant target starting from the A-B, A-B point towards A-B, C-B (the Δ Cue line), Osgood posited that PI generally did not occur. Therefore, proactive effects went from no memory change at the A-B, C-B point to maximum PF at identity. Osgood also speculated how a full surface spanning all cue-target relatedness values would influence later memory (**Fig 1A** [↗](#)), though it was based on scant data.

Osgood largely considered proactive and retroactive effects of memory to be identical (or considered the data too insufficient to assert otherwise) – that is, he used the same surface to describe both effects. Nevertheless, there may be differences between them. First, if reminders retroactively benefit existing memories by evoking rehearsals (Antony et al., 2017 [↗](#); Roediger & Karpicke, 2006 [↗](#); Rowland, 2014 [↗](#)) and promoting associations across episodes, then the conditions that promote reminders (such as instructions to think back or strong relatedness) could benefit existing memories more than new ones. That is, the additional retrieval trial would only occur for old, and not new, associations. Conversely, if the sole mechanism behind these benefits is the interdependence across episodes, then proactive and retroactive effects may be equal. Second, by virtue of being learned second and subject to possible fatigue or prior learning effects (Postman & Keppel, 1977 [↗](#); Tulving & Watkins, 1974 [↗](#); Underwood, 1957 [↗](#)), newer associations may be more susceptible to interference than old associations when probed after a delay (but see Melton & Von Lackum, 1941 [↗](#) and Szpunar et al., 2008 [↗](#) for how this interference is mitigated by intermittent testing, which we use here). If it were the case that interference occurred more strongly for the second-learned list, we would expect worse memory for even the Δ Both or Control pairs in the proactive case than in the retroactive case.

In our prior paper, we examined how the full space of Osgood's predictions resulted in RI versus RF by systematically varying the semantic relatedness of cues, targets, or both. We did this using global vector similarity (GloVe) (Pennington et al., 2014 [↗](#)), a distributional semantic model, and associative strength (AS), collected using free association norms (Nelson et al., 1998 [↗](#)), to obtain quasi-continuous measures of relatedness. We also obtained a large dataset ($N = 1000$ across experiments) so that we could sample each point in the space under two different memory delays and two different average levels of relatedness. Here, we will use a nearly identical approach with a similar dataset to instead investigate *proactive* memory effects ($N = 800$ across the first four experiments and $N = 200$ for a final experiment performed for the purposes of comparing retroactive and proactive effects) (**Fig 1B** [↗](#)).

Briefly, subjects initially learned 36 unrelated word pairs, which we will call *supplemental pairs* (**Fig 1C** [↗](#)). We call the first learned list supplemental pairs because they are less important to our aims – our primary memories of interest were the *base* pairs, which were learned afterwards. Supplemental pairs belonged to four of the five within-participant conditions: pairs in the No Δ condition were the same as later-learned base pairs (e.g., sick-push), whereas other pairs belonged to the Δ Target (e.g., sick-shove), Δ Cue (e.g., ill-push), or Δ Both (e.g., ill-shove) conditions. Next, subjects learned 45 base pairs (e.g., sick-push), which also included pairs from a fifth (Control) condition that had no relationship to any supplemental pairs. In both blocks of learning, subjects learned each pair until they could correctly retrieve it once, after which it dropped out. Across every five subjects, pair assignments were counterbalanced, such that the same base pairs rotated among the five conditions (**Fig 1D** [↗](#)).

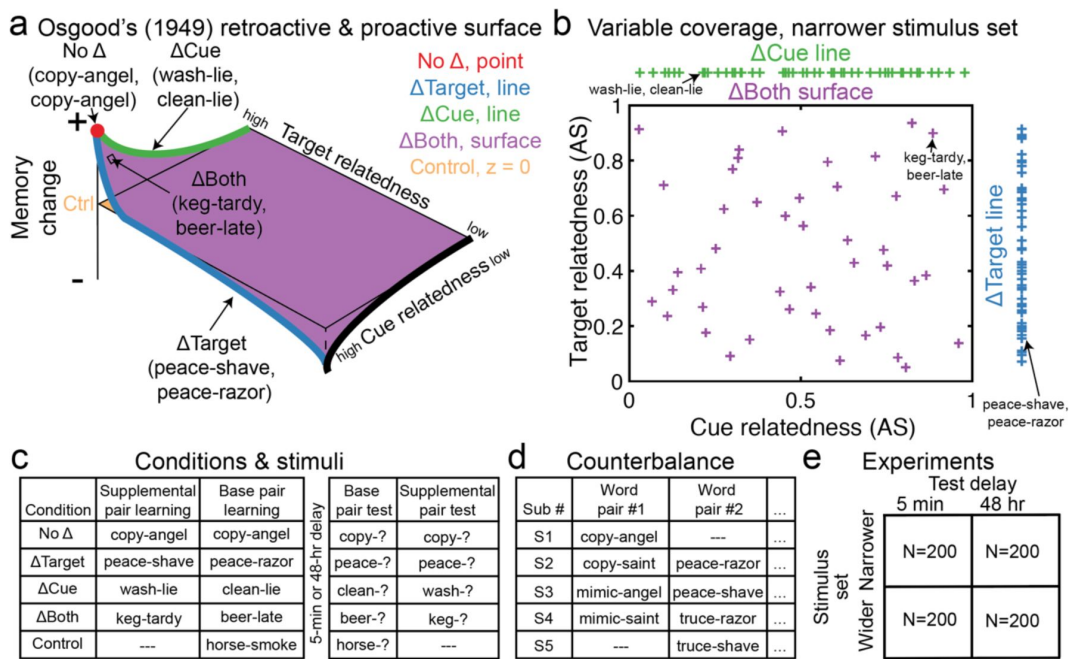


Figure 1

Overview of Osgood's predictions for proactive memory effects, distributions of relatedness values, conditions, stimuli, counterbalance procedures, and experiments. (a) We depict how our design would look if it confirmed Osgood's (1949) proposed surface. Cue and target relatedness are shown along the y- and x-axes, respectively, while memory change is shown along the z-axis, relative to the Control condition, which is simply depicted as the $z = 0$ surface. (b) In the narrower stimulus set, variables covered the full range of associative strength (AS) values. Plus signs along the Δ Cue (green) and Δ Target (blue) lines depict how cue and target relatedness were respectively distributed. Purple plus signs inside the scatterplot show how cue and target relatedness values were distributed in the Δ Both condition. (c) Pairs were divided into four experimental conditions during supplemental pair learning (which was performed first), followed by base pair learning, which included a fifth Control condition. After 5-mins or 48-hrs, base and supplemental pairs were tested (in that order). (d) Each base pair was included in a different condition for every five subjects. (e) The primary four experiments involved crossing delays and stimulus sets. In (a) and (b), example word pairs from (c) are labeled for explanatory purposes. See also **Fig 1-Supp 1** for visualizations using the stimulus set with wider semantic relatedness.

Across the first four experiments, we manipulated two factors: range of relatedness among the pairs and retention interval before the final test. The narrower range of relatedness used direct AS between pairs using free association norms, such that all pairs had between 0.03–0.96 association strength. Though this encompasses what appears to be a full range of relatedness values, pairs with even low AS are still related in the context of all possible associations (e.g., pious-holy has AS = 0.03 but would generally be considered related) (**Fig 1B**). The stimuli using a wider range of relatedness spanned the full range of global vector similarity (Pennington et al., 2014) that included many associations that would truly be considered unrelated (**Fig 1-Supp 1A**). One can see the range of the wider relatedness values in **Fig 1-Supp 1B** and comparisons between narrower and wider relatedness values in **Fig 1-Supp 1C**. Numerous studies have shown that delays affect the extent of interference versus facilitation in retroactive and proactive designs (Antony et al., 2022; Antony & Bennion, 2022; Baran et al., 2010; Chan, 2009; Liu & Ranganath, 2019; Lustig et al., 2004; Ortega et al., 2015; Wahlheim, C.N. et al., submitted; Wixted, 2004). Intriguingly, RI tends to decrease over time (Antony et al., 2022; Antony & Bennion, 2022; Baran et al., 2010; Chan, 2009; Liu & Ranganath, 2019; Lustig et al., 2004; Ortega et al., 2015; Wixted, 2004), whereas PI tends to grow over time via increased intrusions and/or worse memory (Greenberg & Underwood, 1950; Mensink & Raaijmakers, 1988; Postman et al., 1969; Underwood, 1948, 1949). We therefore employed two different delays of 5-min and 48-hr to assess how PI versus PF would change over time with varying levels of semantic relatedness (**Fig 1E**).

Overall, we predicted that relatedness would accelerate learning, benefit later memory, and increase interdependence of linked pairs across lists. Our first analyses focused on overall memory performance, based on proportion of words recalled, across each of five conditions. Following this, we asked whether this changed as a function of semantic relatedness along the respective dimension of each condition (e.g., B-D similarity in the Δ Target condition) and found the relative memorability of each base pair across subjects by comparing memory for that condition against the Control condition. This allowed us to eliminate the incidental memorability of each individual base pair and focus on how each condition and level of relatedness affected memory. We later visualized these effects as Osgood's proposed surface by plotting the No Δ condition at the cue and target identity point, the Δ Target condition as a line of values at cue identity spanning target relatedness, the Δ Cue condition as a line of values at target identity spanning cue relatedness, and the Δ Both condition as a surface spanning bivariate cue and target relatedness (see **Fig 1A** for the full Osgood surface; see **Fig 1B** for a visualization from the narrower stimulus set and **Fig 1-Supp 1** from the wider set). Later, we examined the speed of learning (based on the number of rounds to reach the criterion of one correct recall per trial) across conditions and levels of relatedness. Finally, we conducted a set of analyses comparing our proactive results from this paper to the retroactive results in our prior paper (Antony et al., 2022). While these results diverge from convention in that they compare groups that were not randomly assigned, they do involve the same stimuli, experimental procedure, timing, and subject population, and we use our control conditions to further adjust for any meaningful differences between the studies (other than learning order). We believe these similarities make these comparisons suggestive of important differences between proactive and retroactive effects, and we include appropriate caveats to avoid exceptionally strong conclusions. Additionally, to assess whether all retroactive versus proactive differences could be attributed to the fact that we measured memory for the second-learned (base pair) list first at the final test, we ran a control experiment in which base pairs were learned first and supplemental pairs second (that is, the design from our prior paper) while testing the supplemental pairs first. We then compared this condition to the same learning condition from our prior paper that tested the lists in the original order.

Results

High semantic relatedness resulted in proactive facilitation at a longer delay

Our first analyses focused on how the supplemental pair learning condition affected memory for the later-learned base pairs. In the first two experiments we will discuss, we selected a narrow range of semantic relatedness values using AS from free association norms, or the likelihood that a given word elicits a second word based on a large corpus (Nelson et al., 1998). We found supplemental pairs that predicted words from their respective base pairs using AS values of 0.03 (e.g., pious → holy) to 0.96 (e.g., moo → cow). We then tested memory either 5-min after or 48-hrs after base pair learning. We measured base pair memory performance across five conditions using a one-way (No Δ , Δ Target, Δ Cue, Δ Both, and Control) ANOVA.

Focusing first on the 5-min delay condition, we found a significant difference in performance across some conditions [$F(4,796) = 7.0, p < 0.001$] (Figure 2, top left) but no evidence of PI relative to the Control condition. Follow-up t -test patterns followed a No $\Delta = \Delta$ Cue $> \Delta$ Target = Δ Both = Control pattern (all significant differences indicate adjusted $p < 0.01$; “=” indicates $p > 0.27$). In the 48-hr delay experiment, condition again significantly affected overall memory [$F(4,796) = 116.3, p < 0.001$], and follow-up t -tests showed significant differences across all pairwise conditions, such that No $\Delta > \Delta$ Cue $> \Delta$ Target $> \Delta$ Both $> \Delta$ Control (all adjusted $p < 0.003$) (Figure 2, top right). The finding of higher recall in the conditions with changes relative to the Control condition suggests that the prior learning of semantically related words resulted in PF and that changing the cue benefitted memory more than changing the target. Interestingly, changing both cues and targets resulted in PF, even though there were no words in common across supplemental and base pairs. In tandem, these experiments showed that high semantic relatedness between earlier and more recently learned information created PF at longer delays. They also strongly resemble the retroactive patterns from our prior investigation (Antony et al., 2022) and generally support predictions from Osgood’s proposed surface (Osgood, 1949).

Condition and delay altered proactive effects when using a wider range of semantic relatedness

The above results showed either robust PF or no effect in the experimental conditions. Notably, although the Δ Target condition often generates PI (Anderson & Neely, 1996), we found none. Importantly, in these experiments, though cue- Δ cue and target- Δ target associations varied, they were strongly related overall. Thus, these results could have arisen because even the weakest association (e.g., pious → holy) had an AS of 0.03 and was still more related than the average word pair plucked from the overall semantic space of words. In the next two experiments, we therefore expanded semantic relatedness values to pairs spanning the full range of relatedness in the English language. To assess relatedness in this (wider) stimulus set, we found cosine similarity [$\cos(\theta)$] between GloVe vector embeddings (Pennington et al., 2014), which were trained on word-word co-occurrences in a large collection of texts and correspond with human judgments of similarity (Pennington et al., 2014). We found GloVe values distributed from -0.14 to 0.95, with pairs ranging from unrelated (e.g., sap → laugh) to strongly related (e.g., blue → red). Please note that GloVe is an entirely different relatedness metric and is not a linear transformation of AS (see Fig 1-Supp 1C for how the two stimulus sets compare using the common GloVe metric).

We tested memory for these stimuli using 5-min and 48-hr delays before the final test (Figure 2, bottom). In the 5-min delay experiment, the pattern changed. Base pair memory significantly differed across conditions [$F(4,796) = 8.50, p < 0.001$], and there was PI in the Δ Target condition,

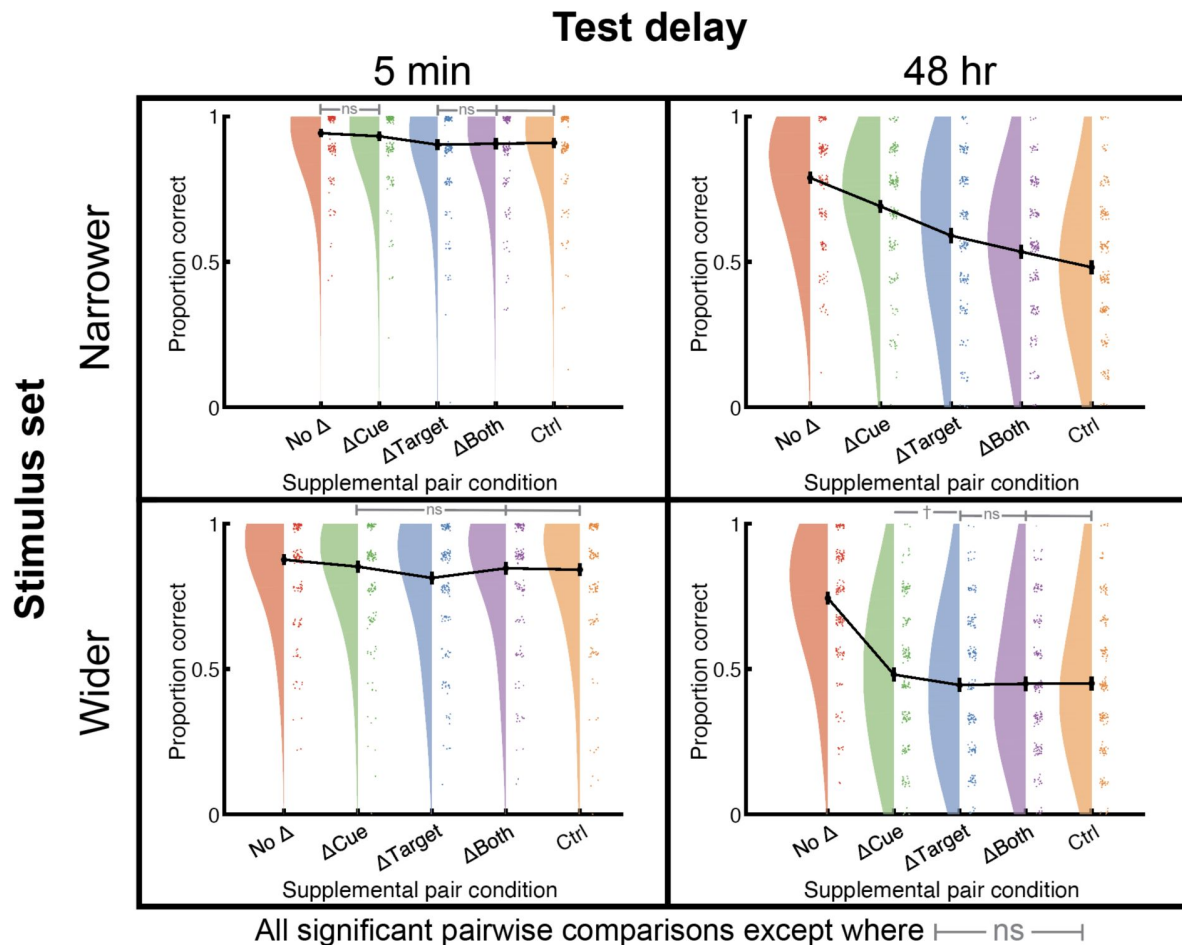


Figure 2

Stimulus set, test delay, and experimental condition all determined whether PF or PI occurred. The narrower stimulus set included only single-step semantic relations between base and supplemental cues and targets (top row), whereas the wider stimulus set included the full range of semantic relationships found in the English language (bottom row). Delays were either 5 min (left column) or 48 hrs (right column). All comparisons were significant except those labeled with connected gray bars and either 'ns' ($p > 0.1$) or † ($0.05 < p < 0.1$), and sets of connected bars indicated all 'ns' differences. Data points from individual subjects were jittered for improved visualization. See [Fig 2-Supp 1](#) for supplemental pair memory.

such that $No \Delta > \Delta Cue = Control = \Delta Both > \Delta Target$ (ΔCue vs. Control, adjusted $p = 0.41$; Control vs $\Delta Both$, $p = 0.63$; all others, $p < 0.05$) (Fig 2, lower left). However, in the 48-hr delay experiment, base pair memory differed strongly across conditions [$F(4,796) = 125.2$, $p < 0.001$], with pairwise t -tests indicating that $No \Delta > \Delta Cue \geq \Delta Target = \Delta Both = Control$ (all significant adjusted $p < 0.001$; ' $\geq$ ' indicates $0.1 > p > 0.05$ for ΔCue against all three later conditions; '=' indicates $p > 0.8$). Therefore, we found PI effects in the $\Delta Target$ condition, but this was only the case with lower average relatedness and a short retention interval. Please see Fig 2-Supp 1 for results from supplemental pair testing in all experiments.

Target relatedness produced proactive facilitation and scaffolded new target learning

The preceding effects are partially consistent with Osgood's predictions (Osgood, 1949) and our prior paper on retroactive effects using the same stimuli (Antony et al., 2022). That is, under conditions of strong overall relatedness, we found long-term PF in each experimental condition, with greater facilitation in the $\Delta Cue > \Delta Target > \Delta Both$ conditions, whereas shorter delays and lower overall relatedness levels resulted in the facilitation being reduced or eliminated (though note that ceiling effects were a concern with shorter delays). Moreover, we found classic PI in the $\Delta Target$ condition when there was both a short delay and lower overall relatedness. This result strikingly resembles the pattern in our prior paper, in which the same combination of experimental condition, timing, and stimulus set was the only one to produce classic retroactive interference (RI) effects (Antony et al., 2022).

Yet, a more detailed way to investigate Osgood's predictions involves measuring PI vs. PF as a function of semantic relatedness. We began these analyses by looking at memorability in the $\Delta Target$ condition. To calculate memorability, we subtracted the proportion of subjects recalling each word pair in the $\Delta Target$ condition minus the Control condition. This produced a memorability value for each pair that indicates PF when positive and PI when negative. We then regressed memorability against the semantic relatedness of the pair, using either AS or GloVe values.

Interestingly, semantic relatedness between targets produced greater memorability only in the wider stimulus set, 48-hr experiment [$r^2(44) = 0.36$, $p < 0.001$; other experiments had non-significant relationships, all $r^2(44) < 0.02$, $p > 0.20$] (Fig 3A). A careful examination of the regression line also shows that, whereas the $\Delta Target$ condition showed no overall PI or PF relative to the Control condition, we found evidence for *both* PI and PF at different levels of semantic relatedness. For pairs with low relatedness, we found a significantly negative y-intercept ($p < 0.001$), indicating PI for these pairs, but for pairs with high relatedness, this trend reversed to PF. This was verified by splitting pair relatedness values into thirds and running one-sample t -tests for all words in each third against 0. The lowest third showed PI [$t(14) = -3.6$, $p = 0.003$], whereas the upper third showed PF [$t(14) = 4.4$, $p < 0.001$].

The lack of effect in the 5-min experiments could be due to ceiling effects, as performance was high in both experiments. However, the lack of effect in the 48-hr, narrow stimulus set was surprising, as the retroactive equivalent was significant in our prior study (Antony et al., 2022). We considered that one important difference between the studies was the directional nature of the AS values (Sahakyan & Goodmon, 2007). In our prior study, we reasoned that we should use AS from supplemental $\rightarrow$ base pairs, such that the supplemental pairs would retroactively act upon the main memory of interest. However, given that the benefits could arise due to reminders, one could reason that this directional process is more important for a later-learned $\rightarrow$ earlier-learned representation than the converse. For this reason, we also investigated the relationship using base $\rightarrow$ supplemental pair AS. This relationship was also not significant [$r^2(44) = 0.007$, $p = 0.41$]. We will return to this surprising result in the Discussion.

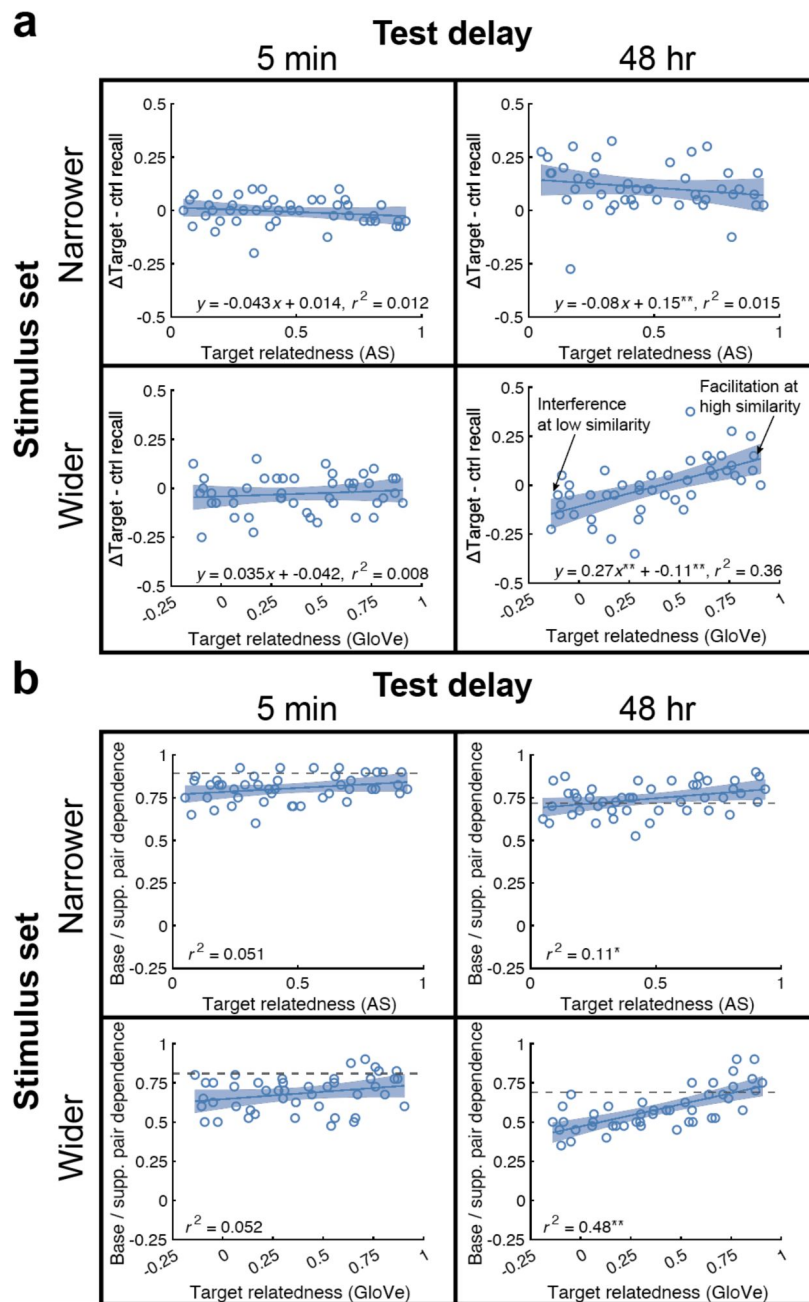


Figure 3

Memory and interdependence in the ΔTarget condition as a function of relatedness. (a) We plotted each pair's cross-subject memorability value against its corresponding target- Δtarget semantic relatedness in all experiments. AS and GloVe values were used in the top and bottom rows, respectively. In the wider relatedness stimulus set, 48-hr experiment, semantic relatedness proactively improved memory. Furthermore, PI occurred under conditions of low relatedness and PF under high relatedness. (b) We plotted memory dependence for each base pair-supplemental pair target duo against its relatedness in all experiments. High dependence means that subjects tended to remember or forget both targets in the duo (relative to recalling one and not the other). Relatedness significantly increased dependence in both 48-hr experiments ($p < 0.05$) but this was marginally significant in the 5min experiments ($0.05 < p < 0.1$). Thick dotted lines show the threshold of dependence levels against all other pairs at the 95th percentile. See also **Fig 3-Supp 1** for intrusion data from this condition.

Under predictions from a reminders-based account of proactive effects of memory (Wahlheim & Jacoby, 2013), recalling a prior pair during new learning can boost memory for pairs by creating interdependence among elements from earlier and more recent pairs. We next investigated the interdependence between corresponding initial (cue-target) and later-learned (cue- Δ target) pairs. We operationalized interdependence as the frequency by which base pair target-supplemental pair target duos across subjects were either both correct or both incorrect across subjects (Horner et al., 2015). That is, interdependence means that recalling the cue-target pair increases the likelihood of recalling the cue- Δ target pair (relative to a threshold of matching up one pair against all other pairs). Next, we correlated interdependence against relatedness. We found that interdependence significantly increased with relatedness in the 48hr experiments [both, $r^2(44) > 0.11$, $p < 0.02$], and this was marginally significant in the 5-min experiments [both $r^2(44) < 0.06$, $0.07 < p < 0.08$] (Fig 3B). Overall, our investigations of longer delays under a wider range in relatedness values showed that target relatedness correlated with base pair memorability and interdependence between base and supplemental pairs. Under other conditions, the relationship was weaker or non-significant, though the relationship between relatedness and interdependence was more reliable than memorability.

Additionally, we investigated whether target relatedness increased intrusions, operationalized here as entering the supplemental pair list response during the base pair test. Importantly, this helped us contrast between predictions from a PI-based account of our results and predictions from a successful reminders account of proactive memory effects. Under the PI-based account, targets could merge into a unified representation that does not allow the subject to differentiate the contexts [e.g., peace-shaverazor (written jointly to emphasize the unification)] (McCrystal, 1970; Postman & Keppel, 1977), as some studies have shown greater intrusion errors with increasing relatedness among stimuli (Dallett, 1962, 1964; Osgood, 1946; Underwood, 1951). This account is not necessarily contradicted by the preceding results, as the merged representation could (on average) help maintain the long-term memory while also leading to occasional source confusions. Conversely, the reminders-based account predicts that successful retrieval of existing memories during new learning scaffolds old and new information (leading to interdependence) while allowing subjects to remember the different contextual associations. We adjudicated between these accounts by correlating across-subject supplemental-into-base pair intrusions against relatedness. Critically, in the wider stimulus set, 48-hr experiment, intrusions significantly decreased with relatedness [$r^2(44) = 0.43$, $p < 0.001$], whereas they did not significantly increase in any experiment [all others, $r^2(44) < 0.02$, $p > 0.26$]. This finding, in combination with highly similar non-significant increases or significant decreases in intrusions in our retroactive study (Antony et al., 2022), supports predictions from the reminders account (Wahlheim et al., 2021) that reminders allow one to build interdependent, yet contextually discriminable representations.

Cue relatedness significantly increased base pair memorability and interdependence

In the above analyses comparing memory in the Δ Cue versus Control conditions, we found PF in all experiments. Next, we asked whether the extent of these PF effects increased with increasing relatedness among cues linked with the same target. We found that relatedness increased memorability in both 48-hr experiments [both $r^2(44) > 0.09$, $p < 0.02$] but neither 5-min experiment [both $r^2(44) < 0.02$, $p > 0.44$] (Fig 4A). Interestingly, in the wider stimulus set, 48-hr experiment, we found a similar effect as in the Δ Target condition analyses, in that there was PI at low relatedness and PF at high relatedness. Splitting relatedness values into thirds, the lowest third demonstrated significant PI [$t(14) = 3.0$, $p = 0.009$], whereas the upper third demonstrated PF [$t(14) = 3.5$, $p = 0.003$]. These results differ somewhat from our prior paper (Antony et al., 2022), which showed similar facilitation in the Δ Cue condition, but in which most correlations between memorability and relatedness were positive yet fell short of significance. Here, these relationships were significant in the 48-hr experiments. Another difference is that, whereas we did not find any

reliable interference in the Δ Cue condition in the prior study, there was significant PI in the low relatedness items in the wider stimulus set, 48-hr experiment. The y-intercept in the identical condition of the prior experiment was significantly above, rather than below, zero.

Additionally, we measured interdependence between target memory in the base and supplemental pair conditions. Interdependence increased in all experiments [all $r^2(44) > 0.08$, $p < 0.04$].

Cue + target relatedness increased memory accuracy and interdependence in the Δ Both condition

In the preceding analyses, we found PF in the Δ Both condition in the narrower stimulus set, 48-hr delay experiment but not in the other experiments. Next, we investigated whether memorability and/or dependence increased as a function of cue + target relatedness. Unlike in our prior retroactive experiment, we found no effects of relatedness on memorability [all $r^2(44) < 0.03$, $p > 0.48$], even though overall memory was better in the Δ Both condition than the Control condition in the narrower stimulus set, 48-hr delay experiment. Conversely, we found that relatedness improved interdependence in all experiments [all $r^2(44) > 0.07$, $p < 0.04$] except the wider stimulus set, 5-min delay experiment [$r^2(44) = 0.02$, $p > 0.85$]. Therefore, although there was no evidence that relatedness strengthened base pair memory, it generally promoted interdependence between base and supplemental pair memory. Overall, our results somewhat resemble the retroactive effects in Antony et al. (2022) [in that both papers showed significant overall facilitative effects and correlations between cue + target relatedness interdependence in the 48-hr, narrower stimulus set experiment](#). However, in the prior study, correlations between cue + target relatedness with memorability were also significant. Note that our prior paper featured a smoothed, 2-D plot with cue and target relatedness as separate predictors and memorability or dependence as the outcome variable (the cue + target relatedness plot was significant but was provided as a supplement). We have provided the smoothed, 2-D plot as **Fig 5, Supp 1** [and as part of the next subsection of the Results](#).

Osgood-style surfaces for proactive memorability and interdependence

One motivating factor for designing this study with such large sample sizes, five conditions, and continuous relatedness ratings was to re-create Osgood's predicted surface in one consolidated figure. We therefore re-plot results from all experimental conditions against the Control condition in **Figure 6** [In these plots, the strength of any pair lies along the x-y plane, which reflects the target and cue relatedness of its corresponding supplemental pair, respectively](#). If there was no preceding supplemental pair (Control condition), the pair remains on the x-y plane (orange). If the base pair is a repeat of the supplemental pair (No Δ condition), it belongs to the target identity, cue identity point (red). If only targets change (Δ Target condition), the pair lies along the cue identity line according to its target relatedness (blue), whereas if only cues change (Δ Cue condition), the pair lies along the target identity line according to its cue relatedness. Lastly, if both cues and targets change (Δ Both condition), the pair belongs on a 2-D surface as a function of both cue and target relatedness (purple).

Using similar formatting, we also plotted memory interdependence between base and supplemental pairs. Generally, relatedness increased with dependence along multiple dimensions in a stronger manner than its effects on proactive memory strengthening. These surfaces attempt to characterize a large range of factors that may proactively impact later memory formation and determine when earlier and more recent memories become linked.

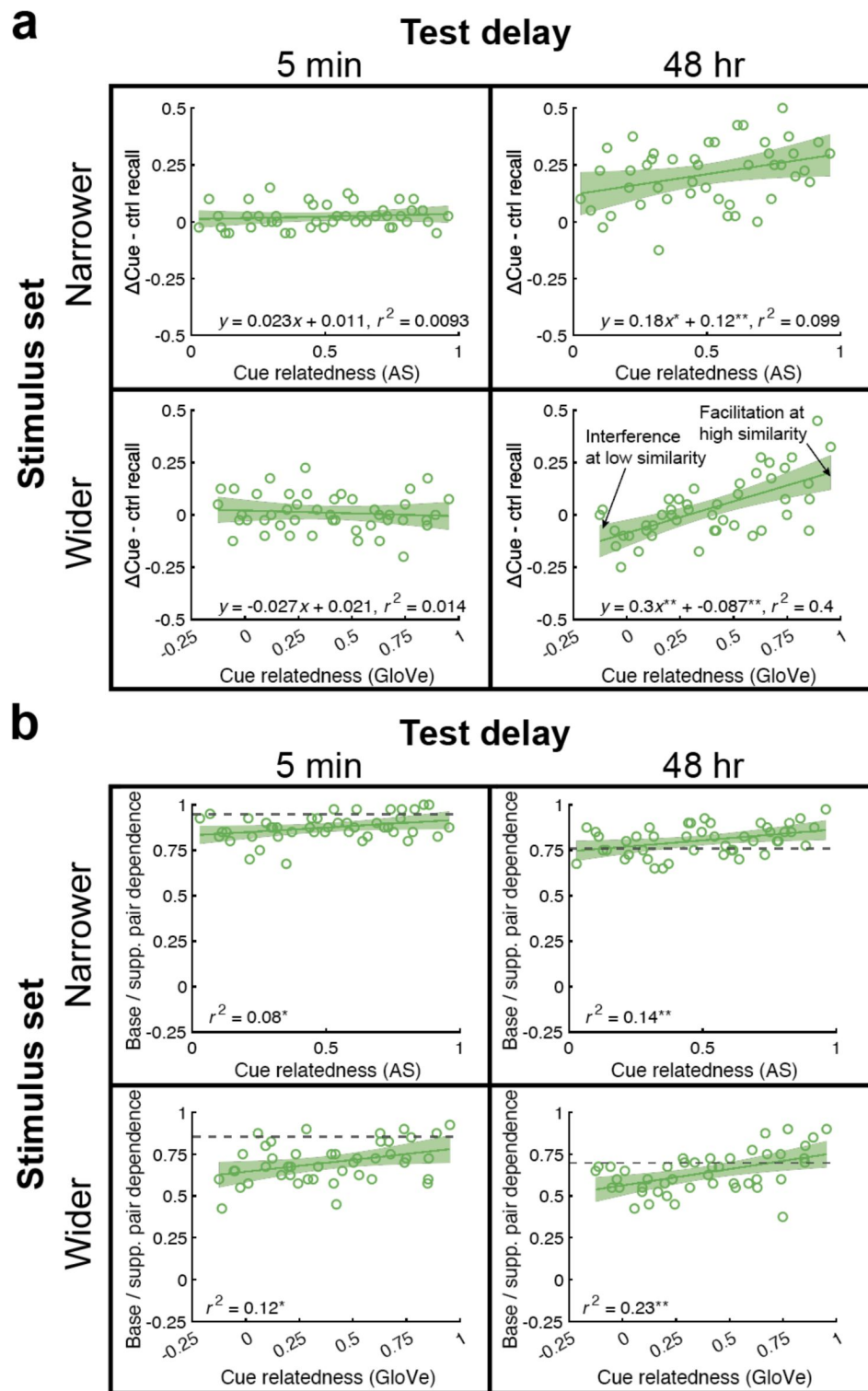


Figure 4

Semantic relatedness between cues benefits memory in the 48-hr experiments and boosts interdependence in all experiments. (a) Across-subject memorability for each pair in the ΔCue minus Control condition benefited from cue relatedness in the 48-hr, but not 5-min, experiments [top: AS; bottom: GloVe]. (b) Relatedness increased memory interdependence in all experiments. Thick dotted lines show the threshold of dependence levels against all other pairs at the 95th percentile.

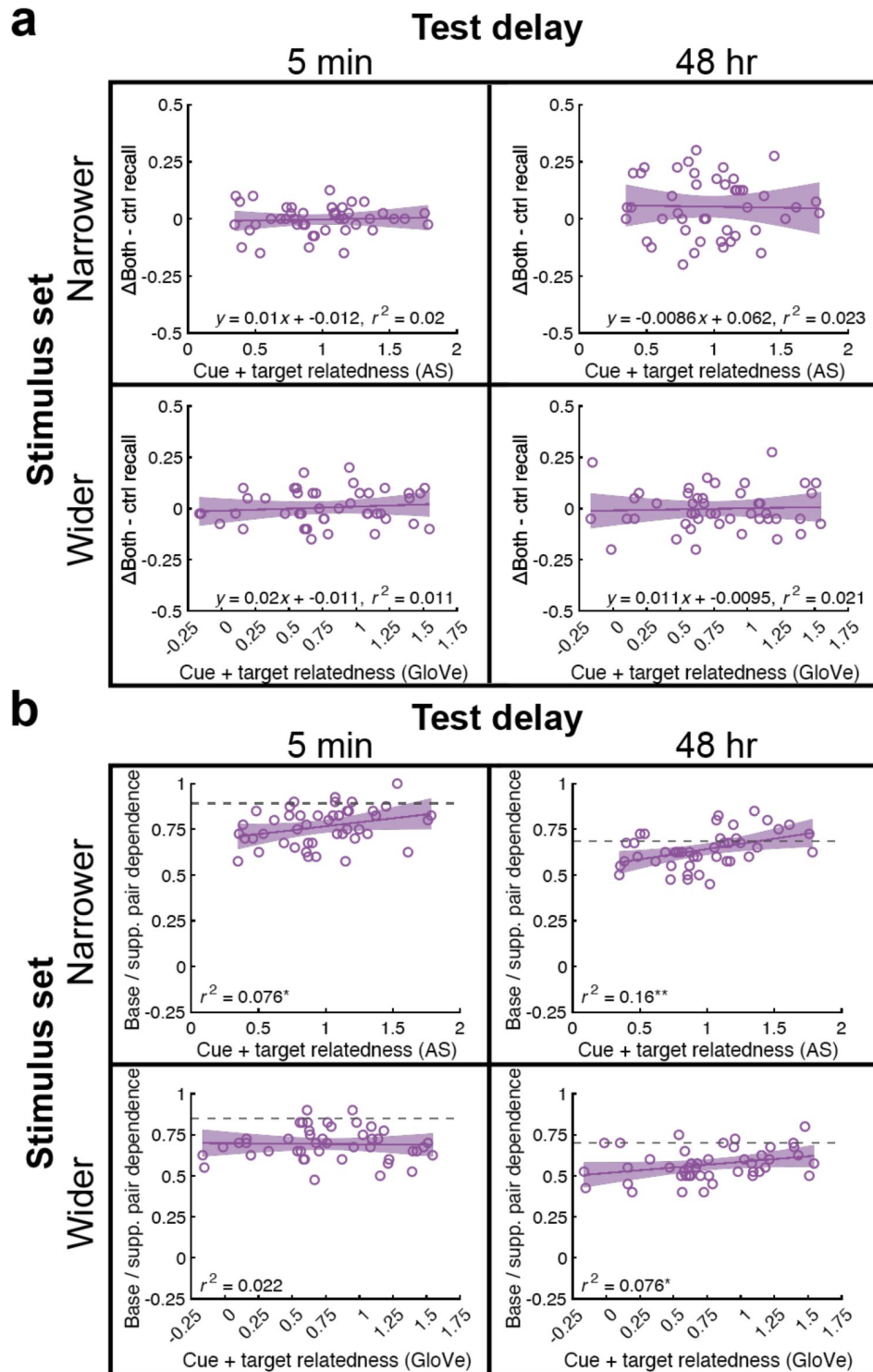


Figure 5

Increasing cue + target relatedness boosted interdependence but not memorability in the ΔBoth condition across experiments. (a) We plotted cue + target relatedness against memorability in each experiment. Memorability did not increase with relatedness in any experiment. (b) Memory interdependence increased with cue + target relatedness in all experiments except the wider stimulus set, 5-min delay experiment. See [Fig 5-Supp 1](#) for smoothed, 2-D plots that plot cue and target relatedness as separate variables predicting memorability and interdependence.

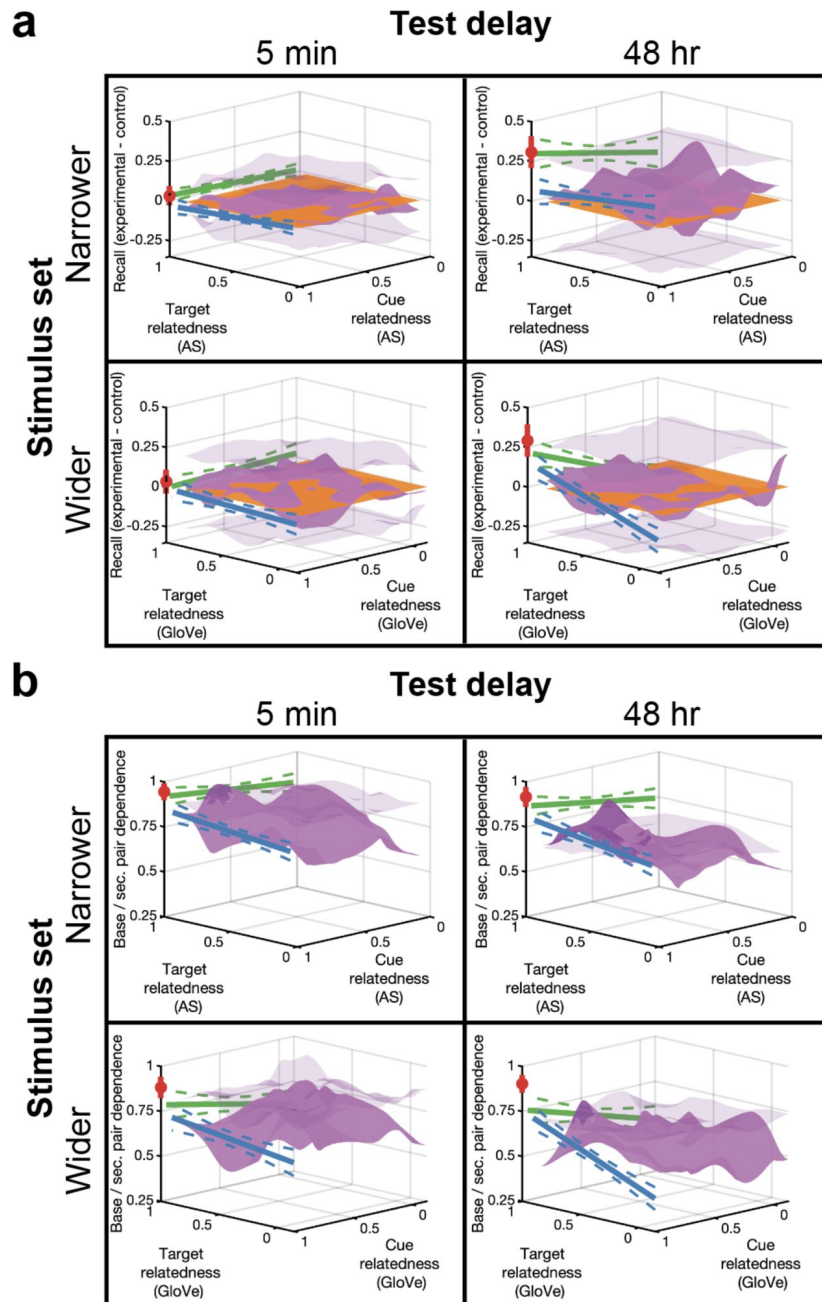


Figure 6

Osgood-style surfaces for proactive effects and interdependence. (a) We plotted memorability from all conditions minus the Control condition from all experiments in 3-D coordinates. These plots had cue and target relatedness on the y- and x-axes, respectively. On the z-axis, we plotted base pair memorability, oriented as PF for positive change and PI for negative change. Memorability for the No Δ ($\pm$ across-pair standard deviation) is at the cue identity, target identity corner point (red circle). We plotted Δ Target condition memorability along the cue identity line against target relatedness (blue) and Δ Cue condition memorability along the target identity line against cue relatedness (green) (both using $\pm$ standard error from the ordinary-least-squares regression fit). We plotted Δ Both condition memorability as a locally smoothed surface as a function of both cue and target relatedness (purple). Transparent surface grids above and below zero represent $p < 0.01$ significance boundaries from permutation tests; significant points on the surface are indicated by a darker shade of purple. (b) Interdependence measurements for all experiments and conditions are formatted similarly to (a).

Re-assessing proactive memory and interdependence effects using a common metric

The above stimulus sets intentionally spanned different ranges of relatedness, but we wanted to ensure our effects held when using a common metric. Therefore, we combined across-subject memorability and interdependence values across stimulus sets in each experimental condition, keeping the results separate for a given test delay. In other words, we combined the two 5-min delay experiments together and the 48hr delay experiments together. Then, we correlated memorability with GloVe relatedness values in a similar fashion to the above analyses (**Fig 6-Supp 1**). In the 5-min delay experiment, there were no significant correlations between memorability and relatedness in the three conditions [all $r^2(44) < 0.011$, $p > 0.2$]. Correlations between interdependence and relatedness were significant in the Δ Cue [$r^2(44) = 0.19$, $p < 0.001$] condition and Δ Target condition [$r^2(44) = 0.14$, $p < 0.001$], but not the Δ Both condition [$r^2(44) = 0.02$, $p = 0.12$]. In the 48-hr delay experiment, correlations between memorability and cue relatedness in the Δ Cue condition [$r^2(44) > 0.29$, $p < 0.001$] and target relatedness in the Δ Target condition [$r^2(44) = 0.2$, $p < 0.001$] were significant, whereas cue+target relatedness in the Δ Both condition was not [$r^2(44) = 0.01$, $p = 0.58$]. In all three conditions, interdependence increased with relatedness [all $r^2(44) > 0.16$, $p < 0.001$].

Semantic relatedness accelerated new learning

Various paradigms have shown that new learning is accelerated when it can be linked with prior information (Barnes & Underwood, 1959; Bartlett, 1932; Bein et al., 2015; Brod et al., 2013; Jarrett & Scheibe, 1963; Metcalfe et al., 1993; Morton et al., 2023; Palermo & Jenkins, 1964; Postman & Parker, 1970; Tse et al., 2007; Underwood, 1951; Van Kesteren et al., 2012; Wimer, 1964; Witherby & Carpenter, 2022; Young, 1955). Indeed, we found this to be the case in our prior experiment, such that relatedness within each condition predicted fewer rounds to achieve learning criterion for later-learned pairs (in that case, supplemental pairs, which were learned after base pairs) (Antony et al., 2022). Here, we had the opportunity to perform the same analyses but also compare results to a later-learned Control condition. Using average number of trials to criterion during base pair learning as an outcome measurement, we evaluated overall condition effects as well as correlations with relatedness. In the narrower stimulus set experiments, the number of trials to criterion followed a $No \Delta < \Delta$ Cue $< \Delta$ Target $< \Delta$ Both $< Control$ condition pattern (5-min delay: all comparisons, adjusted $p < 0.008$; 48-hr delay: all comparisons, adjusted $p < 0.007$) (**Fig 7A**). The wider stimulus set experiments produced a somewhat similar pattern of $No \Delta < \Delta$ Cue $\leq \Delta$ Target $< \Delta$ Both = Control condition pattern (5-min delay: Δ Both vs. Control, adjusted $p = 0.92$; all other comparisons, adjusted $p < 0.03$; 48-hr delay: all comparisons except Δ Both vs. Control and Δ Cue vs. Δ Target, adjusted $p < 0.01$; Δ Both vs. Control, $p = 0.19$; Δ Cue vs. Δ Target, $p = 0.19$). These results resemble the overall memory results in a condition-by-condition fashion, such that faster learning within a condition generally led to better memory after a 48-hr delay. Next, we correlated learnability (average base pair trials to criterion across subjects) as a function of relatedness in the Δ Cue, Δ Target, and Δ Both conditions. We found that higher cue relatedness produced faster base pair learning in every experiment, either significantly or marginally significantly [narrower stimulus set, 48-hr delay experiment: $r^2(44) = 0.064$, $p = 0.052$; all others, $r^2(44) > 0.065$, $p < 0.05$] (**Fig 7B**). Similarly, higher target relatedness produced faster base pair learning in every experiment, either significantly or marginally significantly [narrower stimulus set, 5-min experiment: $r^2(44) = 0.04$, $p = 0.09$; all others, $r^2(44) > 0.11$, $p < 0.02$] (**Fig 7C**). Finally, additive cue + target relatedness generally produced faster base pair learning in the narrower stimulus set experiments [5-min delay experiment: $r^2(44) = 0.08$, $p = 0.04$; 48-hr delay experiment: $r^2(44) = 0.046$, $p = 0.08$] but not the wider stimulus set experiments [5-min delay experiment: $r^2(44) = 0.0045$, $p = 0.37$; 48-hr delay experiment: $r^2(44) = 0.024$, $p = 0.16$] (**Fig 7D**). Overall, it appears that previously learned supplemental pairs scaffold and hasten learning of base pairs depending on their level of relatedness, with most relationships either significant or marginally significant.

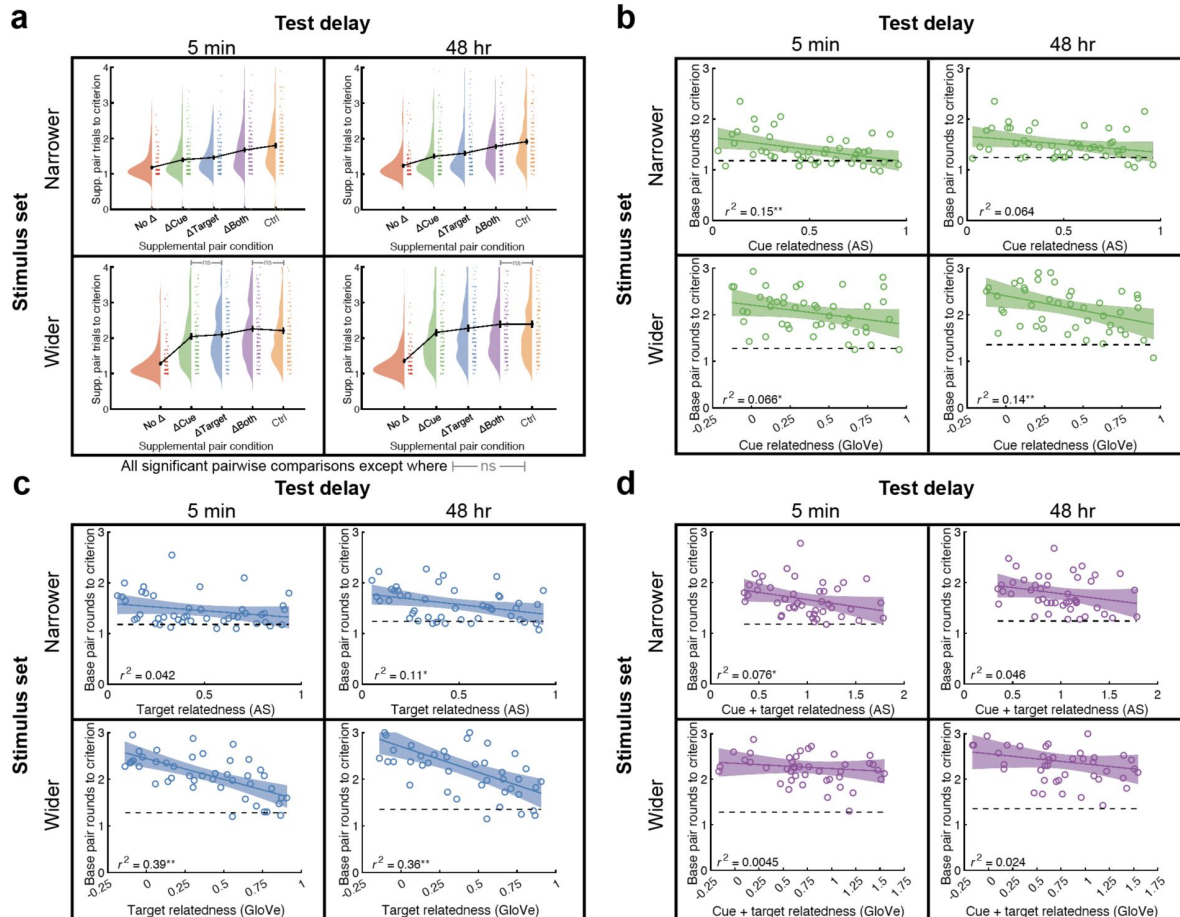


Figure 7

Base pair learning differed based on stimulus set and condition and generally benefited from higher semantic relatedness. (a) In the narrower stimulus set, learning time (mean trials to criterion) followed this trend: No $\Delta < \Delta$ Cue $< \Delta$ Target $< \Delta$ Both $<$ Control. The wider stimulus set tended to show nearly the same pattern: No $\Delta < \Delta$ Cue $\leq \Delta$ Target $< \Delta$ Both = Control, where ' $\leq$ ' indicates a significant difference in one experiment but not the other. All comparisons were significant except those labeled with gray bars and 'ns' ($p > 0.1$). Data points from individual subjects were jittered slightly for visualization, and a few outliers (slower learners) were omitted from the plot for better visualization. (b-d) We correlated relatedness with learning times [top: AS; bottom: GloVe]. (b-c) Learning time across subjects for each word pair generally decreased with increasing cue relatedness in the Δ Cue condition (b) and decreased with increasing target relatedness in the Δ Target stimulus condition (c). (d) In the Δ Both condition, learning time generally decreased with cue + target relatedness in the narrower stimulus set, but not in the wider stimulus set. In (b-d), Pearson correlations are shown in the plots followed by * when $p < 0.05$ and ** when $p < 0.01$.

Comparing retroactive and proactive memory effects

The preceding results accord well with the qualitative results from our prior study examining retroactive effects (Antony et al., 2022 [↗](#)) (with a few exceptions, which we cover more extensively in the Discussion). One conclusion from these converging results could be that relatedness benefits memories in the retroactive and proactive case for similar reasons, namely that by virtue of increasing the likelihood of looking back and associating new learning with existing memories, it helps the connected memories survive due to being embedded in an integrated memory structure (Antony et al., 2024 [↗](#); Antony & Schechtman, 2023 [↗](#); Lee & Chen, 2022 [↗](#); Schlichting et al., 2015 [↗](#); Trabasso & Sperry, 1985 [↗](#)). Our results here and earlier, as well as those from other recent studies showing that memory updating benefits from variables that improve contact between the past and present (e.g., Kemp et al., 2023 [↗](#); Wahlheim et al., 2020 [↗](#), 2023 [↗](#); Wahlheim & Zacks, 2019 [↗](#)), support this idea.

Nonetheless, one could still conclude that there should be asymmetric benefits: reminders are retrieval events, and retrieval practice profoundly affects memory (Antony et al., 2017 [↗](#); Gates, 1917 [↗](#); Roediger & Karpicke, 2006 [↗](#); Rowland, 2014 [↗](#)). Therefore, existing memories may receive the additional benefits of an extra retrieval, while new learning does not receive that benefit. In this final section, we will adjudicate between these predictions by comparing overall base pair memory from the proactive conditions here against the comparable retroactive conditions from Antony et al. (2022) [↗](#). We highlight the caveat that we conducted these experiments at different times. Subjects were therefore not randomly assigned to proactive and retroactive effects conditions. Nevertheless, both studies used the same subject population, learning materials, learning criteria, and timing across experiments, and we will additionally use internal control measurements to make these comparisons as equal as possible.

To compare proactive and retroactive effects directly, we sought a better measurement than raw memory performance in each condition that would allow us to account for within-participant performance. As opposed to results in the preceding sections, we believe here that the best control is the positive control (No Δ condition), because, given that list-based interference has been reported in some studies using A-B, C-D, test C-D designs (Postman & Keppel, 1977 [↗](#)), it is possible that list-based interference still affected new learning in the (negative) Control condition to a small degree. Therefore, we examined memory in all conditions minus the (always superior) No Δ condition, wherein less negative values indicated better memory. We also eschewed the 5-min delay experiments, because in those cases, there is a considerable difference between the timing since base list learning. That is, it was only 5 min in the proactive case, whereas it was 5 min plus the time it took to learn supplemental pairs in the retroactive case; given this difference, one would expect memory differences. However, we include those comparisons in **Fig 8, Supp 2** [↗](#) for those interested.

In comparing retroactive and proactive designs in the narrower stimulus set, 48-hr experiments, we found better memory in the retroactive experiment in the Δ Cue, Δ Target, and Δ Both conditions (all adjusted $p < 0.011$) and not the (negative) Control condition ($p = 0.43$) (**Fig 8** [↗](#), top left). In the wider stimulus set, 48-hr experiments, there were retroactive advantages in the Δ Cue and Δ Target conditions (both $p < 0.005$) but neither the Δ Both ($p = 0.31$) nor the Control conditions ($p = 0.94$). Therefore, all conditions showing long-term memory facilitation also showed greater facilitation in the retroactive than proactive case (Δ Cue, Δ Target, and Δ Both in the narrower stimulus set; Δ Cue and Δ Target in the wider stimulus set) (**Fig 8** [↗](#), top right).

Finally, we wanted to run one more control experiment to rule out a potential difference between the retroactive and proactive designs that the preceding comparisons could not address. Specifically, one design involves subjects first recalling items from a first list and the other a second list. It remains possible that recall is slightly more difficult when mentally flipping around

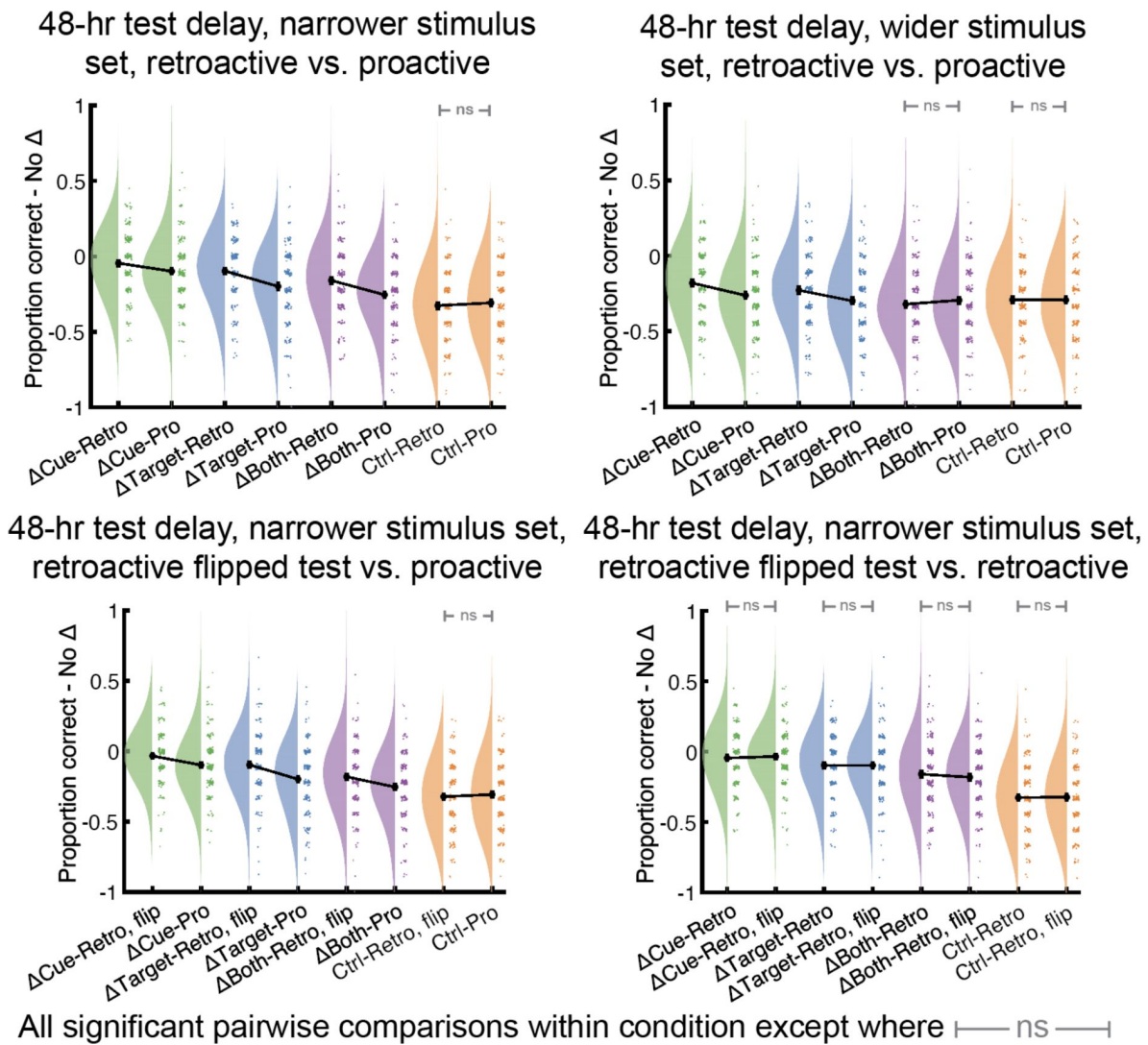


Figure 8

Under otherwise identical conditions, memory benefits after 48 hours were stronger for experimental groups under retroactive versus proactive designs when compared to the No Δ condition, which served as a positive control. Comparisons of the same base pair condition across experiments were performed between the narrower stimulus set, retroactive versus proactive experiments (top left); the wider stimulus set, retroactive versus proactive experiments (top right); the narrower stimulus set, flipped test experiment (in which base pairs were subject to retroactive effects) versus proactive experiment (bottom left); and among the narrower stimulus set retroactive experiment versus flipped experiment (bottom right). In all cases, experimental groups that showed facilitation relative to the Control condition in the overall analyses showed greater retroactive than proactive facilitation.

everything one learned after a certain delay. To address this concern, we ran another experiment ($N = 200$) with the following order: base pair learning, supplemental pair learning, 48-hr delay, supplemental pair testing, base pair testing. We refer to this as the flipped test experiment and it used the narrower stimulus set. In this experiment, we calculated supplemental pair memory in each condition minus the No Δ condition, and we compared it in the flipped test experiment (when tested first after the delay) to the other narrower stimulus set, 48-hr, retroactive experiment (when it was tested second after the delay). Contrary to the idea that probing memory out of order impairs it, we found non-significantly better memory in the flipped test version in the Δ Target and Δ Both (both $0.12 < p < 0.13$) conditions and no differences in the Δ Cue condition ($p = 0.94$).

This experiment also served two other purposes. First, because this flipped test experiment was conducted online, it allowed us to control for experimental setting when comparing the narrower stimulus set, 48-hr experiments, because the original retroactive, narrower stimulus set, 48-hr experiment took place in person (before the Covid-19 pandemic). Second, it replicated the main findings in our prior study (see **Fig 8-Supp 1** [↗](#) for average results from each condition in this flipped test experiment) and allowed us to replicate retroactive versus proactive comparisons described above. We therefore contrasted base pair memory (subtracting No Δ condition memory) in this flipped test experiment against base pair memory in the proactive experiment. We again found better memory in the retroactive experiment in the Δ Cue, Δ Target, and Δ Both conditions (all adjusted $p < 0.002$) and not the (negative) Control condition ($p = 0.52$) (**Fig 8** [↗](#), bottom left). Finally, we directly compared base pair memory in the retroactive, narrower stimulus set, 48-hr delay experiment from our prior study against base pair memory in the flipped test experiment. We found no significant differences in any of the Δ Cue, Δ Target, Δ Both, or Control conditions (all adjusted $p > 0.87$) (**Fig 8** [↗](#), bottom right).

Discussion

We found that, as old and new information became more semantically related, memory for new information increased and the two grew more interdependent. When considering overall memory across experimental conditions, an overall high level of relatedness between old and new information (as in the narrower stimulus set) resulted in profound long-term PF for the Δ Cue, Δ Target, and Δ Both conditions relative to the Control condition. When the overall level of relatedness was lower (in the wider stimulus set), we found classic PI effects in the Δ Target condition in the 5-min delay experiment. Additionally, in the 48hr delay experiment, the overall memory benefits were smaller (Δ Cue condition) or eliminated (Δ Target condition), but we found an intriguing pattern by which PI occurred with low relatedness and PF with high relatedness in both the Δ Cue and Δ Target conditions. We did not find relatedness correlations with memorability in the Δ Both condition of any experiment. Finally, we found benefits of relatedness on interdependence in all experimental conditions of all 48-hr delay experiments but weaker or less reliable effects in 5-minute delay experiments.

These results converged strongly with those from our prior study (Antony et al., 2022 [↗](#)) focusing on retroactive memory effects. In both studies, we found overall benefits of each experimental condition in the narrower stimulus set, 48-hr delay experiment. In both studies, we found correlations between relatedness and memorability in the Δ Cue and Δ Target conditions in the wider stimulus set, 48-hr delay experiment, though this was less reliable in the Δ Cue condition of the prior study. In fact, in the prior study, you could directly see how delay and relatedness combined to determine retroactive facilitation versus interference. In the Δ Target condition of the wider stimulus set experiments, retroactive effects went from interference with low relatedness to no interference with high relatedness in the 5-min experiment to no interference with low relatedness to facilitation with high relatedness in the 48-hr experiment. Here, we found this crossover from PI with low relatedness and PF with high relatedness in the same experiment. Therefore, in both studies, the only condition showing overall retroactive interference (RI) and PI

was in the wider stimulus set, 5-min delay condition, supporting the idea that RI and PI are relatively limited to these cases. Thus, the tendency to study memory with short delays and without interrelationships among study material led to the predominance of this idea in interference studies and theories throughout the 20th century. Additionally, in both studies, we also found correlations between relatedness and interdependence in all experimental conditions in the narrower stimulus set, 48-hr delay experiment and the Δ Cue and Δ Target conditions of the wider stimulus set, 48-hr delay experiments. This finding also supports recent studies showing that interdependence increases with contextual similarity between episodes (Cox et al., 2021 [↗](#); Yu et al., 2023 [↗](#)), which may also be more likely to become integrated during encoding.

There were also some differences between the studies. Here, we found an unexpected null relationship between relatedness and memorability in the Δ Target condition of the 48-hr, narrower stimulus set condition that was significant in the prior experiment. Notably, this was not simply a function of the direction of the associative strength (AS) measure, because flipping around the direction of AS from supplemental target $\rightarrow$ base target to base target $\rightarrow$ supplemental target also produced non-significant results. We note that there's an overall Δ Target benefit in this condition, and the y-intercept of **Fig 3A** [↗](#), top right, is significantly above zero, reflecting memorability benefits even with 0 AS. However, the precise level of relatedness was not predictive of memorability benefits. We also found a weaker effect in the Δ Both condition of the narrower stimulus set, 48-hr delay experiment, such that additive cue+target relatedness did not produce significant memorability benefits (though it still predicted interdependence). Critically, we also found overall differences when directly comparing memory in the otherwise identical retroactive and proactive designs. These comparisons consistently resulted in larger retroactive than proactive memory benefits in each of the experiments showing overall memory facilitation (Δ Cue, Δ Target, and Δ Both in the narrower stimulus set; Δ Cue and Δ Target in the wider stimulus set). With the caveat that these comparisons do not use random assignment, these highly consistent results support the idea that, while reminders produce benefits for both old and new information, they disproportionately benefit old information, likely due to the additional retrieval trial (Carpenter et al., 2008 [↗](#); Gates, 1917 [↗](#); Roediger & Karpicke, 2006 [↗](#); Rowland, 2014 [↗](#)).

We will now offer a cumulative explanation for our results, starting with key points about the recursive reminders hypothesis (Hintzman, 2011 [↗](#)). A series of studies using paired associate learning tasks akin to those here have shown that encouraging subjects to recall existing memories during new learning can reduce PI and promote PF, depending on how well the information is integrated and later recollected (for a review, see Wahlheim et al., 2021 [↗](#)). In these cases, subjects must both notice and subsequently recollect a change between prior and new information to gain these benefits. Similar findings have been observed when using measures of recollecting change at test to infer noticed changes that were not measured during new learning (Wahlheim, 2015 [↗](#)). Additionally, reminders promote interdependence between the old and new information while preserving memory for the specific list contexts (Jacoby et al., 2015 [↗](#); Wahlheim & Jacoby, 2013 [↗](#)). To link our results to this reminders-based account of proactive and retroactive effects of memory, referred to as the memory-for-change framework, we propose that pre-existing relationships between base and supplemental pairs increased the likelihood of reminders, which resulted in retrieval practice benefits to more existing memories and provided more opportunities to integrate them with changed associations during new learning. This account could be verified in future studies including manipulations of semantic relatedness along with overt measures of noticing and recollecting changes.

We next consider these results in more detail for each condition. In the Δ Target condition, we found overall 48-hr benefits (relative to the Control condition) but no correlation with target relatedness using the narrower stimulus set. Conversely, we found no overall longterm benefits using the wider stimulus set, but we did find correlations between target relatedness and memorability. We note the surprising result of a non-significant correlation between relatedness and memorability using the narrower stimulus set, but the analyses combining narrower and

wider stimulus sets at a 48-hr delay produced a positive correlation. No PF occurred in the 5-min experiments, but we found overall PI using the wider stimulus set (**Fig 6-Supp 1**). Altogether, these results show that PI can still occur with low relatedness, like in other studies finding PI in Δ Target (A-B, A-D) paradigms (for a review, see Anderson & Neely, 1996), but PF occurs with higher relatedness. In fact, the absence of low relatedness pairs in the narrower stimulus set likely led to the strong overall PF in this condition across all pairs (positive y-intercept in the upper right of **Fig 3A**). In this instance, there may have been a stronger influence of a binary factor (whether they are related or not), though this remains speculative and is not the case for other analyses in our paper. We believe these effects arise from the competing processes of impairments between competing responses at retrieval that have not been integrated versus retrieval benefits when that integration has occurred (which occurs especially often with high target relatedness). These types of competing processes appear operative in various associative learning paradigms such as retrieval-induced forgetting (Anderson & McCulloch, 1999; Carroll et al., 2007), and the fan effect (Moeser, 1979; Reder & Anderson, 1980). This is exactly the proposal of the memory-for-change framework, which has received support from studies showing that PF occurs when changes (integrated representations) are recollected and PI when such changes are not recollected (for a review, see Wahlheim et al., 2021).

In the Δ Cue condition, we generally found PF in every experiment. Moreover, PF increased with cue relatedness at a 48-hour delay using the wider stimulus set; the only PI in any experiment was for low relatedness items using the wider stimulus set at a 48-hour delay, wherein memorability values from pairs (averaged across subjects) in the lowest third of relatedness were recalled worse than the Control condition. These results are consistent with Osgood's observation that interference tends to be absent or low when targets remain constant and cues change as opposed to when targets change and cues remain constant (e.g., the Δ Target condition). The general facilitation and lack of interference effects are consistent with the idea that each independent cue does not become overloaded and can elicit the same (and only) target with which it was paired (Anderson & Neely, 1996; Antony et al., 2022; Watkins & Watkins, 1975). Additionally, repeating the target, even with another cue, may increase its accessibility (at minimum) or evoke a reminding, leading to overall benefits (Estes, 1979; Martin, 1965). The prior literature on results for this condition has not been particularly reliable or theoretically extensive; we speculate that the selective PI we see could arise via confusion over whether the same target was indeed assigned twice or due to a very weak form of interference involving weight changes after targets become assigned to a new memory. Benefits could indicate that subjects could mentally link Δ Cue $\rightarrow$ cue $\rightarrow$ target, which would link both cues together via the reminding event; the overall correlations between relatedness and interdependence support this idea. In sum, the Δ Cue condition largely results in more consistent PF than the Δ Target condition, perhaps because benefits can arise via the same mechanism (reminders-enabled interdependence) or even enhanced target response availability while not suffering the same drawbacks of competition that occurs when multiple non-integrated targets are linked to the same cue.

Next, we will discuss proactive effects in the Δ Both condition, for which the prior literature is virtually non-existent. In the narrower stimulus set, 48-hr delay experiment, we found overall PF effects. These findings resemble those from our prior retroactive study (Antony et al., 2022), such that this was the only experiment in which we found overall memory effects. However, unlike that prior study, here there was no correlation between cue+target relatedness and memorability. These findings also resemble general benefits of relatedness in other conditions and experiments, with the difference being that changes to both cues and targets mean that both relatedness values need to be relatively high for reminders (and their associated strengthening) to occur. That is, it is much easier for a subject to consider a new pair with both a cue related to an old cue and a target related to an old target to simply be a new word pair, causing fewer reminders overall. Interestingly, in three of the experiments, we found significant correlations between cue+target relatedness and interdependence. It is less clear why we found

interdependence in some experiments with an absence of memory boost; we speculate that in some cases, subjects may recognize the link between related pairs in the Δ Both condition only at final retrieval.

Putting these conditions together, we consider our findings to largely support Osgood's key ideas. That is, Osgood had the bold idea that the memory surface predicting facilitation vs. interference applied in both the retroactive and proactive directions. We propose that semantic relatedness, by increasing the likelihood of reminders, provides a mechanism by which existing memories can become more accessible and associated with new learning as part of integrated memory representations that preserve targets and their contexts. Specifically, along both the cue and target identity lines of Osgood's surface, the increasing proactive benefits strongly support his proposal that PI becomes PF with high enough relatedness. However, he did not predict PI in the Δ Cue condition, and additionally, we found weak support for PF inching towards the cue and target identity point along the 2-D part of the surface in the Δ Both condition [which was present in the retroactive narrower stimulus set, 48-hr experiment of our prior paper].

While Osgood predicted that retroactive and proactive memory effects worked in both directions, he never explicitly stated whether they should be symmetric. By showing stronger retroactive than proactive benefits here, we propose that the retroactive benefits are greater, presumably because target responses receive retrieval practice during reminders. Indeed, we have recently shown similar results in between-subjects comparisons of retroactive and proactive memory effects in A-B, A-D paradigms (Wahlheim, C.N. et al., submitted). However, our assertion awaits additional empirical verification using experiments that directly compare retroactive and proactive effects within-subjects and at various retention intervals. Additionally, these effects partially resemble those on spontaneous recovery, whereby original associations tend to face interference after new, conflicting learning, but slowly recover over time (either absolutely or relative to the new learning) and often eventually eclipse memory for the new information (Barnes & Underwood, 1959 [↗](#); Postman et al., 1969 [↗](#); Wheeler, 1995 [↗](#)). In both cases, original associations appear more robust to change over time, though it is unclear whether these similar outcomes stem from similar mechanisms.

Our experiment had a few limitations. First, although we showed highly consistent results across retroactive and proactive comparisons and used the same subject populations and procedure, these comparisons across experiments did not involve random assignment (i.e., all experiments assessing retroactive comparisons were completed before those assessing proactive comparisons). Second, we largely explained our results using reminders-based accounts, but we did not directly manipulate reminders or measure such reminders during new learning (cf. Jacoby et al., 2015 [↗](#)). These limitations could be considered in future research. Yet another promising future direction would be to investigate how learning in this paradigm affects different aspects of the memory trace, either by investigating perceptual versus conceptual changes (Lifanov et al., 2021 [↗](#); Melega & Sheldon, 2023 [↗](#)) or the semantic representations themselves (Walsh & Rissman, 2023 [↗](#)). Finally, future studies could ensure these effects are not limited to these stimuli and generalize to other word stimuli in addition to testing other domains (Baek & Papaj, 2024 [↗](#); Holding, 1976 [↗](#)).

Our goal in this paper was to examine fundamental questions about learning and memory for related pieces of information. Our prior work examined the retroactive impact of relatedness on existing memories (Antony et al., 2022 [↗](#)), whereas, here, we examined how relatedness proactively affected memory for new learning. Both studies were inspired by Osgood's proposal that relatedness facilitates memory strength in both directions – and both largely support his proposal. Here we offer support to the idea that, by increasing the likelihood of reminders, both old and new information can mutually benefit from the newly integrated memories, so long as they promote later recollection. Nevertheless, because retrieved memories enjoy the benefits of retrieval practice, retroactive effects appear to produce stronger benefits for memories recalled at

a 48-hr delay. Therefore, to return to the example about vermouth from the introduction, if one first learns to make a martini using dry vermouth and then, while learning about making Manhattans using sweet vermouth, they recall the martini ingredients, both memories should benefit from this mental integration, but the martini memory could benefit more. We believe these findings have important ramifications for how linking old and new associations can influence human decision-making (Nicholas et al., 2023 [↗](#); Shohamy & Daw, 2015 [↗](#); Wimmer & Shohamy, 2012 [↗](#)) and reinforce and clarify the significance of building up complex networks of associations for learning and memory theory.

Methods

Subjects.

We collected data online from 200 subjects in each experiment to amply power memorability measures from each condition. We used a 5-way counterbalance, indicating that there would be $200 / 5 = 40$ data points going into each memorability value. All subjects were undergraduate students who received psychology course credit for participating and self-attested to having normal or corrected-to-normal vision. We excluded subjects with memory performance < 4 standard deviations below the mean; we ran additional subjects until again reaching 200 without exclusions. Additionally, numerous subjects did not return or complete the final test. The final breakdowns were as follows: narrower stimulus set, 5-min delay: $N = 212$, 127 female, 3 non-binary; narrower stimulus set, 48-hr delay: $N = 205$, 159 female, 0 non-binary; wider stimulus set, 5-min delay: $N = 203$, 138 female, 3 non-binary; wider stimulus set, 48-hr delay: $N = 214$, 141 female, 6 non-binary; flipped-order experiment: $N = 223$, 141 female, 2 non-binary. All subjects were recruited via an online scheduling software or Google document set up for a psychology course. Subjects provided informed consent, and all procedures were in accordance with the California Polytechnic State University, San Luis Obispo Institutional Review Board. Note that in the prior paper examining retroactive memory effects (Antony et al., 2022 [↗](#)), all experiments were similarly run online except for the narrower stimulus set, 48-hr delay experiment, wherein subjects used lab computers (before the COVID-19 pandemic). Note also that in that prior study, what we refer to here as ‘supplemental pairs’ were called ‘secondary pairs’; we changed the language to avoid confusion, as these were learned first in this study. All subjects engaged in each step of the experiment with the guidance of research assistants; that is, all sessions were conducted synchronously online.

Stimuli & relatedness metrics.

We created our experiments with the memorability measure in mind to control for the actual 3–5 letter words recalled across individuals. Therefore, the main memories of interest (base pairs) involved the same 45 pairs for everyone within an experiment, meaning that we altered only the experimental condition of the supplemental pairs and the level of semantic relatedness between base and supplemental pairs across subjects. See Antony et al. (2022) [↗](#) for a detailed discussion of how stimuli were selected and allocated to counterbalance conditions across subjects. Note that for the narrower stimulus set, AS values are directed, which means that the strength between a given base $\rightarrow$ supplemental pair word was not equal to the same supplemental $\rightarrow$ base pair word. We designed the studies to systematically vary the AS from supplemental $\rightarrow$ base pair word, but in the prior study, this involved looking backward in time, whereas here, it involved looking forward in time. We bring attention to both directed AS values in the Results section to ensure this factor did not alone explain differences between the studies. For the wider stimulus set experiments, GloVe relationships were undirected, and therefore the base-supplemental pair decisions were arbitrary. All relatedness metrics were identical to those in Antony et al. (2022) [↗](#).

Procedure.

The four main experiments in this paper followed this order: supplemental pair learning, base pair learning, a 5-min or 48-hr delay, base pair testing, and supplemental pair testing (**Figure 1A**). For supplemental pair learning, subjects first viewed 36 pairs followed by multiple rounds of testing. During encoding, pairs were preceded by a one-second fixation cross before appearing for four seconds. Cue and target words were shown centrally in the horizontal dimension and just above and below the vertical center of the screen, respectively. During retrieval-to-criterion testing, a one-second fixation cross preceded the appearance of the cue word; one second after the cue word, a blank prompt appeared in which subjects were asked to type in the target response. Subjects had no time limits for their response, and both cue and target words were shown as feedback after the response. Correct pairs were dropped from later rounds of testing until all pairs were correctly retrieved once. We did not use spell checking during learning, meaning that in some cases pairs could have been essentially retrieved more than once. However, we do not believe this would differ across conditions to affect learning results.

Before base pair learning, subjects were told they would learn a new list of pairs that may or may not change between lists. Then base pair learning proceeded identically to supplemental pair learning, except that 45 rather than 36 pairs were learned.

At final test, we asked subjects to first recall all words from the base list (which we called the 2nd list). Test trials involved a one-second fixation cross followed by the cue word and unlimited time to respond. Subjects received no feedback during this test. Following base pair testing, we asked subjects to recall all words from the supplemental list (which we called the 1st list), which had the same format. Finally, subjects were debriefed and allowed to leave.

Statistics.

In each experiment, comparisons across conditions used one-way (condition: No Δ , Δ Target, Δ Cue, Δ Both, Control), repeated-measures ANOVAs. When significant, these were followed up using pairwise, within-subject, FDR-corrected (Benjamini & Hochberg, 1995) *t*-tests.

Memorability, memory dependence, and bootstrapping analyses were performed identically to our prior paper (Antony et al., 2022), as were methods for creating scatterplots and 2-D and 3-D surface plots. Briefly, for the memorability correlation analyses, we calculated the proportion of subjects recalling each base pair for each experimental condition minus the proportion in the Control condition. Next, we ran linear regression between the experimental – Control memorability of that pair and its relatedness value. In the Δ Both condition, we added cue + target relatedness values together. The y-intercept of the ordinary least squares line was theoretically meaningful in some analyses – that is, 0 AS means 0 subjects might endorse a word in a free association task, but 0 GloVe cosine similarity is not necessarily more meaningful than other low GloVe values. The slope indicated the relationship between relatedness and memorability. Two MATLAB functions were used in the memorability plots. The slope and intercept were added to each plot based on *p* values from ‘fitlm’, and best-fit lines were added along with the confidence error output from ‘polypredci’ (Strider, 2021).

Across-experiment memory comparisons involved subtracting the No Δ from all other conditions and running between-subject corrected *t*-tests within the same condition. Results from these tests were corrected for multiple comparisons using FDR correction.

Acknowledgements

The authors would especially like to thank Rebecca Slagle Luenser and Anthony Vierra for their work coordinating data collection on these experiments, as well as Valency Jarvis and Rachel Saconi for extensive data collection efforts across multiple terms. The authors would also like to thank the following individuals (all Cal Poly undergraduate students) for running participants across the five experiments: Nicole Attridge, Leigh Ann Bardman, Nicole Brault, Kylie Capella, Priyanka Chandar, Macey Coffman, Sadie Cooper, Rasha Demeter, Caprice Depetro, Chloe Fleischer, Daniel Frazier, Samantha Garrett, Trevor Guerra, Olivia Harrington, McKenzie Harrison, Amanda Hill, Erika Holloway, Sarah Howie, An Huynh, Rebecca Iniguez, Katie Johansen, Emma Klebaner, Carlo Lopiccolo, Jayelin Lombard, Angelo Lozano, Corinne Lykins, Rachel Nebel, Jordyn Niemiec, Sahar Oliaei, Suhana Patel, Katherine Preston, Pilar Reyes, Hector Reyes, Elli Rouse, Macy Rowe, Nami Saito, Lily Sanz, Michael Scarpa, Colin Schmitt, Christina Schwake, Evan Schweitzer, Jessica Shaver, Michela Smith, Luke Smith, Isabella Strawn, Aly Tierney, Sarah Tung, Jacob Van Dam, Mia Venturini, and Chase Volz.

Declaration of interests

The authors declare no competing financial interests.

Author note

Code / data for this project can be found at <https://osf.io/hmj8b/> .

Supplementary figures

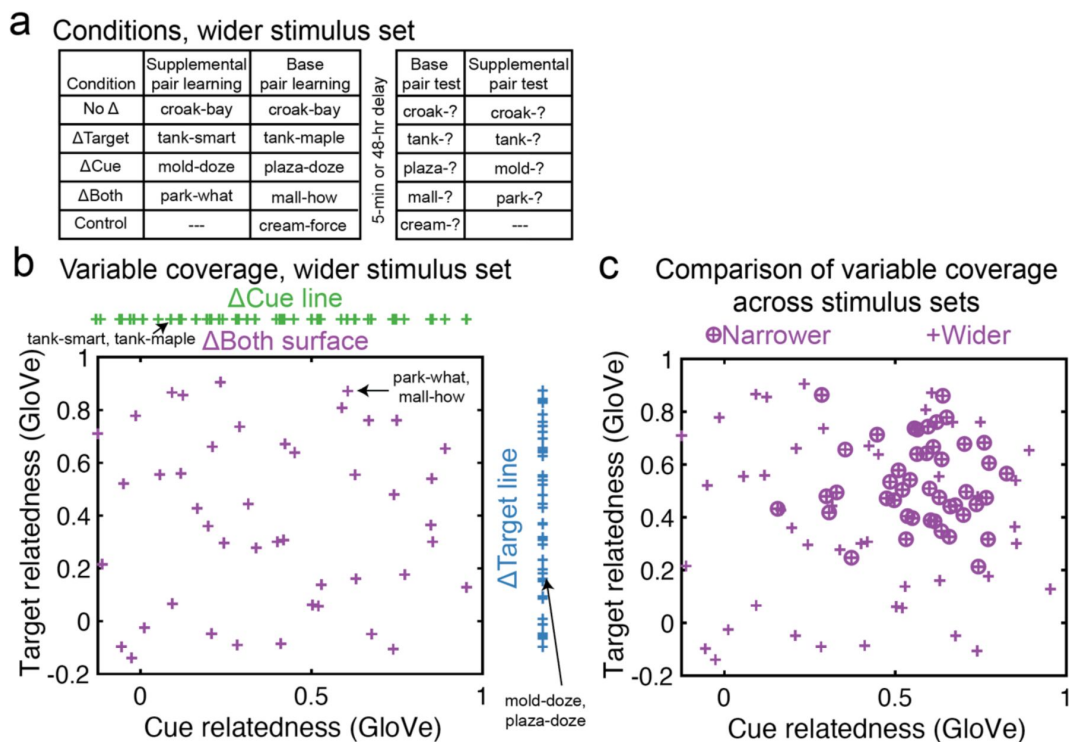


Figure 1-Supp 1

Wider stimulus set examples and coverage. (a) Examples of stimuli from the wider stimulus set for all conditions. (b) Variable coverage. Distributions of cue and target GloVe values are shown along the target identity (green) and cue identity (blue) lines, respectively. Distributions of GloVe values are shown as purple points inside the 2-D surface (scatterplot) for the Δ Both condition. (c) Comparison of variable coverage in narrower and wider stimulus sets using the common GloVe metric.

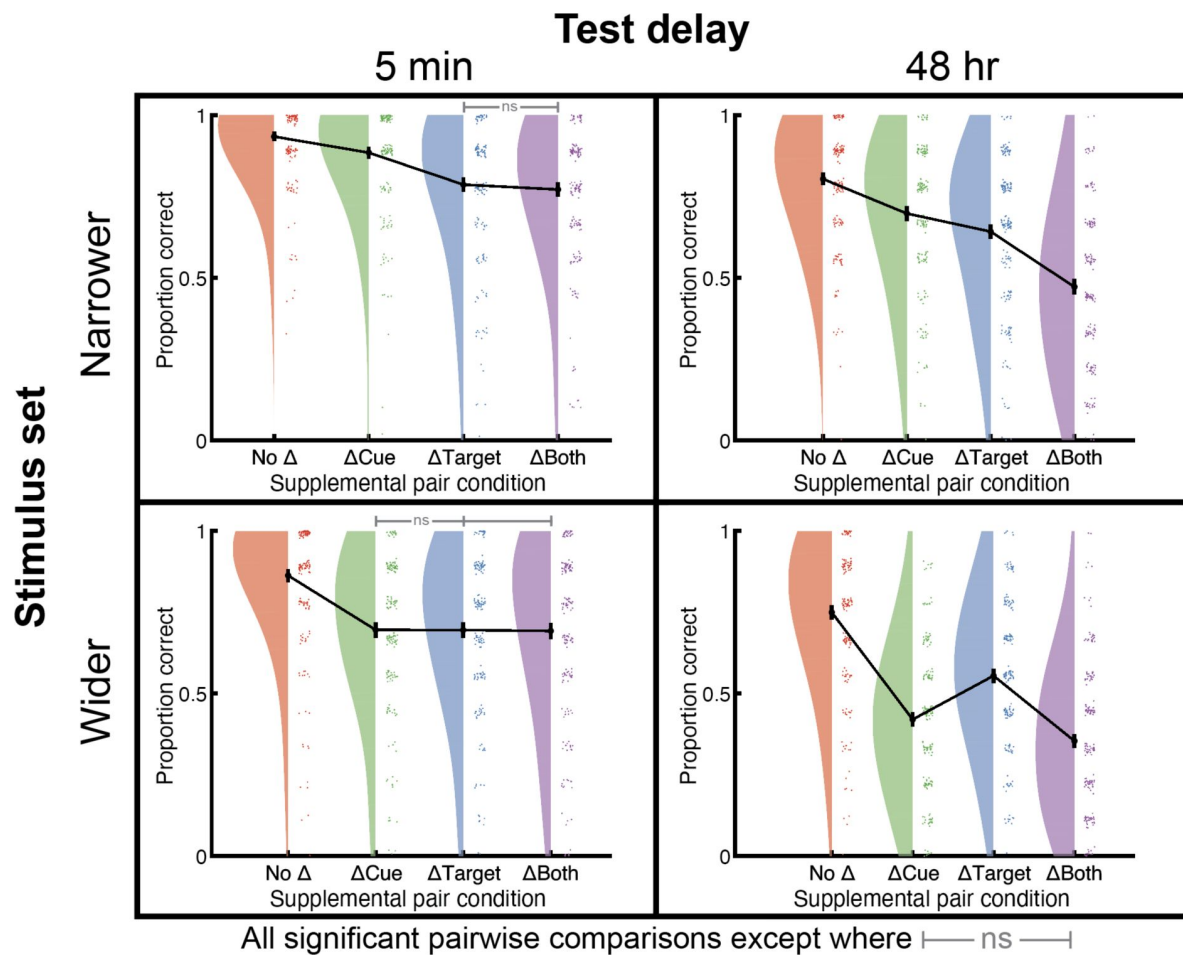


Figure 2-Supp 1

Supplemental pair memory differed as a function of overall stimulus set relatedness, delay, and word pair condition. All comparisons were significant except where labeled with gray bars and 'ns' ($p > 0.1$), and data points from individual subjects were jittered slightly for better visualization.

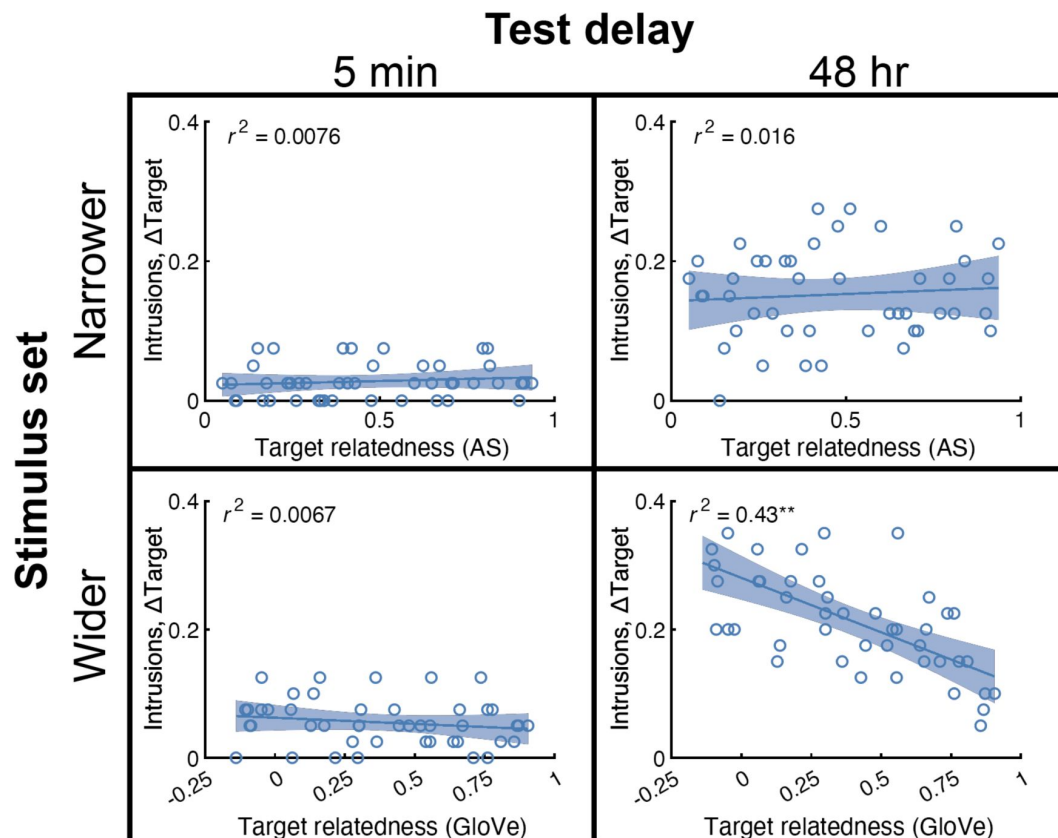


Figure 3-Supp 1

Intrusions did not increase with target relatedness in the Δ Target condition in any experiment. We plotted across-subject intrusion rates of reporting supplemental pairs during base pair testing against target relatedness [Top: AS; bottom: GloVe]. There were no positive relationships between intrusions and target relatedness, though there was a significantly negative relationship in the wider stimulus set, 5-min delay experiment.

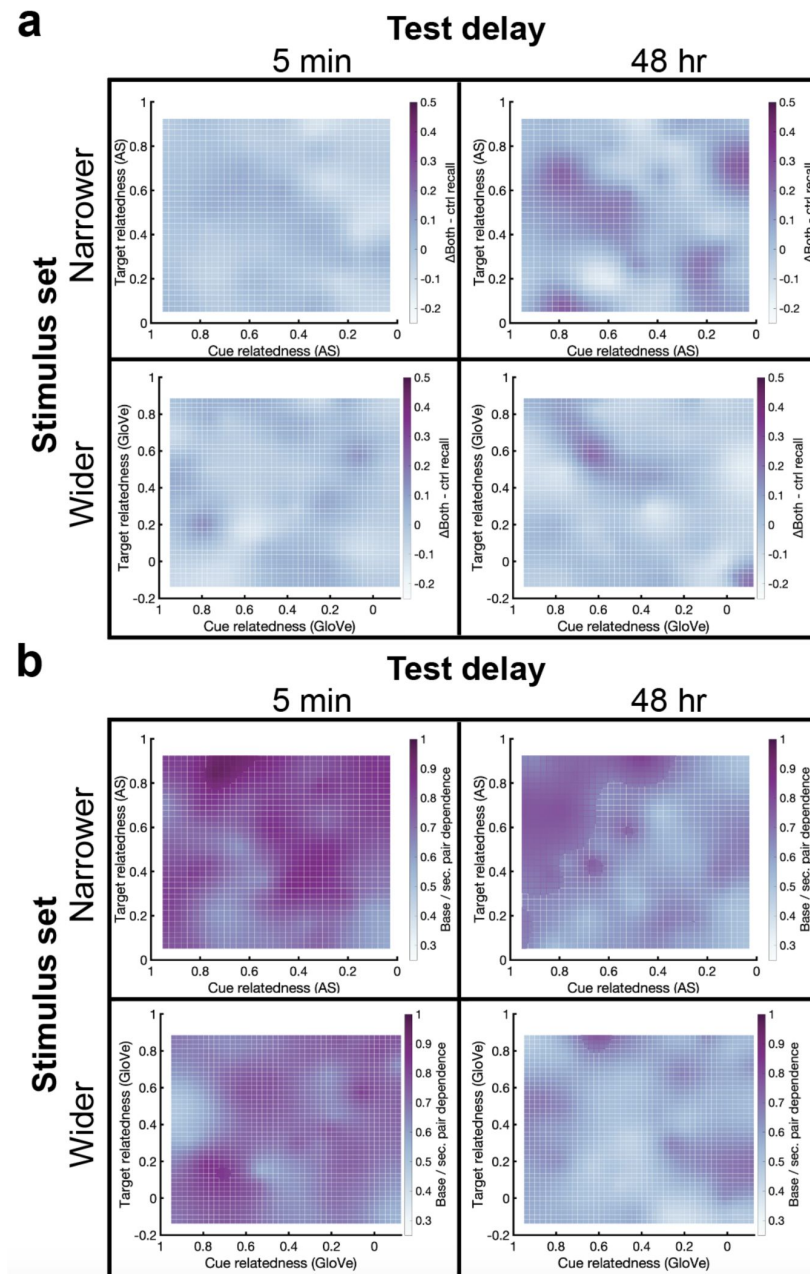


Figure 5-Supp 1

Bivariate cue and target relatedness had no effect on memorability, though it did significantly predict interdependence in the narrower stimulus set in the 48-hr delay experiment. These plots are created by putting a white 95th percentile threshold of bootstrapped samples in white grid marks above the purple (true) data, so the presence of purple indicates significance at the $p < 0.01$ level. (a) Across-subject memorability for each base pair – control was plotted against the bivariate level of cue and target relatedness [top: AS; bottom: GloVe]. (b) Across-subject dependence for each base pair target-supplemental pair target duo was plotted against cue and target relatedness. High levels of cue and target relatedness increased dependence in the narrower stimulus set, 48-hr delay experiment.

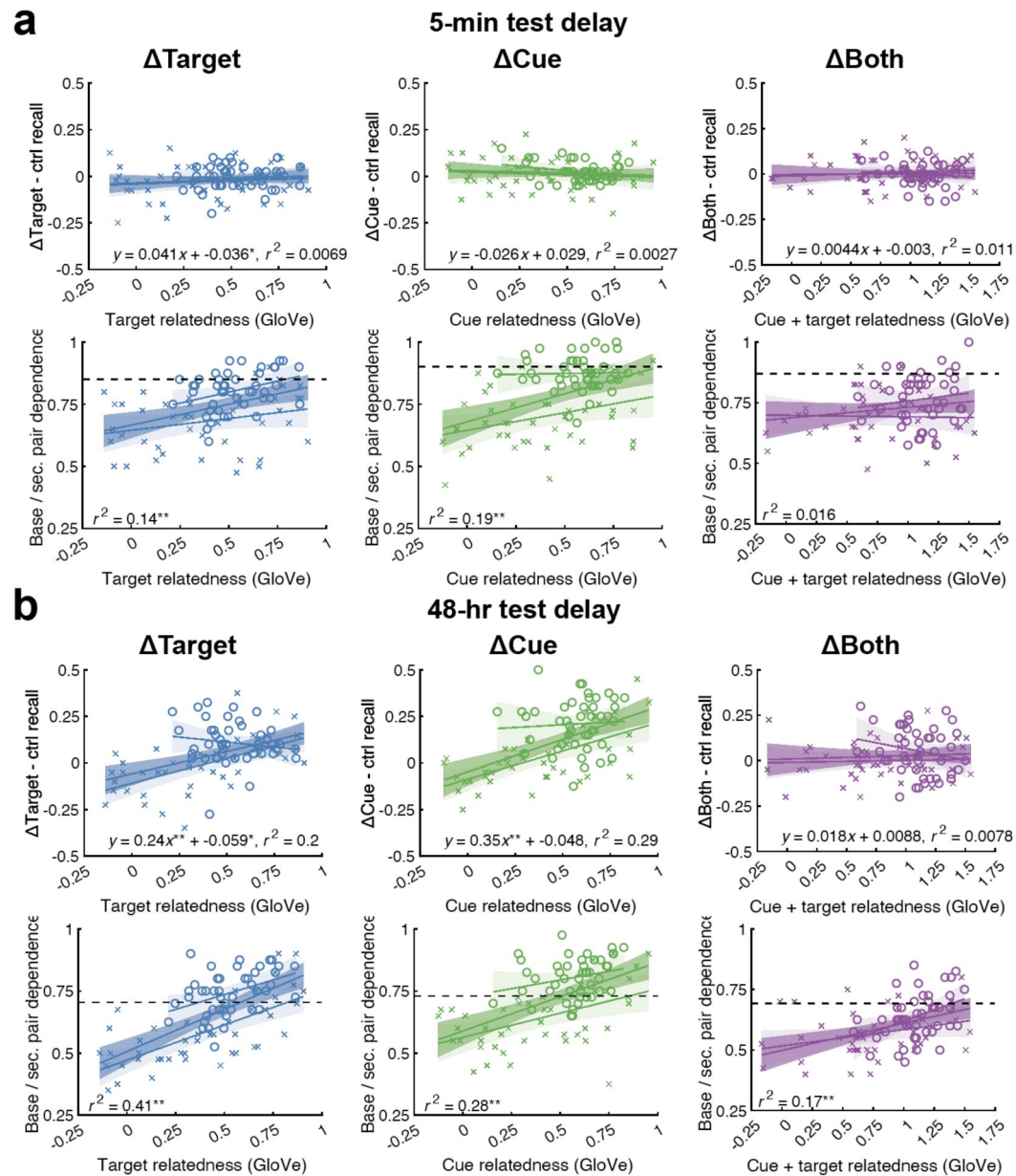


Figure 6-Supp 1

Memory and interdependence data from preceding analyses plotted using a common GloVe metric. (a-b) Across-subject memorability (top) and dependence (bottom) for the Δ Target (left), Δ Cue (middle), and Δ Both (right) conditions were plotted after combining narrower and wider stimulus sets within a delay. In the 5-min delay experiments (a), we found no significant relationships between relatedness and memorability, though there were relationships between relatedness and interdependence in the Δ Target and Δ Cue conditions. In the 48-hr delay experiments, relatedness predicted memorability in the Δ Target and Δ Cue conditions, and relatedness predicted interdependence in all conditions. Correlations between the combined correlation are plotted in darker colors, with each individual correlation between stimulus set and relatedness replotted in lighter colors for visual comparison. Text labels reflect correlation values for the combined datasets.

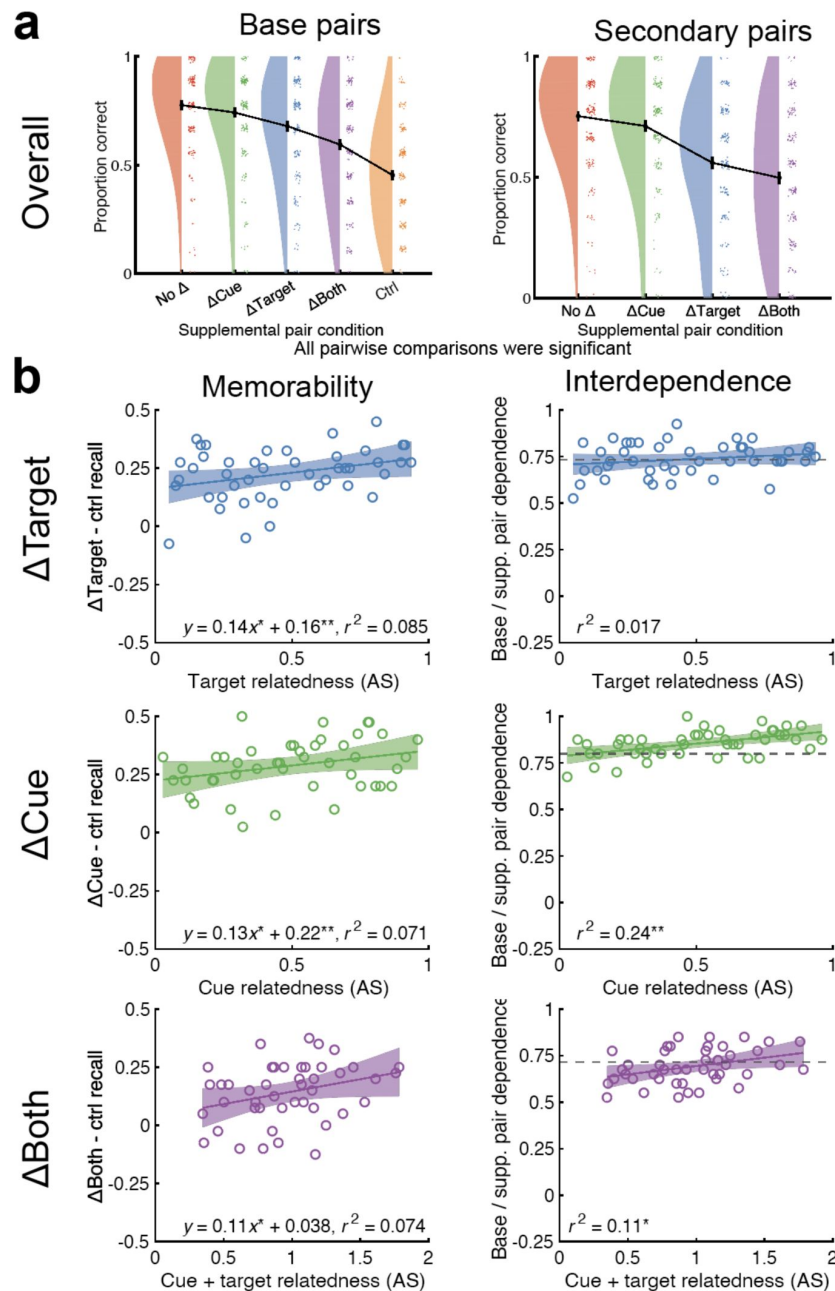


Figure 8-Supp 1

Memory and interdependence data in a paradigm in which we used base pair learning followed by supplemental pair learning and then flipped the order of the tests (supplemental pairs followed by base pairs). (a) Overall memory performance for base (left) and supplemental pairs (right) by condition. All comparisons were significant in both plots, and data points from individual subjects were jittered slightly for better visualization. (b) Correlations between retroactive memory effects (left) and base-supplemental pair dependence (right) in the Δ Target (top), and Δ Cue (middle) conditions. All correlations between memorability and relatedness were significant, whereas correlations between relatedness and interdependence were significant in the Δ Cue and Δ Both conditions. Pearson correlations are provided within the plots, followed by * when $p < 0.05$ and ** when $p < 0.01$.

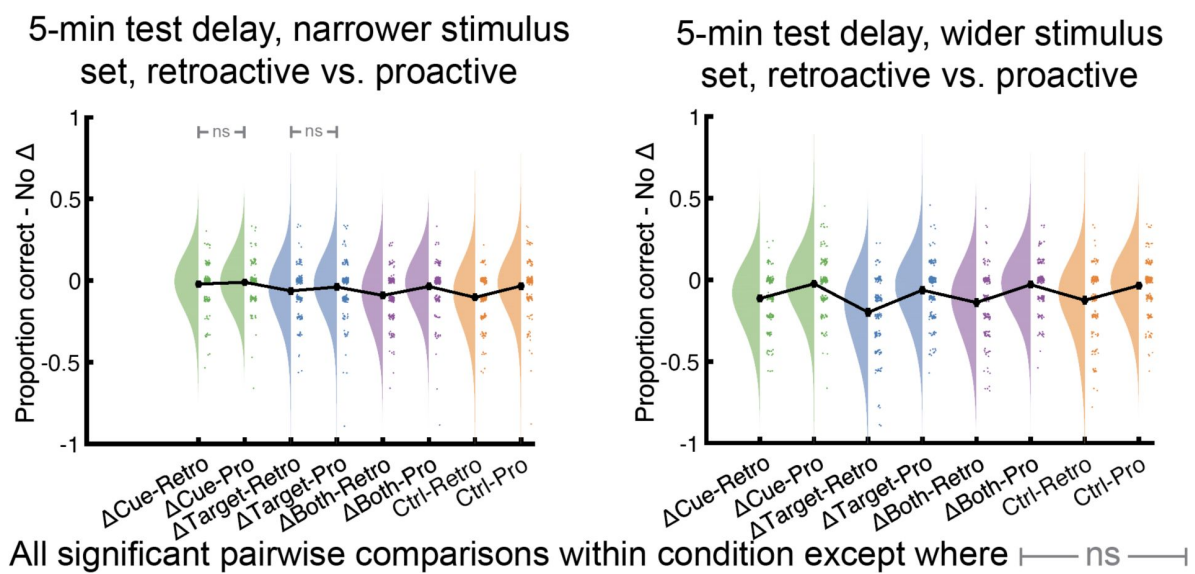


Figure 8-Supp 2

Direct comparisons of the same base pair condition across 5-minute delay experiments were performed between the narrower stimulus set, retroactive versus proactive experiments (left) and the wider stimulus set, retroactive versus proactive experiments (right).

References

- Amundson J. C., Escobar M., Miller R. R. (2003) **Proactive interference between cues trained with a common outcome in first-order Pavlovian conditioning** *Journal of Experimental Psychology: Animal Behavior Processes* **29**
- Anderson M. C., McCulloch K. C. (1999) **Integration as a general boundary condition on retrieval-induced forgetting** *Journal of Experimental Psychology: Learning Memory and Cognition* **25**:608–629
- Anderson M. C., Neely J. H. (1996) **Interference and inhibition in memory retrieval** *Memory Academic press* :237–313
- Antony J. W., Bennion K. A. (2022) **Semantic associates create retroactive interference on an independent recall task** *Journal of Experimental Psychology: Learning, Memory, and Cognition* :701–713 <https://doi.org/10.1037/xlm0001216>
- Antony J. W., Ferreira C. S., Norman K. A., Wimber M. (2017) **Retrieval as a fast route to memory consolidation** *Trends in Cognitive Sciences* **21**:573–576
- Antony J. W., Lozano A., Dhoat P., Chen J., Bennion K. A. (2024) **Causal and chronological relationships predict memory organization for nonlinear narratives** *Journal of Cognitive Neuroscience* :1–18
- Antony J. W., Romero A., Vierra A. H., Luenser R. S., Hawkins R. D., Bennion K. A. (2022) **Semantic relatedness retroactively boosts memory and promotes memory interdependence across episodes** *eLife* **11**
- Antony J. W., Schechtman E. (2023) **Reap while you sleep: Consolidation of memories differs by how they were sown** *Hippocampus* :922–935 <https://doi.org/10.1002/hipo.23526>
- Baddeley A. D., Dale H. C. A. (1966) **The effect of semantic similarity on retroactive interference in long- and short-term memory** *Journal of Verbal Learning and Verbal Behavior* **5**:417–420
- Baek M., Papaj D. R. (2024) **The Role of Similarity of Stimuli and Responses in Learning by Nectar-Foraging Bumble Bees: A Test of Osgood’s Model** <https://doi.org/10.1101/2024.04.25.591145>
- Baran B., Wilson J., Spencer R. M. C. (2010) **REM-dependent repair of competitive memory suppression** *Experimental Brain Research* **203**:471–477
- Barnes J. M., Underwood B. J. (1959) **“Fate” of first-list associations in transfer theory** *Journal of Experimental Psychology* :97–105 <https://doi.org/10.1017/9781316477045.005>
- Bartlett F. C. (1932) **Remembering** Cambridge University Press
- Bein O., Livneh N., Reggev N., Gilead M., Goshen-Gottstein Y., Maril A. (2015) **Delineating the effect of semantic congruency on episodic memory: The role of integration and relatedness** *PLoS ONE* **10**

- Benjamin A. S., Tullis J. (2010) **What makes distributed practice effective?** *Cognitive Psychology* **61**:228–247
- Benjamini Y., Hochberg Y. (1995) **Controlling the false discovery rate: A practical and powerful approach to multiple testing** *Journal of the Royal Statistical Society. Series B* **57**:289–300
- Bower G. H., Thompson-Schill S., Tulving E. (1994) **Reducing retroactive interference: An interference analysis** *Journal of Experimental Psychology: Learning, Memory, and Cognition* **20**:51–66
- Briggs G. E. (1954) **Acquisition, extension, and recovery functions in retroactive inhibition** *Journal of Experimental Psychology* **47**:285–293
- Brod G., Werkle-Bergner M., Shing Y. L. (2013) **The influence of prior knowledge on memory: A developmental cognitive neuroscience perspective** *Frontiers in Behavioral Neuroscience* <https://doi.org/10.3389/fnbeh.2013.00139>
- Bruce R. W. (1933) **Conditions of transfer of training** *Journal of Experimental Psychology* **16**:343–361
- Brunmair M., Richter T. (2019) **Similarity matters: A meta-analysis of interleaved learning and its moderators** *Psychological Bulletin* **145**
- Bugelski B. R., Cadwallader T. C. (1956) **A reappraisal of the transfer and retroaction surface** *Journal of Experimental Psychology* **52**:360–366
- Caplan J. B., Rehani M., Andrews J. C. (2014) **Associations compete directly in memory** *Quarterly Journal of Experimental Psychology* **67**:955–978
- Carpenter S. K., Pashler H., Wixted J. T., Vul E. (2008) **The effects of tests on learning and forgetting** *Memory & Cognition* **36**:438–448
- Carroll M., Campbell-Ratcliffe J., Murnane H., Perfect T. (2007) **Retrieval-induced forgetting in educational contexts: Monitoring, expertise, text integration, and test format** *European Journal of Cognitive Psychology* :580–606 <https://doi.org/10.1080/09541440701326071>
- Castro L., Ortega N., Matute H. (2002) **Proactive Interference in Human Predictive Learning** *International Journal of Comparative Psychology* <https://doi.org/10.46867/C4DG6B>
- Chan J. C. K. (2009) **When does retrieval induce forgetting and when does it induce facilitation? Implications for retrieval inhibition, testing effect, and text processing** *Journal of Memory and Language* :153–170 <https://doi.org/10.1016/j.jml.2009.04.004>
- Cox W. R., Dobbelaar S., Meeter M., Kindt M., van Ast V. A. (2021) **Episodic memory enhancement versus impairment is determined by contextual similarity across events** *Proceedings of the National Academy of Sciences of the United States of America* <https://doi.org/10.1073/pnas.2101509118>
- Dallett K. M. (1962) **The transfer surface re-examined** *Journal of Verbal Learning and Verbal Behavior* **1**:91–94
- Dallett K. M. (1964) **Proactive and retroactive inhibition in the A-B, A-B' paradigm** *Journal of Experimental Psychology* **68**:190–200

- Eich J. M. (1982) **A composite holographic associative recall model** *Psychological Review* **89**:627–661
- Estes W. K. (1979) **Role of response availability in the effects of cued-recall tests on memory** *Journal of Experimental Psychology: Human Learning and Memory* :567–573 <https://doi.org/10.1037/0278-7393.5.6.567>
- Garlitch S. M., Wahlheim C. N. (2020) **The role of reminding in retroactive effects of memory for older and younger adults** *Psychology and Aging* :697–709 <https://doi.org/10.1037/pag0000427>
- Gates A. I. (1917) **Recitation as a factor in memorizing** *Archives of Psychology* **6**:1–104
- Greenberg R., Underwood B. J. (1950) **Retention as a function of stage of practice** *Journal of Experimental Psychology* **40**
- Hamilton R. J. (1943) **Retroactive facilitation as a function of degree of generalization between tasks** *Journal of Experimental Psychology* **32**:363–376
- Hintzman D. L. (2011) **Research strategy in the study of memory: Fads, fallacies, and the search for the “coordinates of truth.”** *Perspectives on Psychological Science* **6**:253–271
- Hintzman D. L., Summers J. J., Block R. A. (1975) **Spacing judgments as an index of study-phase retrieval** *Journal of Experimental Psychology: Learning, Memory, and Cognition* :31–40 <https://doi.org/10.1037/0278-7393.1.1.31>
- Holding D. H. (1976) **An approximate transfer surface** *Journal of Motor Behavior* :1–9 <https://doi.org/10.1080/00222895.1976.10735049>
- Horner A. J., Bisby J. A., Bush D., Lin W.-J., Burgess N. (2015) **Evidence for holistic episodic recollection via hippocampal pattern completion** *Nature Communications* **6**:7462–7462
- Houston J. P. (1966) **Stimulus recall and experimental paradigm** *Journal of Experimental Psychology* :619–621 <https://doi.org/10.1037/h0023766>
- Jacoby L. L., Wahlheim C. N. (2013) **On the importance of looking back: The role of recursive reminders in recency judgments and cued recall** *Memory and Cognition* **41**:625–637
- Jacoby L. L., Wahlheim C. N., Kelley C. M. (2015) **Memory consequences of looking back to notice change: Retroactive and proactive facilitation** *Journal of Experimental Psychology: Learning Memory and Cognition* **41**:1282–1297
- Jarrett R. F., Scheibe K. E. (1963) **Association chains and paired-associate learning** *Journal of Verbal Learning and Verbal Behavior* :264–268 [https://doi.org/10.1016/S0022-5371\(63\)80005-5](https://doi.org/10.1016/S0022-5371(63)80005-5)
- Jung J. (1962) **Transfer of training as a function of degree of first-list learning** *Journal of Verbal Learning and Verbal Behavior* :197–199 [https://doi.org/10.1016/S0022-5371\(62\)80030-9](https://doi.org/10.1016/S0022-5371(62)80030-9)
- Jung J. (1967) **Differentiation and negative transfer as a function of degree of first- and second-list learning** *Psychological Reports* :73–80 <https://doi.org/10.2466/pr0.1967.21.1.73>
- Kanungo R. (1967) **Meaning mediation in verbal transfer** *British Journal of Psychology* **58**:205–212

- Kemp P. L., Loaiza V. M., Wahlheim C. N. (2023) **Testing can enhance episodic memory updating in younger and older adults** *Psychology and Aging* :656–669 <https://doi.org/10.1037/pag0000776>
- Keppel G., Underwood B. J. (1962) **Retroactive inhibition of R-S associations** *Journal of Experimental Psychology* :400–404 <https://doi.org/10.1037/h0040424>
- Lee H., Chen J. (2022) **Predicting memory from the network structure of naturalistic events** *Nature Communications* <https://doi.org/10.1038/s41467-022-31965-2>
- Lifanov J., Linde-Domingo J., Wimber M. (2021) **Feature-specific reaction times reveal a semanticisation of memories over time and with repeated remembering** *Nature Communications* <https://doi.org/10.1038/s41467-021-23288-5>
- Liu X. L., Ranganath C. (2019) **Resurrected Memories: Sleep-dependent memory consolidation saves memories from competition induced by retrieval practice** *PsyArXiv*
- Lustig C., Konkel A., Jacoby L. L. (2004) **Which route to recovery? Controlled retrieval and accessibility bias in retroactive interference** *Psychological Science* **15**:729–735
- Mandler G., Heinemann S. H. (1956) **Effect of overlearning of a verbal response on transfer of training** *Journal of Experimental Psychology* **52**
- Martin E. (1965) **Transfer of verbal paired associates** *Psychological Review* **72**:327–343
- Martin R. B., Dean S. J. (1964) **Implicit and explicit mediation paired-associate learning** *Journal of Experimental Psychology* **68**:21–27
- McCrystal T. J. (1970) **List differentiation as a function of time and test order** *Journal of Experimental Psychology* **83**
- McGeoch J. A., McDonald W. T. (1931) **Meaningful relation and retroactive inhibition** *The American Journal of Psychology* **43**:579–588
- McGeoch J. A., McGeoch G. O. (1937) **Studies in retroactive inhibition: X. The influence of similarity of meaning between lists of paired associates** *Journal of Experimental Psychology* **21**:320–329
- McKinley G. L., Benjamin A. S. (2020) **The role of retrieval during study: Evidence of reminding from overt rehearsal** *Journal of Memory and Language* <https://doi.org/10.1016/j.jml.2020.104128>
- McKinley G. L., Ross B. H., Benjamin A. S. (2019) **The role of retrieval during study: Evidence of reminding from self-paced study time** *Memory & Cognition* :877–892 <https://doi.org/10.3758/s13421-019-00897-6>
- Mehler J., Miller G. A. (1964) **Retroactive interference in the recall of simple sentences** *British Journal of Psychology* :295–301 <https://doi.org/10.1111/j.2044-8295.1964.tb00913.x>
- Melega G., Sheldon S. (2023) **Conceptual relatedness promotes memory generalization at the cost of detailed recollection** *Scientific Reports* <https://doi.org/10.1038/s41598-023-40803-4>

- Melton A. W., Von Lackum W. J. (1941) **Retroactive and proactive Inhibition in retention: Evidence for a two-factor theory of retroactive inhibition** *The American Journal of Psychology* <https://doi.org/10.2307/1416789>
- Mensink G.-J., Raaijmakers J. G. (1988) **A model for interference and forgetting** *Psychological Review* **95**:434–455
- Metcalfe J., Schwartz B. L., Joaquim S. G. (1993) **The cue-familiarity heuristic in metacognition** *Journal of Experimental Psychology: Learning, Memory, and Cognition* :851–861 <https://doi.org/10.1037/0278-7393.19.4.851>
- Mills J. A. (1973) **Effects of interitem associative strength, rehearsal, and proactive inhibition on the retention of paired-associate lists** *The Journal of Psychology* **85**:313–321
- Moeser S. D. (1979) **The role of experimental design in investigations of the fan effect** *Journal of Experimental Psychology: Human Learning and Memory* **5**:125–134
- Morgan R. L., Underwood B. J. (1950) **Proactive inhibition as a function of response similarity** *Journal of Experimental Psychology* **40**:592–603
- Morton N. W., Zippi E. L., Preston A. R. (2023) **Memory reactivation and suppression modulate integration of the semantic features of related memories in hippocampus** *Cerebral Cortex* :9020–9037 <https://doi.org/10.1093/cercor/bhad179>
- Nelson D. L., McEvoy C. L., Schreiber T. A. (1998) **The University of South Florida word association, rhyme, and word fragment norms**
- Nicholas J., Daw N. D., Shohamy D. (2023) **Proactive and reactive construction of memory-based preferences** *bioRxiv* <https://doi.org/10.1101/2023.12.10.570977>
- Ortega A., Gómez-Ariza C. J., Morales J., Bajo M. T. (2015) **Low involvement of preexisting associations makes retrieval-induced forgetting long lasting** *Cognitive Processing* :121–130 <https://doi.org/10.1007/s10339-015-0650-8>
- Osgood C. E. (1946) **Meaningful similarity and interference in learning** *Journal of Experimental Psychology* **36**:277–301
- Osgood C. E. (1949) **The similarity paradox in human learning: A resolution** *Psychological Review* **56**:132–143
- Palermo D. S., Jenkins J. J. (1964) **Paired-associate learning as a function of the strength of links in the associative chain** *Journal of Verbal Learning and Verbal Behavior* :406–412 [https://doi.org/10.1016/S0022-5371\(64\)80010-4](https://doi.org/10.1016/S0022-5371(64)80010-4)
- Pennington J., Socher R., Manning C. D. (2014) **GloVe: Global vectors for word representation** *EMNLP 2014 - 2014 Conference on Empirical Methods in Natural Language Processing, Proceedings of the Conference* :1532–1543
- Postman L., Cofer B. S., Musgrave Charles N. (1961) **The present status of interference theory** *Verbal learning and verbal behavior* McGraw-Hill Book Company :152–196
- Postman L. (1962) **Retention as a function of degree of overlearning** *Science* **135**:666–667

- Postman L., Keppel G. (1977) **Conditions of cumulative proactive inhibition** *Journal of Experimental Psychology. General* **106**:376–403
- Postman L., Parker J. (1970) **Maintenance of first-list associations during transfer** *The American Journal of Psychology* **83**:171–188
- Postman L., Stark K., Henschel D. (1969) **Conditions of recovery after unlearning** *Journal of Experimental Psychology* **82**
- Postman L., Underwood B. J. (1973) **Critical issues in interference theory** *Memory & Cognition* :19–40 <https://doi.org/10.3758/BF03198064>
- Reder L. M., Anderson J. R. (1980) **A partial resolution of the paradox of interference: The role of integrating knowledge** *Cognitive Psychology* :447–472 [https://doi.org/10.1016/0010-0285\(80\)90016-X](https://doi.org/10.1016/0010-0285(80)90016-X)
- Roediger H. L., Karpicke J. D. (2006) **The power of testing memory: Basic research and implications for educational practice** *Perspectives on Psychological Science* **1**:181–210
- Rowland C. A. (2014) **The effect of testing versus restudy on retention: A meta-analytic review of the testing effect** *Psychological Bulletin* :1432–1463 <https://doi.org/10.1037/a0037559>
- Russell W. A., Storms L. H. (1955) **Implicit verbal chaining in paired-associate learning** *Journal of Experimental Psychology* :287–293 <https://doi.org/10.1037/h0042642>
- Sahakyan L., Goodmon L. B. (2007) **The influence of directional associations on directed forgetting and interference** *Journal of Experimental Psychology: Learning, Memory, and Cognition* :1035–1049 <https://doi.org/10.1037/0278-7393.33.6.1035>
- Saltz E., Hamilton H. (1967) **Spontaneous recovery of list 1 responses in the A-B, A'-C paradigm** *Journal of Experimental Psychology* **75**:267–273
- Schlichting M. L., Mumford J. a, Preston A. R. (2015) **Learning-related representational changes reveal dissociable integration and separation signatures in the hippocampus and prefrontal cortex** *Nature Communications* <https://doi.org/10.1038/ncomms9151>
- Schlichting M. L., Preston A. R., Hannula D. E., Duff M. C. (2017) **The Hippocampus and Memory Integration: Building Knowledge to Navigate Future Decisions** *The Hippocampus from Cells to Systems: Structure, Connectivity, and Functional Contributions to Memory and Flexible Cognition* Springer International Publishing :405–437 https://doi.org/10.1007/978-3-319-50406-3_13
- Shohamy D., Daw N. D. (2015) **Integrating memories to guide decisions** *Current Opinion in Behavioral Sciences* :85–90 <https://doi.org/10.1016/j.cobeha.2015.08.010>
- Siegel L. L., Kahana M. J. (2014) **A retrieved context account of spacing and repetition effects in free recall** *Journal of Experimental Psychology: Learning, Memory, and Cognition* :755–764 <https://doi.org/10.1037/a0035585>
- Strider S. (2021) **Polypredci [Computer software]**

- Szpunar K. K., McDermott K. B., Roediger H. L. (2008) **Testing during study insulates against the buildup of proactive interference** *Journal of Experimental Psychology. Learning, Memory, and Cognition* :1392–1399 <https://doi.org/10.1037/a0013082>
- Thios S. J., D’Agostino P. R. (1976) **Effects of repetition as a function of study-phase retrieval** *Journal of Verbal Learning and Verbal Behavior* **15**:529–536
- Trabasso T., Sperry L. L. (1985) **Causal relatedness and importance of story events** *Journal of Memory and Language* **24**:595–611
- Traxler A. J. (1973) **Retroactive and proactive inhibition in young and elderly adults using an unpaced modified free recall test** *Psychological Reports* :215–222 <https://doi.org/10.2466/pr0.1973.32.1.215>
- Tse D., Langston R. F., Kakeyama M., Bethus I., Spooner P. a, Wood E. R., Witter M. P., Morris R. G. M. (2007) **Schemas and memory consolidation** *Science* :76–82 <https://doi.org/10.1126/science.1135935>
- Tullis J. G., Benjamin A. S., Ross B. H. (2014) **The reminding effect: Presentation of associates enhances memory for related words in a list** *Journal of Experimental Psychology: General* :1526–1540 <https://doi.org/10.1037/a0036036>
- Tulving E., Watkins M. J. (1974) **On negative transfer: Effects of testing one list on the recall of another** *Journal Of Verbal Learning And Verbal Behavior* **13**:181–193
- Twedt H. M., Underwood B. J. (1959) **Mixed vs. Unmixed lists in transfer studies** *Journal of Experimental Psychology* :111–116 <https://doi.org/10.1037/h0043795>
- Underwood B. J. (1948) **Retroactive and proactive inhibition after five and forty-eight hours** *Journal of Experimental Psychology* **38**
- Underwood B. J. (1949) **Proactive inhibition as a function of time and degree of prior learning** *Journal of Experimental Psychology* **39**
- Underwood B. J. (1951) **Associative transfer in verbal learning as a function of response similarity and degree of first-list learning** *Journal of Experimental Psychology* :44–53 <https://doi.org/10.1037/h0059484>
- Underwood B. J. (1957) **Interference and forgetting** *Psychological Review* **64**:49–60
- Underwood B. J. (1969) **Attributes of memory** *Psychological Review* **76**:559–573
- Van Kesteren M. T. R., Ruiter D. J., Fernández G., Henson R. N. (2012) **How schema and novelty augment memory formation** *Trends in Neurosciences* :211–219 <https://doi.org/10.1016/j.tins.2012.02.001>
- Wahlheim C. N. (2014) **Proactive effects of memory in young and older adults: The role of change recollection** *Memory and Cognition* :950–964 <https://doi.org/10.3758/s13421-014-0411-4>
- Wahlheim C. N. (2015) **Testing can counteract proactive interference by integrating competing information** *Memory & Cognition* **43**:27–38

- Wahlheim C. N., Alexander T. R., Peske C. D. (2020) **Reminders of Everyday Misinformation Statements Can Enhance Memory for and Beliefs in Corrections of Those Statements in the Short Term** *Psychological Science* :1325–1339 <https://doi.org/10.1177/0956797620952797>
- Wahlheim C. N., Garlitch S. M., Kemp P. L. (2021) **Context differentiation and reminders in episodic memory updating** *Psychology of Learning and Motivation* Elsevier :245–277 <https://doi.org/10.1016/bs.plm.2021.06.001>
- Wahlheim C. N., Jacoby L. L. (2013) **Remembering change: The critical role of recursive reminders in proactive effects of memory** *Memory and Cognition* **41**:1–15
- Wahlheim C. N., Smith S. T., Garlitch S. M., Wiley R. W. (2023) **Interpolated retrieval retroactively increases recall and promotes cross-episode memory interdependence** *Learning & Memory* :151–163 <https://doi.org/10.1101/lm.053782.123>
- Wahlheim C. N., Zacks J. M. (2019) **Memory guides the comprehension of event changes for older and younger adults** *Journal of Experimental Psychology: General* **148**:30–50
- Wahlheim C.N., Antony J.W., Kelley C.M. **Subsequent memory effects of mnemonic discrimination on episodic memory updating**
- Walsh C. R., Rissman J. (2023) **Behavioral representational similarity analysis reveals how episodic learning is influenced by and reshapes semantic memory** *Nature Communications* <https://doi.org/10.1038/s41467-023-42770-w>
- Watkins O. C., Watkins M. J. (1975) **Buildup of proactive inhibition as a cue-overload effect** *Journal of Experimental Psychology: Human Learning & Memory* **1**:442–452
- Wheeler M. A. (1995) **Improvement in recall over time without repeated testing: Spontaneous recovery revisited** *Journal of Experimental Psychology: Learning, Memory, and Cognition* :173–184 <https://doi.org/10.1037//0278-7393.21.1.173>
- Wimer R. (1964) **Osgood’s transfer surface: Extension and test** *Journal of Verbal Learning and Verbal Behavior* **3**:274–279
- Wimmer G. E., Shohamy D. (2012) **Preference by association: How memory mechanisms in the hippocampus bias decisions** *Science* :270–273 <https://doi.org/10.1126/science.1223252>
- Witherby A. E., Carpenter S. K. (2022) **The rich-get-richer effect: Prior knowledge predicts new learning of domain-relevant information** *Journal of Experimental Psychology: Learning, Memory, and Cognition* :483–498 <https://doi.org/10.1037/xlm0000996>
- Wixted J. T. (2004) **The psychology and neuroscience of forgetting** *Annual Review of Psychology* **55**:235–269
- Young R. K. (1955) **Retroactive and proactive effects under varying conditions of response similarity** *Journal of Experimental Psychology* **50**:113–119
- Yu W., Duncan K., Schlichting M. (2023) **Using retrieval contingencies to understand memory integration and inference** <https://doi.org/10.31234/osf.io/eafsh>
- Zou F., Wanjia G., Allen E. J., Wu Y., Charest I., Naselaris T., Kay K., Kuhl B. A., Hutchinson J. B., DuBrow S. (2023) **Re-expression of CA1 and entorhinal activity patterns preserves temporal context memory at long timescales** *Nature Communications* **14**

Author information

Kelly A Bennion *

Department of Psychology and Child Development, California Polytechnic State University,
San Luis Obispo, United States

For correspondence: kbennion@calpoly.edu

*These authors contributed equally to this work

Jade Phong

Department of Psychology and Child Development, California Polytechnic State University,
San Luis Obispo, United States

Mytien Le

Department of Psychology and Child Development, California Polytechnic State University,
San Luis Obispo, United States

Kunhua Cheng

Department of Psychology and Child Development, California Polytechnic State University,
San Luis Obispo, United States

Christopher N Wahlheim

Department of Psychology, University of North Carolina at Greensboro, Greensboro, United
States

James W Antony *

Department of Psychology and Child Development, California Polytechnic State University,
San Luis Obispo, United States

For correspondence: jwantony@calpoly.edu

*These authors contributed equally to this work

Editors

Reviewing Editor

Laura Colgin

University of Texas at Austin, Austin, United States of America

Senior Editor

Laura Colgin

University of Texas at Austin, Austin, United States of America

Reviewer #1 (Public review):

Summary:

Bennion and colleagues present a careful examination of how an earlier set of memories can either interfere with or facilitate memories formed later. This impressive work is a companion piece to an earlier paper by Antony and colleagues (2022) in which a similar

experimental design was used to examine how a later set of memories can either interfere with or facilitate memories formed earlier. This study makes contact with an experimental literature spanning 100 years, which is concerned with the nature of forgetting, and the ways in which memories for particular experiences can interact with other memories. These ideas are fundamental to modern theories of human memory, for example, paired-associates studies like this one are central to the theoretical idea that interference between memories is a much bigger contributor to forgetting than any sort of passive decay.

Strengths:

At the heart of the current investigation is a proposal made by Osgood in the 1940s regarding how paired associates are learned and remembered. In these experiments one learns a pair of items, A-B (cue-target), and then later learns another pair that is related in some way, either A'-B (changing the cue, delta-cue), or A-B' (changing the target, delta-target), or A'-B' (changing both, delta-both), where the prime indicates that item has been modified, and may be semantically related to the original item. The authors refer to the critical to-be-remembered pairs as base pairs. Osgood proposed that when the changed item is very different from the original item there will be interference, and when the changed item is similar to the original item there will be facilitation. Osgood proposed a graphical depiction of his theory in which performance was summarized as a surface, with one axis indicating changes to the cue item of a pair and the other indicating changes to the target item, and the surface itself necessary to visualize the consequences of changing both.

In the decades since Osgood's proposal, there have been many studies examining slivers of the proposal, e.g., just changing targets in one experiment, just changing cues in another experiment. Because any pair of experiments use different methods, this has made it difficult to draw clear conclusions about the effects of particular manipulations.

The current paper is a potential landmark, in that they manipulate multiple fundamental experimental characteristics using the same general experimental design. Importantly, they manipulate the semantic relatedness of the changed item to the original item, the delay between the study experience and the test, and which aspect of the pair is changed. Furthermore, they include both a positive control condition (where the exact same pair is studied twice), and a negative control condition (where a pair is only studied once, in the same phase as the critical base pairs). This allows them to determine when the prior learning exhibits an interfering effect relative to the negative control condition, and also allows them to determine how close any facilitative effects come to matching the positive control.

The results are interpreted in terms of a set of existing theories, most prominently the memory-for-change framework, which proposes a mechanism (recursive reminding) potentially responsible for the facilitative effects examined here. One of the central results is the finding that a stronger semantic relationship between a base pair and an earlier pair has a facilitative effect on both the rate of learning of the base pair and the durability of the memory for the base pair. This is consistent with the memory-for-change framework, which proposes that this semantic relationship prompts retrieval of the earlier pair, and the two pairs are integrated into a common memory structure that contains information about which pair was studied in which phase of the experiment. When semantic relatedness is lower, they more often show interference effects, with the idea being that competition between the stored memories makes it more difficult to remember the base pair.

This work represents a major methodological and empirical advance for our understanding of paired-associates learning, and it sets a laudably high bar for future work seeking to extend this knowledge further. By manipulating so many factors within one set of experiments, it fills a gap in the prior literature regarding the cognitive validity of an 80-year-old proposal by Osgood. The reader can see where the observed results match Osgood's theory and where they are inconclusive. This gives us insight, for example, into the necessity

of including a long delay in one's experiment, to observe potential facilitative effects. This point is theoretically interesting, but it is also a boon for future methodological development, in that it establishes the experimental conditions necessary for examining one or another of these facilitation or interference effects more closely.

The authors were exceptionally responsive to the suggestions of the reviewers, and the revisions have improved the theoretical clarity of the paper. I think the value of this work will grow with time, as memory researchers and theorists use it as a benchmark for new theory development. For example, the data from these experiments will undoubtedly be used to develop and constrain a new generation of computational models of paired-associates learning.

Weaknesses:

One minor weakness of the work is that the overarching theoretical framing does not necessarily specify the expected result for each and every one of the many effects examined. For example, with a narrower set of semantic associations being considered (all of which are relatively high associations) and a long delay, varying the semantic relatedness of the target item did not reliably affect the memorability of that pair. However, the same analysis showed a significant effect when the wider set of semantic associations was used. The positive result is consistent with the memory-for-change framework, but the null result isn't clearly informative to the theory. However, research is never done; comparing the results with the two sets of semantic associations is informative from a methodological perspective, in that it establishes the degree to which semantic relatedness must be altered to affect behavioral performance in a paired-associates task.

<https://doi.org/10.7554/eLife.95480.2.sa3>

Reviewer #2 (Public review):

Summary:

The study focuses on how relatedness with existing memories affects the formation and retention of new memories. Of core interest were the conditions that determine when prior memories facilitate new learning or interfere with it. Across a set of experiments that varied the degree of relatedness across memories as well as retention interval, the study compellingly shows that relatedness typically leads to proactive facilitation of new learning, with interference only observed under specific conditions and immediate test and being thus an exception rather than a rule.

Strengths:

The study uses a well-established word-pair learning paradigm to study interference and facilitation of overlapping memories. It however goes more in depth than a typical interference study in the systematic variation of several factors: (1) which elements of an association are overlapping and which are altered (change target, change cue, change both, change neither); (2) how much the changed element differs from the original (word relatedness, with two ranges of relatedness considered); (3) retention period (immediate test, 2-day delay). Furthermore, each experiment has a large N sample size, so both significant effects as well as null effects are robust and informative.

The results show the benefits of relatedness, but also replicate interference effects in the "change target" condition when the new target is not related to the old target and when test is immediate. This provides reconciliation of some existing seemingly contradictory results on the effect of overlap on memory. Here, the whole range of conditions is mapped to

convincingly show how the direction of the effect can flip across the surface of relatedness values.

Additional strength comes from supporting analyses, such as analyses of learning data, demonstrating that relatedness leads to both better final memory and also faster initial learning.

More broadly, the study informs our understanding of memory integration, demonstrating how interdependence of memory for related information increases with relatedness. Together with a prior study of retroactive interference and facilitation, the results provide new insights into the role of reminding in memory formation.

In summary, this is a highly rigorous body of work that sets a great model for future studies and improves our understanding of memory organization.

Weaknesses:

The evidence for the proactive facilitation driven by relatedness is very convincing. However, in the finer scale results, the continuous relationship between the degree of relatedness and the degree of proactive facilitation/interference is less clear. The relationship was only found in the wider stimulus set, where some pairs were unrelated and other pairs related, and only when GloVe metric for measuring relatedness was used. The absence of a relationship between relatedness and memory in the narrow stimulus set (where all pairs were related to some degree) suggests this could be potentially an all-or-none effect (facilitation for related) rather than a matter of degree. Furthermore, a different metric of relatedness, associative strength AS, did not show the same relationship. The discrepancy between the metrics is not fully resolved. This is less of a problem with interdependence analyses where the results are more converging across narrow and wider range as well as the two metrics.

A smaller weakness, acknowledged by the authors, is generalizability beyond the word set used here. Using a carefully crafted stimulus set and repeating the same word pairings across participants and conditions was important for memorability calculations and some of the other analyses. However, highlighting the inherently noisy item-by-item results, especially in the Osgood-style surface figures, makes it challenging to imagine how the results would generalize to new stimuli, even within the same relatedness ranges as the current stimulus sets.

<https://doi.org/10.7554/eLife.95480.2.sa2>

Reviewer #3 (Public review):

Summary:

Bennion et al. investigate how semantic relatedness proactively benefits the learning of new word pairs. The authors draw predictions from Osgood (1949), which posits that the degree of proactive interference (PI) and proactive facilitation (PF) of previously learned items on to-be-learned items depends on the semantic relationships between the old and new information. In the current study, participants subjects learn a set of word pairs ("supplemental pairs"), followed by a second set of pairs ("base pairs"), in which the cue, target or both words are changed, or the pair was identical. Pairs were drawn from either a narrower or wider stimulus set and were tested after either a 5 minute or 48 hour delay. The results show that semantic relatedness overwhelmingly produces PF and greater memory interdependence between base and supplemental pairs, except in the case of unrelated pairs in a wider stimulus set after a short delay, which produced PI. In their final analyses, the authors compare their current results to previous work from their group studying the analogous retroactive effects of semantic relatedness on memory. These comparisons show

generally similar, if slightly weaker, patterns of results. The authors interpret their results in the framework of recursive reminders (Hintzman, 2011), which posits that the semantic relationships between new and old word pairs promotes reminders of the old information during the learning of the new to-be-learned information. These reminders help to integrate the old and new information and result in additional retrieval practice opportunities that in turn improve later recall.

Strengths:

Overall, I thought that the analyses were thorough and well-thought-out and the results were incredibly well-situated in the literature, especially with the additional clarification and framing that the authors have made in response to reviewer comments. In particular, I found that the large sample size, inclusion of a wide range of semantic relatedness across the two stimulus sets, variable delays and the ability to directly compare the current results to their prior results on the retroactive effects of semantic relatedness were particular strengths of the authors' approach and make this an impressive contribution to the existing literature. I thought that their interpretations and conclusions were mostly reasonable and included appropriate caveats (where applicable).

Weaknesses:

The changes and additional analyses that the authors have made have addressed my concerns about their analyses. Including the additional Fig 1- Supp 1, panel C greatly helps with the interpretability across stimulus sets, and the additional analyses the authors have performed teasing apart whether cue and target similarity separately influence memorability and interdependence seem to support the rest of their conclusions.

<https://doi.org/10.7554/eLife.95480.2.sa1>

Author response:

The following is the authors' response to the original reviews.

Public Reviews:

Reviewer #1 (Public Review):

Summary:

Bennion and colleagues present a careful examination of how an earlier set of memories can either interfere with or facilitate memories formed later. This impressive work is a companion piece to an earlier paper by Antony and colleagues (2022) in which a similar experimental design was used to examine how a later set of memories can either interfere with or facilitate memories formed earlier. This study makes contact with an experimental literature spanning 100 years, which is concerned with the nature of forgetting, and the ways in which memories for particular experiences can interact with other memories. These ideas are fundamental to modern theories of human memory, for example, paired-associate studies like this one are central to the theoretical idea that interference between memories is a much bigger contributor to forgetting than any sort of passive decay.

Strengths:

At the heart of the current investigation is a proposal made by Osgood in the 1940s regarding how paired associates are learned and remembered. In these experiments, one learns a pair of items, A-B (cue-target), and then later learns another pair that is

related in some way, either A'-B (changing the cue, delta-cue), or A-B' (changing the target, delta-target), or A'-B' (changing both, delta-both), where the prime indicates that item has been modified, and may be semantically related to the original item. The authors refer to the critical to-be-remembered pairs as base pairs. Osgood proposed that when the changed item is very different from the original item there will be interference, and when the changed item is similar to the original item there will be facilitation. Osgood proposed a graphical depiction of his theory in which performance was summarized as a surface, with one axis indicating changes to the cue item of a pair and the other indicating changes to the target item, and the surface itself necessary to visualize the consequences of changing both.

In the decades since Osgood's proposal, there have been many studies examining slivers of the proposal, e.g., just changing targets in one experiment, just changing cues in another experiment. Because any pair of experiments uses different methods, this has made it difficult to draw clear conclusions about the effects of particular manipulations.

The current paper is a potential landmark, in that the authors manipulate multiple fundamental experimental characteristics using the same general experimental design. Importantly, they manipulate the semantic relatedness of the changed item to the original item, the delay between the study experience and the test, and which aspect of the pair is changed. Furthermore, they include both a positive control condition (where the exact same pair is studied twice), and a negative control condition (where a pair is only studied once, in the same phase as the critical base pairs). This allows them to determine when the prior learning exhibits an interfering effect relative to the negative control condition and also allows them to determine how close any facilitative effects come to matching the positive control.

The results are interpreted in terms of a set of existing theories, most prominently the memory-for-change framework, which proposes a mechanism (recursive reminding) potentially responsible for the facilitative effects examined here. One of the central results is the finding that a stronger semantic relationship between a base pair and an earlier pair has a facilitative effect on both the rate of learning of the base pair and the durability of the memory for the base pair. This is consistent with the memory-for-change framework, which proposes that this semantic relationship prompts retrieval of the earlier pair, and the two pairs are integrated into a common memory structure that contains information about which pair was studied in which phase of the experiment. When semantic relatedness is lower, they more often show interference effects, with the idea being that competition between the stored memories makes it more difficult to remember the base pair.

This work represents a major methodological and empirical advance for our understanding of paired-associates learning, and it sets a laudably high bar for future work seeking to extend this knowledge further. By manipulating so many factors within one set of experiments, it fills a gap in the prior literature regarding the cognitive validity of an 80-year-old proposal by Osgood. The reader can see where the observed results match Osgood's theory and where they are inconclusive. This gives us insight, for example, into the necessity of including a long delay in one's experiment, to observe potential facilitative effects. This point is theoretically interesting, but it is also a boon for future methodological development, in that it establishes the experimental conditions necessary for examining one or another of these facilitation or interference effects more closely.

We thank the reviewer for their thorough and positive comments -- thank you so much!

Weaknesses:

One minor weakness of the work is that the overarching theoretical framing does not necessarily specify the expected result for each and every one of the many effects examined. For example, with a narrower set of semantic associations being considered (all of which are relatively high associations) and a long delay, varying the semantic relatedness of the target item did not reliably affect the memorability of that pair. However, the same analysis showed a significant effect when the wider set of semantic associations was used. The positive result is consistent with the memory-for-change framework, but the null result isn't clearly informative to the theory. I call this a minor weakness because I think the value of this work will grow with time, as memory researchers and theorists use it as a benchmark for new theory development. For example, the data from these experiments will undoubtedly be used to develop and constrain a new generation of computational models of paired-associates learning.

We thank the reviewer for this constructive critique. We agree that the experiments with a narrower set of semantic associations are less informative; in fact, we thought about removing these experiments from the current study, but given that we found results in the Δ Both condition in Antony et al. (2022) using these stimuli that we did NOT find in the wider set, we thought it was worth including for a thorough comparison. We hope that the analyses combining the two experiment sets (Fig 6-Supp 1) are informative for contextualizing the results in the ‘narrower’ experiments and, as the reviewer notes, for informing future researchers.

Reviewer #2 (Public Review):

Summary:

The study focuses on how relatedness with existing memories affects the formation and retention of new memories. Of core interest were the conditions that determine when prior memories facilitate new learning or interfere with it. Across a set of experiments that varied the degree of relatedness across memories as well as retention interval, the study compellingly shows that relatedness typically leads to proactive facilitation of new learning, with interference only observed under specific conditions and immediate test and being thus an exception rather than a rule.

Strengths:

The study uses a well-established word-pair learning paradigm to study interference and facilitation of overlapping memories. However it goes more in-depth than a typical interference study in the systematic variation of several factors: (1) which elements of an association are overlapping and which are altered (change target, change cue, change both, change neither); (2) how much the changed element differs from the original (word relatedness, with two ranges of relatedness considered); (3) retention period (immediate test, 2-day delay). Furthermore, each experiment has a large N sample size, so both significant effects as well as null effects are robust and informative.

The results show the benefits of relatedness, but also replicate interference effects in the "change target" condition when the new target is not related to the old target and when the test is immediate. This provides a reconciliation of some existing seemingly contradictory results on the effect of overlap on memory. Here, the whole range of conditions is mapped to convincingly show how the direction of the effect can flip across the surface of relatedness values.

Additional strength comes from supporting analyses, such as analyses of learning data, demonstrating that relatedness leads to both better final memory and also faster initial learning.

More broadly, the study informs our understanding of memory integration, demonstrating how the interdependence of memory for related information increases with relatedness. Together with a prior study of retroactive interference and facilitation, the results provide new insights into the role of reminding in memory formation.

In summary, this is a highly rigorous body of work that sets a great model for future studies and improves our understanding of memory organization.

We thank their reviewer for their thorough summary and very supportive words!

Weaknesses:

The evidence for the proactive facilitation driven by relatedness is very convincing. However, in the finer scale results, the continuous relationship between the degree of relatedness and the degree of proactive facilitation/interference is less clear. This could be improved with some additional analyses and/or context and discussion. In the narrower range, the measure used was AS, with values ranging from 0.03-0.98, where even 0.03 still denotes clearly related words (pious - holy). Within this range from "related" to "related a lot", no relationship to the degree of facilitation was found. The wider range results are reported using a different scale, GloVe, with values from -0.14 to 0.95, where the lower end includes unrelated words (sap - laugh). It is possible that any results of facilitation/interference observed in the wider range may be better understood as a somewhat binary effect of relatedness (yes or no) rather than the degree of relatedness, given the results from the narrower condition. These two options could be more explicitly discussed. The report would benefit from providing clearer information about these measures and their range and how they relate to each other (e.g., not a linear transformation). It would be also helpful to know how the values reported on the AS scale would end up if expressed in the GloVe scale (and potentially vice-versa) and how that affects the results. Currently, it is difficult to assess whether the relationship between relatedness and memory is qualitative or quantitative. This is less of a problem with interdependence analyses where the results converge across a narrow and wider range.

We thank the reviewer for this point. While other analyses do show differences across the range of AS values we used, we agree in the case of the memorability analysis in the narrower stimulus set, 48-hr experiment (or combining across the narrower and wider stimulus sets), there could be a stronger influence of binary (yes/no) relatedness. We have now made this point explicitly (p. 26):

“Altogether, these results show that PI can still occur with low relatedness, like in other studies finding PI in Δ Target (A-B, A-D) paradigms (for a review, see Anderson & Neely, 1996), but PF occurs with higher relatedness. In fact, the absence of low relatedness pairs in the narrower stimulus set likely led to the strong overall PF in this condition across all pairs (positive y-intercept in the upper right of Fig 3A). In this particular instance, there may have been a stronger influence of a binary factor (whether they are related or not), though this remains speculative and is not the case for other analyses in our paper.”

Additionally, we have also emphasized that the two relatedness metrics are not linear transforms of each other. Finally, as in addressing both your and reviewer #3's comment below, we now graph relatedness values under a common GloVe metric in Fig 1-Supp 1C (p. 9):

“Please note that GloVe is an entirely different relatedness metric and is not a linear transformation of AS (see Fig 1-Supp 1C for how the two stimulus sets compare using the common GloVe metric).”

A smaller weakness is generalizability beyond the word set used here. Using a carefully crafted stimulus set and repeating the same word pairings across participants and conditions was important for memorability calculations and some of the other analyses. However, highlighting the inherently noisy item-by-item results, especially in the Osgood-style surface figures, makes it challenging to imagine how the results would generalize to new stimuli, even within the same relatedness ranges as the current stimulus sets.

We thank the reviewer for this critique. We have added this caveat in the limitations to suggest that future studies should replicate these general findings with different stimulus sets (p. 28):

“Finally, future studies could ensure these effects are not limited to these stimuli and generalize to other word stimuli in addition to testing other domains (Baek & Papaj, 2024; Holding, 1976).”

Reviewer #3 (Public Review):

Summary:

Bennion et al. investigate how semantic relatedness proactively benefits the learning of new word pairs. The authors draw predictions from Osgood (1949), which posits that the degree of proactive interference (PI) and proactive facilitation (PF) of previously learned items on to-be-learned items depends on the semantic relationships between the old and new information. In the current study, participants learn a set of word pairs ("supplemental pairs"), followed by a second set of pairs ("base pairs"), in which the cue, target, or both words are changed, or the pair is identical. Pairs were drawn from either a narrower or wider stimulus set and were tested after either a 5-minute or 48-hour delay. The results show that semantic relatedness overwhelmingly produces PF and greater memory interdependence between base and supplemental pairs, except in the case of unrelated pairs in a wider stimulus set after a short delay, which produced PI. In their final analyses, the authors compare their current results to previous work from their group studying the analogous retroactive effects of semantic relatedness on memory. These comparisons show generally similar, if slightly weaker, patterns of results. The authors interpret their results in the framework of recursive reminders (Hintzman, 2011), which posits that the semantic relationships between new and old word pairs promote reminders of the old information during the learning of the new to-be-learned information. These reminders help to integrate the old and new information and result in additional retrieval practice opportunities that in turn improve later recall.

Strengths:

Overall, I thought that the analyses were thorough and well-thought-out and the results were incredibly well-situated in the literature. In particular, I found that the large sample size, inclusion of a wide range of semantic relatedness across the two stimulus sets, variable delays, and the ability to directly compare the current results to their prior results on the retroactive effects of semantic relatedness were particular strengths of the authors' approach and make this an impressive contribution to the existing literature. I thought that their interpretations and conclusions were mostly reasonable and included appropriate caveats (where applicable).

We thank the reviewer for this kind, effective summary and highlight of the paper's strengths!

Weaknesses:

Although I found that the paper was very strong overall, I have three main questions and concerns about the analyses.

My first concern lies in the use of the narrow versus wider stimulus sets. I understand why the initial narrow stimulus set was defined using associative similarity (especially in the context of their previous paper on the retroactive effects of semantic similarity), and I also understand their rationale for including an additional wider stimulus set. What I am less clear on, however, is the theoretical justification for separating the datasets. The authors include a section combining them and show in a control analysis that there were no directional effects in the narrow stimulus set. The authors seem to imply in the Discussion that they believe there are global effects of the lower average relatedness on differing patterns of PI vs PF across stimulus sets (lines 549-553), but I wonder if an alternative explanation for some of their conflicting results could be that PI only occurs with pairs of low semantic relatedness between the supplemental and base pair and that because the narrower stimulus set does not include the truly semantically unrelated pairs, there was no evidence of PI.

We agree with the reviewer's interpretation here, and we have now directly stated this in the discussion section (p. 26):

"Altogether, these results show that PI can still occur with low relatedness, like in other studies finding PI in Δ Target (A-B, A-D) paradigms (for a review see, Anderson & Neely, 1996), but PF occurs with higher relatedness. In fact, the absence of low relatedness pairs in the narrower stimulus set likely led to the strong overall PF in this condition across all pairs (positive y-intercept in the upper right of Fig 3A)."

As for the remainder of this concern, please see our response to your elaboration on the critique below.

My next concern comes from the additive change in both measures (change in Cue + change in Target). This measure is simply a measure of overall change, in which a pair where the cue changes a great deal but the target doesn't change is treated equivalently to a pair where the target changes a lot, but the cue does not change at all, which in turn are treated equivalently to a pair where the cue and target both change moderate amounts. Given that the authors speculate that there are different processes occurring with the changes in cue and target and the lack of relationship between cue+target relatedness and memorability, it might be important to tease apart the relative impact of the changes to the different aspects of the pair.

We thank the reviewer for this great point. First, we should clarify that we only added cue and target similarity values in the Δ Both condition, which means that all instances of equivalence relate to non-zero values for both cue and target similarity. However, it is certainly possible cue and target similarity separately influence memorability or interdependence. We have now run this analysis separately for cue and target similarity (but within the Δ Both condition). For memorability, neither cue nor target similarity independently predicted memorability within the Δ Both condition in any of the four main experiments (all $p > 0.23$). Conversely, there were some relationships with interdependence. In the narrower stimulus set, 48-hr delay experiment, both cue and target similarity significantly or marginally predicted base-secondary pair interdependence (Cue: $r = 0.30$, $p = 0.04$; Target: $r = 0.29$, $p = 0.054$). Notably, both survived partial correlation analyses partialing out the other factor (Cue: $r = 0.33$, $p = 0.03$; Target: $r = 0.32$, $p = 0.04$). In the wider stimulus set, 48-hr delay experiment, only target similarity predicted interdependence (Cue: $r = 0.09$, $p = 0.55$; Target: $r = 0.34$, $p = 0.02$), and target similarity also predicted interdependence after partialing out cue similarity ($r = 0.34$, $p = 0.02$). Similarly, in the narrower stimulus set, 5-min

delay experiment, only target similarity predicted interdependence (Cue: $r = 0.01$, $p = 0.93$; Target: $r = 0.41$, $p = 0.005$), and target similarity also predicted interdependence after partialing out cue similarity ($r = 0.42$, $p = 0.005$). Neither predicted interdependence in the wider stimulus set, 5-min delay experiment (Cue: $r = -0.14$, $p = 0.36$; Target: $r = 0.09$, $p = 0.54$). We have opted to leave this out of the paper for now, but we could include it if the reviewer believes it is worthwhile.

Note that we address the multiple regression point raised by the reviewer in the critique below.

Finally, it is unclear to me whether there was any online spell-checking that occurred during the free recall in the learning phase. If there wasn't, I could imagine a case where words might have accidentally received additional retrieval opportunities during learning - take for example, a case where a participant misspelled "razor" as "razer." In this example, they likely still successfully learned the word pair but if there was no spell-checking that occurred during the learning phase, this would not be considered correct, and the participant would have had an additional learning opportunity for that pair.

We did not use online spell checking. We agree that misspellings would be considered successful instances of learning (meaning that for those words, they would essentially have successful retrieval more than once). However, we do not have a reason to think that this would meaningfully differ across conditions, so the main learning results would still hold. We have included this in the Methods (p. 29-30):

“We did not use spell checking during learning, meaning that in some cases pairs could have been essentially retrieved more than once. However, we do not believe this would differ across conditions to affect learning results.”

Recommendations for the authors:

Reviewer #1 (Recommendations For The Authors):

In terms of the framing of the paper, I think the paper would benefit from a clearer explication of the different theories at play in the introductory section. There are a few theories being examined. Memory-for-change is described in most detail in the discussion, it would help to describe it more deliberately in the intro. The authors refer to a PI account, and this is contrasted with the memory-for-change account, but it seems to me that these theories are not mutually exclusive. In the discussion, several theories are mentioned in passing without being named, e.g., I believe the authors are referring to the fan effect when they mention the difference between delta-cue and delta-target conditions. Perhaps this could be addressed with a more detailed account of the theory underlying Osgood's predictions, which I believe arise from an associative account of paired-associates memory. Osgood's work took place when there was a big debate between unlearning and interference. The current work isn't designed to speak directly to that old debate. But it may be possible to develop the theory a bit more in the intro, which would go a long way towards scaffolding the many results for the reader, by giving them a better sense up front of the theoretical implications.

We thank the reviewer for this comment and the nudge to clarify these points. First, we have now made the memory-for-change and reminders accounts more explicit in the introduction, as well as the fact that we are combining the two in forming predictions for the current study (p. 3):

“Conversely, in favor of the PF account, we consider two main, related theories. The first is the importance of “reminders” in memory, which involve reinstating representations from an earlier study phase during later learning (Hintzman, 2011). This idea centers study-phase

retrieval, which involves being able to mentally recall prior information and is usually applied to exact repetitions of the same material (Benjamin & Tullis, 2010; Hintzman et al., 1975; Siegel & Kahana, 2014; Thios & D'Agostino, 1976; Zou et al., 2023). However, reminders can occur upon the presentation of related (but not identical) material and can result in better memory for both prior and new information when memory for the linked events becomes more interdependent (Hintzman, 2011; Hintzman et al., 1975; McKinley et al., 2019; McKinley & Benjamin, 2020; Schlichting & Preston, 2017; Tullis et al., 2014; Wahlheim & Zacks, 2019). The second is the memory-for-change framework, which builds upon these ideas and argues that humans often retrieve prior experiences during new learning, either spontaneously by noticing changes from what was learned previously or by instruction (Jacoby et al., 2015; Jacoby & Wahlheim, 2013). The key advance of this framework is that recollecting changes is necessary for PF, whereas PI occurs without recollection. This framework has been applied to paradigms including stimulus changes, including common paired associate paradigms (e.g., A-B, A-D) that we cover extensively later. Because humans may be more likely to notice and recall prior information when it is more related to new information, these two accounts would predict that semantic relatedness instead promotes successful reminders, which would create PF and interdependence among the traces.”

Second, as the reviewer suggests, we were referring to the fan effect in the discussion, and we have now made that more explicit (p. 26):

“We believe these effects arise from the competing processes of impairments between competing responses at retrieval that have not been integrated versus retrieval benefits when that integration has occurred (which occurs especially often with high target relatedness). These types of competing processes appear operative in various associative learning paradigms such as retrieval-induced forgetting (Anderson & McCulloch, 1999; Carroll et al., 2007), and the fan effect (Moeser, 1979; Reder & Anderson, 1980).”

Finally, our reading of Osgood’s proposal is as an attempt to summarize the qualitative effects of the scattered literature (as of 1949) and did not discuss many theories. For this reason, we generally focus on the directional predictions relating to Osgood’s surface, but we couch it in theories proposed since then.

It strikes me that the advantage seen for items in the retroactive study compared to the proactive study is consistent with classic findings examining spontaneous recovery. These classic studies found that first-learned materials tended to recover to a level above second-learned materials as time passed. This could be consistent with the memory-for-change proposal presented in the text. The memory-for-change proposal provides a potential cognitive mechanism for the effect, here I'm just suggesting a connection that could be made with the spontaneous recovery literature.

We thank the reviewer for this suggestion. Indeed, we agree there is a meaningful point of connection here. We have added the following to the Discussion (p. 27):

“Additionally, these effects partially resemble those on spontaneous recovery, whereby original associations tend to face interference after new, conflicting learning, but slowly recover over time (either absolutely or relative to the new learning) and often eventually eclipse memory for the new information (Barnes & Underwood, 1959; Postman et al., 1969; Wheeler, 1995). In both cases, original associations appear more robust to change over time, though it is unclear whether these similar outcomes stem from similar mechanisms.”

Minor recommendations

Line 89: relative existing -> relative to existing.

Line 132: "line from an unrelated and identical target" -> from an unrelated to identical target (take a look, just needs rephrasing).

Line 340: (e.g. peace-shaverazor) I wasn't clear whether this was a typographical error, or whether the intent was to typographically indicate a unified representation.

Line 383: effects on relatedness -> effects of relatedness.

We think the reviewer for catching these errors. We have fixed them, and for the third comment, we have clarified that we indeed meant to indicate a unified representation (p. 12):

“[e.g., peace-shaverazor (written jointly to emphasize the unification)]”

Page 24: Figure 8. I think the statistical tests in this figure are just being done between the pairs of the same color? Like in the top left panel, delta-cue pro and delta-target retro are adjacent and look equivalent, but there is no n.s. marking for this pair. Could consider keeping the connecting line between the linked conditions and removing the connecting lines that span different conditions.

Indeed, we were only comparing conditions with the same color. We have changed the connecting lines to reflect this.

Page 26 line 612: I think this is the first mention that the reminders account is referred to as the memory-for-change framework, consider mentioning this in the introduction.

Thank you – we have now mentioned this in the introduction.

Lines 627-630. Is this sentence referring to the fan effect? If so it could help the reader to name it explicitly.

We have now named this explicitly.

Reviewer #2 (Recommendations For The Authors):

This is a matter of personal preference, but I would prefer PI and PF spelled out instead of the abbreviations. This was also true for RI and RF which are defined early but then not used for 20 pages before being re-used again. In contrast, the naming of the within-subject conditions was very intuitive.

We appreciate this perspective. However, we prefer to keep the terms PI and PF for the sake of brevity. We now re-introduce terms that do not return until later in the manuscript.

Osgood surface in Figure 1A could be easier to read if slightly reformatted. For example, target and cue relatedness sides are very disproportional and I kept wondering if that was intentional. The z-axis could be slightly more exaggerated so it's easier to see the critical messages in that figure (e.g., flip from + to - effect along the one dimension). The example word pairs were extremely helpful.

Figures 1C and 1D were also very helpful. It would be great if they could be a little bigger as the current version is hard to read.

Figure 1B took a while to decipher and could use a little more anticipation in the body of the text. Any reason to plot the x-axis from high to low on this figure? It is confusing (and not done in the actual results figures). I believe the supplemental GloVe equivalent in the supplement also has a confusing x-axis.

Thank the reviewer for this feedback. We have modified Figure 1A to reduce the disproportionality and accentuate the z-axis changes. We have also made the text in C and D larger. Finally, we have flipped around the x-axis in B and in the supplement.

The description of relatedness values was rather confusing. It is not intuitive to accept that AS values from 0.03-0.96 are "narrow", as that seems to cover almost the whole theoretical range. I do understand that 0.03 is still a value showing relatedness, but more explanation would be helpful. It is also not clear how the GloVe values compare to the AS values. If I am understanding the measures and ranges correctly, the "narrow" condition could also be called "related only" while the "wide" condition could be called "related and unrelated". This is somewhat verbalized but could be clearer. In general, please provide a straightforward way for a reader to explicitly or implicitly compare those conditions, or even plot the "narrow" condition using both AS values and GloVe values so one can really compare narrow and wider conditions comparing apples with apples.

We thank the reviewer for this critique. First, we have now sought to clarify this in the Introduction (p. 11-12):

“Across the first four experiments, we manipulated two factors: range of relatedness among the pairs and retention interval before the final test. The narrower range of relatedness used direct AS between pairs using free association norms, such that all pairs had between 0.03-0.96 association strength. Though this encompasses what appears to be a full range of relatedness values, pairs with even low AS are still related in the context of all possible associations (e.g., pious-holy has AS = 0.03 but would generally be considered related) (Fig 1B). The stimuli using a wider range of relatedness spanned the full range of global vector similarity (Pennington et al., 2014) that included many associations that would truly be considered unrelated (Fig 1-Supp 1A). One can see the range of the wider relatedness values in Fig 1-Supp 1B and comparisons between narrower and wider relatedness values in Fig 1-Supp 1C.”

Additionally, as noted in the text above, we have added a new subfigure to Fig 1-Supp 1 that compares the relatedness values in the narrower and wider stimulus sets using the common GloVe metric.

Considering a relationship other than linear may also be beneficial (e.g., the difference between AS of 0.03 and 0.13 may not be equal to AS of .83 and .93; same with GloVe). I am assuming that AS and GloVe are not linear transforms of each other. Thus, it is not clear whether one should expect a linear (rather than curvilinear or another monotonic) relationship with both of them. It could be as simple as considering rank-order correlation rather than linear correlation, but just wanted to put this out for consideration. The linear approach is still clearly fruitful (e.g., interdependence), but limits further the utility of having both narrow and wide conditions without a straightforward way to compare them.

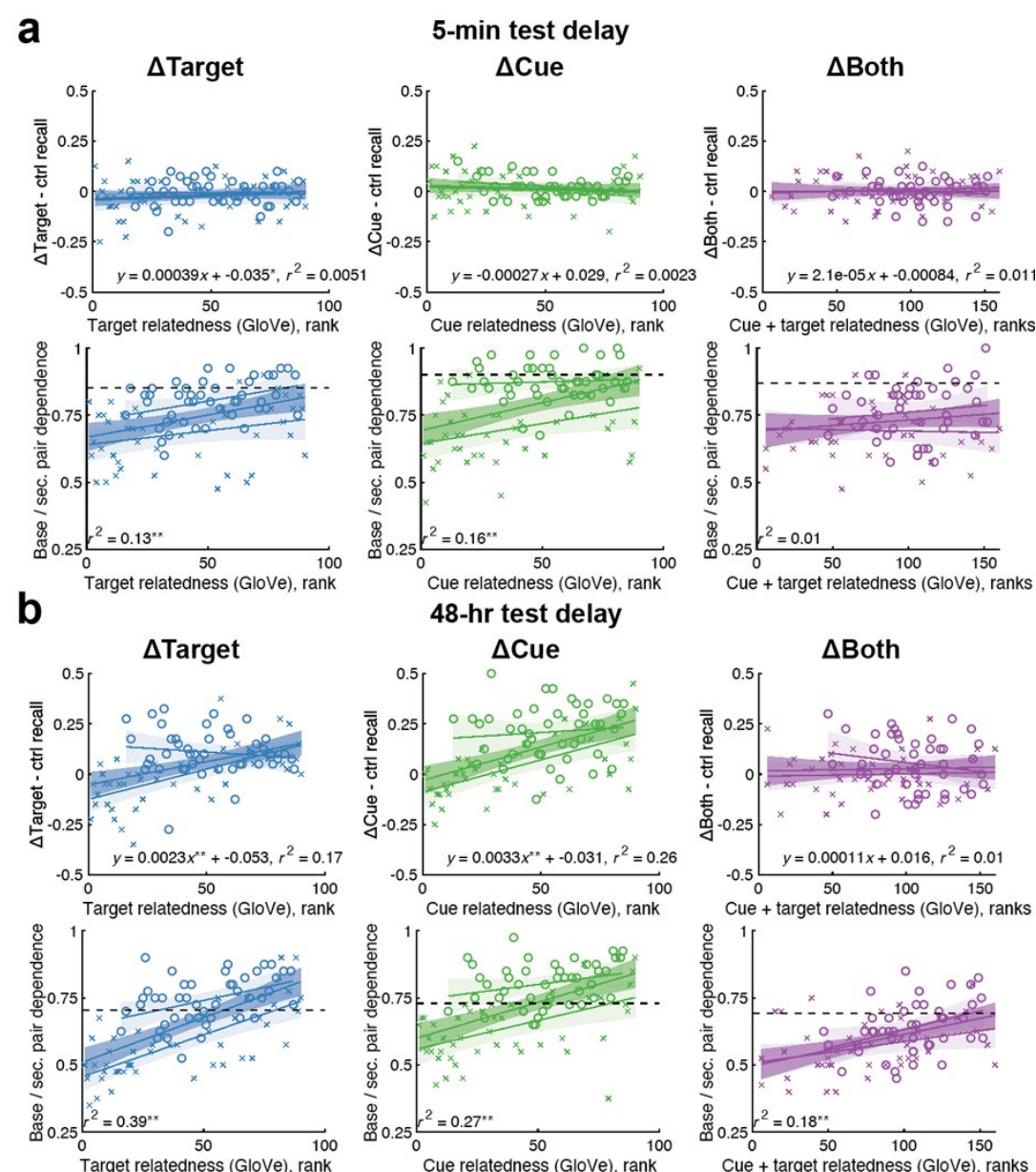
We thank the reviewer for this point. Indeed, AS and GloVe are not linear transforms of each other, but metrics derived from different sources (AS comes from human free associations; GloVe comes from a learned vector space language model). (We noted this in the text and in our response to your above comment.) However, we do have the ability to put all the word pairs into the GloVe metric, which we do in the Results section, “Re-assessing proactive memory and interdependence effects using a common metric”. In this analysis, we used a linear correlation that combined data sets with a similar retention interval and replicated our main findings earlier in the paper (p. 5):

“In the 48-hr delay experiment, correlations between memorability and cue relatedness in the Δ Cue condition [$r^2(44) > 0.29, p < 0.001$] and target relatedness in the Δ Target condition [$r^2(44) = 0.2, p < 0.001$] were significant, whereas cue+target relatedness in the Δ Both condition was not [$r^2(44) = 0.01, p = 0.58$]. In all three conditions, interdependence increased with relatedness [all $r^2(44) > 0.16, p < 0.001$].”

Following the reviewer suggestion to test things out using rank order, we also re-created the combined analysis using rank order based on GloVe values rather than the raw GloVe values. The ranks now span 1-90 (because there were 45 pairs in each of the narrower and wider stimulus sets). All results qualitatively held.

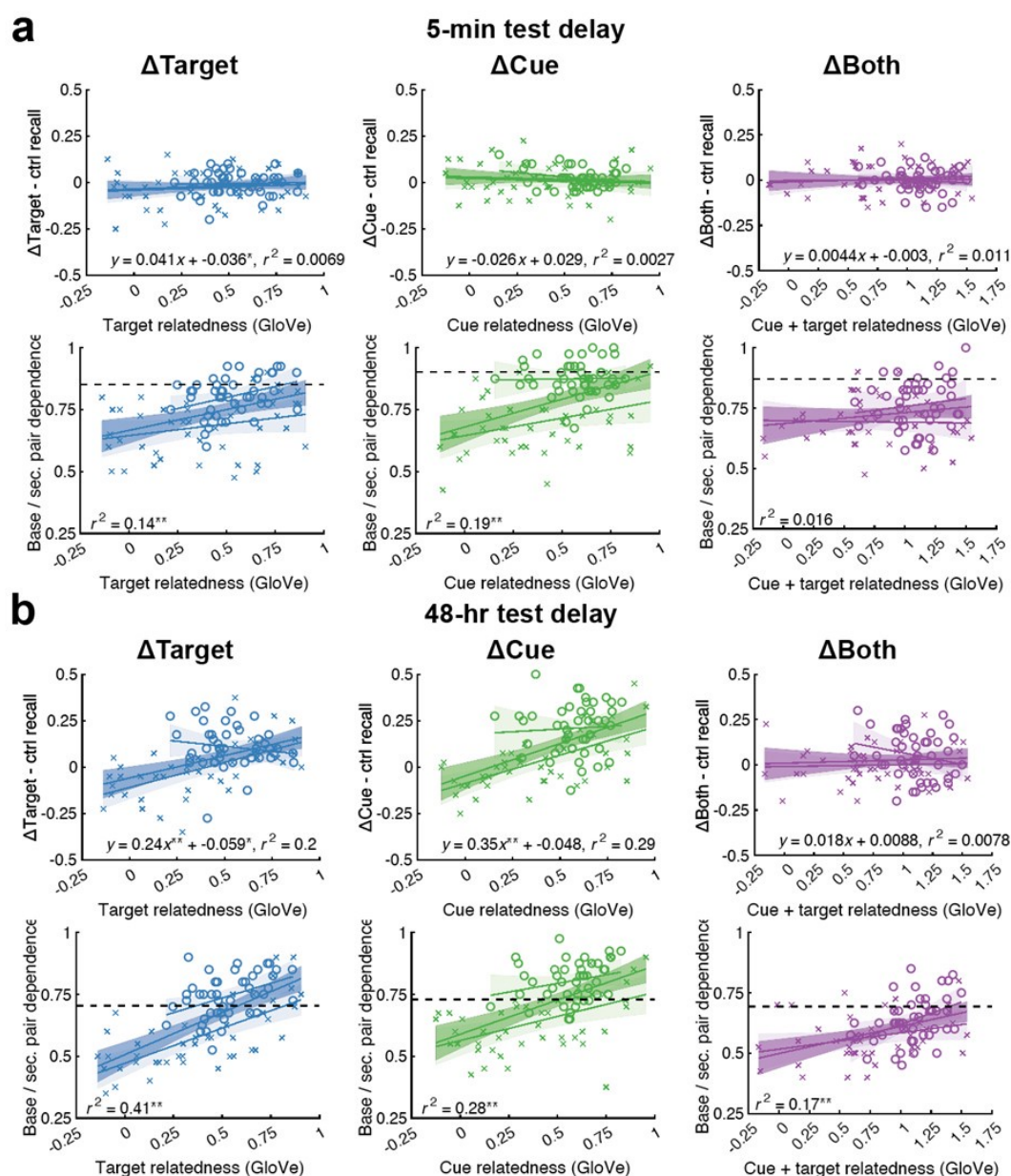
Author response image 1.

Rank order results.



Author response image 2.

And the raw results in Fig 6-Supp 1 (as a reference).



Reviewer #3 (Recommendations For The Authors):

In regards to my first concern, the authors could potentially test whether the stimulus sets are different by specifically looking at pairs from the wider stimulus set that overlap with the range of relatedness from the narrow set and see if they replicate the results from the narrow stimulus set. If the results do not differ, the authors could simplify their results section by collapsing across stimulus sets (as they did in the analyses presented in Figure 6 - Supplementary Figure 1). If the authors opt to keep the stimulus sets separate, it would be helpful to include a version of Figure 1b/Figure 1 - Supplementary Figure 1 where the coverage of the two stimulus sets are plotted on the same figure using GloVe similarity so it is easier to interpret the results.

We have conducted this analysis in two ways, though we note that we will eventually settle upon keeping the stimulus sets separate. First, we examined memorability between the data sets by removing one pair at a time from the wider stimulus set until there was no significant difference ($p > 0.05$). We did this at the long delay because that was more informative for most of our analyses. Even after reducing the wider stimulus set, the narrow stimulus set still had significantly or marginally higher memorability in all three conditions ($p < 0.001$ for Δ Cue; $p < 0.001$ for Δ Target; $p = 0.08$ for Δ Both). We reasoned that this was likely because the AS values still differed (all, $p < 0.001$), which would present a clear way for participants to associate words that may not be as strongly similar in vector space (perhaps due to polysemy for individual words). When we ran the analysis a different way that equated AS, we no longer found significant memorability differences ($p = 0.13$ for Δ Cue; $p = 0.50$ for Δ Target; $p = 0.18$ for Δ Both). However, equating the two data sets in this analysis required us to drop so many pairs to equate the wider stimulus data set (because only a few only had a direct AS connection; there were 3, 5, and 1 pairs kept in the Δ Cue, Δ Target, and Δ Both conditions) that we would prefer not to report this result.

Additionally, we now plot the two stimulus sets on the same plot (Reviewer 2 also suggested this).

*In regards to my second concern, one potential way the authors could disambiguate the effects of change in cue vs change in target might be to run a multiple linear regression with change in Cue, change in Target, and the change in Cue*change in Target interaction (potentially with random effects of subject identity and word pair identity to combine experiments and control for pair memorability/counterbalancing), which has the additional bonus of potentially allowing the authors to include all word pairs in a single model and better describe the Osgood-style spaces in Figure 6.*

This is a very interesting idea. We set this analysis up as the reviewer suggested, using fixed effects for Δ Cue, Δ Target, and Δ Cue* Δ Target, and random effects for subject and word ID. Because we had a binary outcome variable, we used mixed effects logistic regression. For a given pair, if it had the same cue or target, the corresponding change column received a 0, and if it had a different cue or target, it received a graded value (1 - GloVe value between the new and old cue or target). For this analysis, because we designed this analysis to indicate a treatment away from a repeat (as in the No Δ condition, which had no change for either cues and targets), we omitted control items. For items in the Δ Both condition, we initially used positive values in both the Cue and Target columns too, with the multiplied Δ Cue* Δ Target value in its own column. We focused these analyses on the 48-hr delay experiments. In both experiments, running it this way resulted in highly significant negative effects of Δ Cue and Δ Target (both $p < 0.001$), but positive effects of Δ Cue* Δ Target ($p < 0.001$), presumably because after accounting for the negative independent predictions of both Δ Cue and Δ Target, Δ Cue* Δ Target values actually were better than expected.

We thought that those results were a little strange given that generally there did not appear to be interactions with Δ Cue* Δ Target values, and the positive result was simply due to the other predictors in the model. To show that this is the case, we changed the predictors so that items in the Δ Both condition had 0 in Δ Cue and Δ Target columns alongside their Δ Cue* Δ Target value. In this case, all three factors negatively predicted memory (all $p < 0.001$).

We don't necessarily see this second approach as better, partly because it seems clear to us that any direction you go from identity is just hurting memory, and we felt the need to drop the control condition. We next flipped around the analysis to more closely resemble how we ran the other analyses, using similarity instead of distance. Here, identity along any dimension indicated a 1, a change in any part of the pair involved using that pair's GloVe value (rather than the 1 - the GloVe value from above), and the control condition simply had

zeros in all the columns. In this case, if we code the cue and target similarity values as themselves in the Δ Both condition, in both 48-hr experiments, cue and target similarity significantly positively predicted memory (narrower set: cue similarity had $p = 0.006$, target similarity had $p < 0.001$; wider set: both $p < 0.001$) and the interaction term negatively predicted memory ($p < 0.001$ in both). If we code cue and target similarity values as 0s in the Δ Both condition, all three factors tend to be positive (narrower, Cue: $p = 0.11$, Target and Interaction: $p < 0.001$; wider, Cue and Target $p < 0.001$; Interaction: $p = 0.07$).

Ultimately, we would prefer to leave this out of the manuscript in the interest of simplicity and because we largely find that these analyses support our prior conclusions. However, we could include them if the reviewer prefers.

<https://doi.org/10.7554/eLife.95480.2.sa0>