Investigating the Effects of Mediated Associations on Judgment of Learning Reactivity

Nicholas P. Maxwell1 & Mark J. Huff 2

1 Midwestern State University, 2 The University of Southern Mississippi

**Author Note**

Correspondence concerning this article should be addressed to Nicholas P. Maxwell, Department of Psychology, Midwestern State University, 3410 Taft Blvd, Wichita Falls, TX, United States. Study materials, data files, and *R* code used for analyses have been made available via OSF (https://osf.io/mfbnz/).

Abstract

Judgments of learning (JOLs) are often reactive on cue-target pairs. This pattern, however, is moderated by relatedness, as related but not unrelated pairs generally show a memorial benefit when compared to a no-JOL control group. According to Soderstrom et al.’s (2015) cue-strengthening account, JOLs direct attention towards intrinsic cues which aid retrieval. However, reactivity may instead reflect specific processing of cue-target associations, which is applied selectively only when semantic associations are available. The present study tested these accounts using mediated associates (e.g., lion-stripes) which are unrelated to each other and indirectly related through a non-presented mediator (e.g., tiger). Based on a cue-strengthening account, no reactivity would be expected for mediated associates. In contrast, a relational account, predicts a memory benefit whenever pairs are semantically related, even when these pairs are indirectly related through mediators. Overall, reactivity extended to mediated associates in cued recall (Experiment 1) and recognition tests (Experiments 2 and 3). Interestingly, JOL reactivity was consistently found on recognition of non-mediated unrelated pairs (Experiments 2-4). Thus, positive reactivity on related pairs for cued-recall testing likely reflects increased activation of cue-target associations. However, because recognition is based on familiarity cues, reactivity occurs globally for all pair types, regardless of cue-target relations.

Word Count: 201

*Keywords*: Judgments of Learning; Reactivity; Mediated Associates; Cued-Recall; Recognition

Investigating the Effects of Mediated Associations on Judgment of Learning Reactivity

Metamemory, or the processes by which individuals monitor their and adjust their memory abilities, is critical for understanding the learning process. Metamemorial processes are complex, such as deciding whether an information has been sufficiently learned or should be restudied , whether an individual lacks basic knowledge to learn higher-order concepts, or determining whether certain materials may better lend themselves to long-term retention than others (see Nelson & Narens, 1990; Schwartz & Metcalfe, 2017, for review). One method to investigate metamemory processes is to have participants make judgments of learning, which predict whether studied materials will be later remembered. In a typical JOL experiment, participants study items (often cue-target paired associates) while rating their ability to correctly recall the target word in the presence of the cue word on a later test. While JOLs can be elicited via various scales (e.g., Hanczakowski, Zawadzka, Pasek, & Higham, 2013), they are often framed as the likelihood of successfully recalling a pair’s target if prompted by the cue at test using a probability scale (i.e., 0% - 100% ratings). By measuring changes in JOLs, including changes due to material type, delays, and other encoding conditions, researchers can assess metamemory accuracy in a simple comparison process which can be informative for learning (e.g., Koriat & Bjork, 2005; Maxwell & Huff, 2021; Rhodes & Castel, 2008; see Rhodes, 2016).

Despite widespread use by metamemory researchers, early studies often regarded JOLs as having no direct effect on memory performance and instead focused on factors influencing their accuracy (e.g., associative direction, Koriat & Bjork, 2005; Maxwell & Huff, 2021; font size, Rhodes & Castel, 2008; JOL timing, Dunlosky & Nelson, 1994; Nelson & Dunlosky, 1991). A growing body of evidence, however, indicates that JOLs that are provided concurrently with or immediately following encoding, are *reactive* on learning, particularly when participants provide them while studying cue-target pairs (see Double, Birney, & Walker, 2018, for review). Thus, merely providing JOLs at encoding influences participants’ later memory for studied items, likely by directing participants’ attention to aspects of the stimuli that would have otherwise been overlooked (see Ericsson & Simon, 1993). These memory changes can manifest as either memory benefits (i.e., *positive reactivity*) or costs (i.e., *negative reactivity*). Testing for reactivity effects simply involves comparing memory performance for participants making JOLs to a separate group of participants who do not provide JOLs (a no-JOL control task such as silent reading). However, as previous research often focused on factors influencing JOL accuracy, a no-JOL control was often omitted.

For those studies that do investigate JOL reactivity effects, a common test variable is the relatedness of the word pairs. These studies have consistently found that reactivity is moderated by pair relatedness such that providing JOLs generally results in positive reactivity on related pairs (e.g., cat – dog) while not being reactive on unrelated pairs (e.g., cat – sky; Janes, Rivers, & Dunlosky, 2018; Maxwell & Huff, 2022; Soderstrom, Clark, Halamish, & Bjork, 2015; etc.; but see Mitchum, Kelly, & Fox, 2016). Additionally, a meta-analysis conducted by Double, Birney, and Walker (2018) analyzing results from 17 JOL studies showed strong evidence of positive reactivity on related pairs but no evidence of negative reactivity on unrelated pairs. Thus, making JOLs modifies memory for cue-target pairs, specifically by improving recall of related but not unrelated pairs.

To explain the effects of relatedness on JOL reactivity, Soderstrom et al. (2015) proposed a cue-strengthening account. Per this account, JOL reactivity will occur whenever two requirements are met. First, providing JOLs must direct participants’ attention towards specific aspects of the study pairs which might otherwise be overlooked. For instance, when making JOLs, participants use intrinsic properties of the stimuli as indicators of future recall ability (i.e., a perceived pair relatedness; see Koriat, 1997). Because perceptions of pair relatedness are strong predictors of later recall, participants use this cue to inform their JOLs. In doing so, the act of making JOLs strengthens these relatedness cues, but only for related pairs in which these cues are easily perceived at encoding. Because unrelated pairs lack perceptible relatedness cues, cue-strengthening does not occur for this pair type. Second, any cues strengthened at encoding must additionally be easily accessible at test. Therefore, based on a cue strengthening account, positive reactivity would be expected to occur whenever cue-target pairs contain visible relatedness cues and when the test type used is sensitive to these cues.

Soderstrom et al.’s (2015) cue-strengthening account aligns with the general pattern of reactivity observed on cue-target pairs (i.e., positive reactivity on related pairs, no reactivity on unrelated pairs; e.g., Janes et al., 2018; Maxwell & Huff, 2022; Rivers, Janes, & Dunlosky, 2021). Furthermore, previous research supports Soderstrom et al.’s claim that reactivity only occurs when the test emphasizes the cues strengthened at encoding. For example, Myers, Rhodes, and Hausman (2020) compared reactivity effects between cued-recall and free-recall testing (Experiments 1 and 2) and recognition testing (Experiments 3 and 4). Overall, reactivity observed on related pairs with cued-recall testing extended to recognition testing but not free-recall testing in which cues are absent at test. These findings were replicated by Chang and Brainard (2023; Experiment 3), who similarly found no positive reactivity on related pairs when free-recall testing was used. Thus, it is evident that reactivity effects require a match between cues strengthened at encoding and the type of test used to assess memory.

**Cue-Strengthening and Relational Encoding**

As previously noted, several studies have tested the cue-strengthening account by manipulating the type of test participants complete at retrieval (e.g., Myers et al., 2020; Chang & Brainard, 2023). However, fewer studies have assessed the *specific* cues which JOLs are purported to strengthen. Instead, previous studies have often assumed that JOLs enhance recall of cue-target pairs specifically by strengthening relatedness cues, rather than strengthening other intrinsic cues which participants could also potentially utilize when forming their JOLs (e.g., concreteness, item frequency, relatedness, etc., see Dunlosky & Matvey, 2001; Koriat, 1997, for reviews). This is because reactivity studies often use mixed lists of related and unrelated pairs. While cue-target pairs contain several intrinsic cues, relatedness is typically the most salient. Thus, relatedness is easily perceived at encoding, particularly for strong associates. Furthermore, the presence or absence of cue-target relations provides a highly salient marker of difficulty which can be used to inform the magnitude of JOLs (Mueller, Tauber, & Dunlosky, 2013). As a result, the presence of relatedness cues may obscure other intrinsic cues which could potentially be strengthened.

Given the link between relatedness and reactivity, recent work has explored how relatedness contributes to reactivity, often by manipulating pair types and encoding tasks. Maxwell and Huff (2022) investigated relatedness effects on reactivity using three types of paired associates (forward, backward, and symmetrical pairs) and unrelated pairs by comparing JOLs with two additional judgment tasks—judgments of associative memory (JAMs; Maki, 2007; Valentine & Buchanan, 2013) and frequency of co-occurrence judgments—each of which similarly emphasized cue-target relations while removing the metacognitive component associated with JOLs (i.e., neither judgment required making a memory prediction). Positive reactivity occurred on all related pair types, and importantly, JOLs, JAMs and frequency judgments each produced reactivity patterns mirroring JOLs (i.e., positive reactivity on related pairs, no reactivity on unrelated pairs; e.g., Janes et al., 2018; Soderstrom et al., 2015). These findings suggest that JOLs selectively encourage participants to process cue-target relations based on their preexisting relations, benefiting memory for related but not unrelated pairs. Consistent with this account, ratings for JAMs and frequency judgments were moderately-to-strongly correlated with JOLs on related and unrelated pairs (*r*s ≥ .70 and .41, respectively), suggesting that participants in each group based their judgment on similar cues. The authors concluded that JOL reactivity likely reflects the implementation of a relational encoding strategy, which is triggered whenever participants are asked to judge aspects of cue-target pairs pertaining to relatedness. However, because unrelated pairs lack preexisting relationships, they receive no memorial benefit. Thus, providing JOLs at encoding only benefits related cue-target pairs.

Separately, Halamish and Undorf (2023) assessed the influence of pair relatedness on JOL reactivity by comparing reactivity for related, unrelated, and identical cue-target pairs. Importantly, the authors also had participants complete a relatedness judgment task at retrieval in which participants indicated at test whether the cue item was presented alongside a related, unrelated, or identical target. Consistent with previous findings, participants making JOLs demonstrated positive reactivity on related pairs but not unrelated pairs. Furthermore, Halamish and Undorf demonstrated that positive reactivity additionally extended to identical cue-target pairs, providing further evidence that relatedness is a requisite for JOL reactivity. Regarding the relatedness judgments, making JOLs improved accuracy on related pairs, but not identical or unrelated pairs. Thus, findings from Halamish and Undorf provide further evidence that making JOLs causes participants to process cue-target relations but only on related cue-target pairs. Considered alongside findings from Maxwell and Huff (2022), there is converging evidence that JOL reactivity on cue-target pairs occurs via relational processing, with JOLs directing participants to process cue-target associations to a greater extent relative to silent reading.

**The Present Study**

While it is evident that JOL reactivity is contingent on cue-target relations, it remains unclear the extent to which obvious relatedness cues are required for immediate JOLs to trigger positive reactivity on cue-target pairs. For example, previous research investigating whether reactivity patterns on forward associates extend to backward associates (e.g., *mouse* – *cheese* vs. *cheese – mouse*; Maxwell & Huff, 2022; Maxwell & Huff, 2023; Mitchum et al., 2016) has yielded mixed results. Unlike forward associates, backward associates appear related at encoding, yet because the target is not a common response to the cue, relatedness cues utilized at encoding provide little benefit when memory is assessed via cued-recall testing (e.g., Koriat & Bjork, 2005; Maxwell & Huff, 2021). A cue-strengthening account therefore predicts that JOLs would produce no benefit on this pair type. Consistent with this account, Mitchum et al. found no differences in recall of backward associates between a JOL and a control group. However, in two studies, Maxwell and Huff (2022; 2023) reported positive reactivity on backward associates, and furthermore, found that these patterns extended to other judgment types which emphasized cue-target relations (Maxwell & Huff, 2022) and various list constructions (i.e., mixed vs. pure lists, Maxwell & Huff, 2023). Thus, the observation of positive reactivity on backward associates suggests that JOL reactivity may instead reflect the strengthening of implicit cue-target relations versus explicit relatedness cues as posited by the cue-strengthening account.

While the presence of positive reactivity on backward associates suggests that JOL reactivity is based on cue-target associations, a more complete test of this account would be to compare reactivity on forward pairs to a pair type that contains cue-target relations but lacks obvious relatedness cues at encoding. In doing so, this would provide a test of whether reactivity depends upon observable relatedness cues at encoding or if the mere presence of cue-target relations is sufficient to facilitate memory. The present study tested whether reactivity would extend to *mediated paired-associates* (e.g., *lion* – *stripes*). Unlike traditional forward associates, mediated associates are not directly related via traditional measures of word association (e.g., forward association strength; FAS; Nelson, McEvoy, & Schreiber, 2004). Instead, this type of cue-target relation reflects an indirect relationship between concepts, as paired items are directly unrelated yet are related via a non-presented item which links the two concepts (e.g., *lion* – *stripes* is mediated through *tiger*; see Huff & Hutchison, 2011). When participants encounter mediated associates, the non-presented mediator would be expected to be activated via spreading activation (see Balota & Lorch, 1986; Jones 2010). Thus, if JOL reactivity only requires that intrinsic relatedness cues are perceptible at encoding (i.e., cue-strengthening), no reactivity would be expected to occur on this pair type. However, if JOLs also strengthen pre-existing cue-target associations via relational processing, positive reactivity would be expected to occur on mediated associates. Thus, by comparing forward and mediated associates, the present study provided a stronger test of the relational account of JOL reactivity.

**Experiment 1: Mediated Associates and Cued-Recall Testing**

The goal of Experiment 1 was to test the relational and cue-strengthening account of JOL reactivity. In doing so, we compared cued-recall performance between JOL and no-JOL groups using forward and mediated associates and unrelated pairs. Based on previous reactivity studies, we expected any observed reactivity would be moderated by pair type. Specifically, providing JOLs should produce positive reactivity, but only on related pairs. For unrelated pairs, no reactivity was expected. Regarding mediated associates, the relational and cue-strengthening accounts lead to diverging predictions. First, the relational account predicts that JOLs would be reactive on mediated pairs, as providing JOLs should strengthen the pre-existing links between mediated pairs, improving memory for this pair type versus a no-JOL control group. The cue-strengthening account, however, predicts a memory improvement on forward associates but no reactivity on mediated associates. This is because mediated associates lack obvious relatedness cues at encoding, making them appear similar to unrelated pairs. Thus, any positive reactivity on mediated associates would be taken as evidence in favor of a relational account of JOL reactivity.

**Method**

**Participants**

Participant recruitment was based on an a priori power analysis conducted using *G\*Power 3.1* (Faul, Erdfelder, Lang, & Buchner, 2007), which indicated that 106 participants would be required to detect small main effects/interactions or larger (Cohen’s *d* ≥ 0.25). However, because we conducted Experiment 1 online, we extended participant recruitment due to an anticipated increase in response variability. Participant recruitment occurred simultaneously via two platforms. Eighty-four undergraduate psychology students were recruited from the University of Southern Mississippi and completed the study in exchange for partial course credit. An additional 48 participants completed Experiment 1 via Prolific (www.prolific.co) and were compensated at a rate of $4.00/half-hour. To be eligible for participation, participants were required to be native English speakers, and Prolific participants were additionally required to have obtained at least a high school degree or equivalent. For both recruitment sources, participants were randomly assigned to either the JOL or no-JOL encoding groups. Data from 12 participants were omitted due to either low recall rates (i.e., recall < 5%, which suggested participants were distracted at encoding), recall rates exceeding 95% (which suggested cheating at test), or for providing JOLs that consistently anchored on scale extremes (i.e., JOLs of all 0 or 100), which suggested that participants were not following encoding directions. As a result, our final sample contained 120 participants (JOL group *n* = 60; no-JOL group *n* = 60).

**Materials**

To create the stimuli pairs, 60 cue-target word pairs were generated using the University of South Florida Free Association norms. These pairs included 30 forward associates (e.g., mouse – cheese; Nelson, McEvoy, & Schreiber, 2004) and 30 unrelated pairs (e.g., muffin – floor). Next, an additional 30 mediated associates (e.g., beach – box) were taken from Balota and Lorch (1986) and Jones (2010), resulting in a total of 90 cue-target pairs the final stimuli set. The stimuli were then split into two lists, which contained 15 pairs of each type (i.e., forward, unrelated, and mediated). This resulted in 45 cue-target pairs per study list. For both lists, pair types were matched on several lexical variables which could potentially influence recall, including SUBTLEX frequency (Brysbaert & New, 2009), concreteness values derived from the English Lexicon Project (Balota et al., 2007), and word length (see Appendix Table A1 for lexical properties for all pair types). Additionally, forward pairs within each list were matched on FAS (see Table A2). To account for primacy and recency effects, study lists were arranged such that the 45 tested pairs were always preceded and followed by an additional five non-tested buffer items. Thus, each list contained a total of 55 pairs. The final set of cue-target pairs is available via OSF (https://osf.io/mfbnz/). Finally, two cued-recall tests were created by taking the cue items from the 45 tested pairs and replacing the target item with a question mark (e.g., mouse – ?).

**Procedure**

Experiment 1 was administered online using Collector, an open-source program for presenting web-based cognitive psychology experiments (Garcia & Kornell, 2015). Following informed consent, all participants were told that they would be studying a series of cue-target word pairs and that their memory for the target item in each pair would later be tested. After receiving this initial set of instructions, participants in the JOL group were further informed that while studying, they would be asked to rate their likelihood of later recalling the target item if prompted by the cue. Specifically, JOL participants were instructed to provide their ratings using a 0-100 scale and to think of these ratings as the probability of recalling the target item on a memory test. Participants in the no-JOL group were instructed to read each pair silently. Following the instructions, participants were presented with the first study list. For both groups, encoding was self-paced, with participants pressing the ENTER key to move to the next pair. Participants in the JOL group provided their ratings concurrently with study, such that JOLs were elicited while the pair was displayed on the computer screen. List items were randomized for each participant, with the caveat that each list always began and ended with the same five buffer items across participants.

Following the first list, participants completed a two-minute filler task which required them to alphabetize the 50 US states. Once the time-limit was reached, participants immediately began the cued-recall test, which individually presented them with the first word from each of the previously studied cue-target pairs (minus buffer items) in a randomized order. Each item was structured as a cue-target pair, with the missing target item represented by a question mark. Participants were instructed to type the missing word for each pair and were additionally informed that if they could not retrieve an item, they could advance to the next pair. The cued-recall test was self-paced, with participants typing their response and pressing the ENTER key to advance to the next pair. Following this test, participants immediately began the second block, which was structured identically as the first. Thus, participants completed two study/test cycles. To account for potential block effects, block order was counterbalanced across participants. Following completion of the second study/test block, participants were debriefed. For both groups, the experiment took approximately 30 minutes to complete.

**Results**

For all analyses,significance was set at the *p* < .05 level. For all significant analyses of variance (ANOVAs) and t-tests, we report partial eta squared (*ηp*2) and Cohen’s *d* effect size measures, respectively. Finally, for all non-significant main effects, interactions, and post-hoc comparisons, provide a Bayesian estimate regarding the strength of evidence in support of the null hypothesis (see Masson, 2011; Wagenmakers, 2007). This supplemental analysis compares a model in which a significant effect is assumed to a secondary model assuming a null effect and returns an estimated probability of the null hypothesis being retained (termed *p*BIC; Bayesian information criterion). Because *p*BIC is sensitive to sample size, this measure provides additional confidence in null effects reported. Prior to running our analyses, participants’ recall responses were scored in *R* using the *lrd* package, which provides tools for automated scoring of recall data while accounting for potential spelling and grammatical errors (Maxwell, Huff, & Buchanan, 2022). This scoring process followed Maxwell et al.’s (2022) guidelines for processing cued-recall data, such that participant responses were allowed to vary by one character before being counted as incorrect.

Figure 1 plots mean cued-recall rates for participants in the JOL and no-JOL groups. For completeness, all comparisons are reported in the Appendix (Table A3). To test for JOL reactivity effects, data was analyzed using a 2 (Encoding Group: JOL vs. No-JOL) × 3 (Pair Type: Forward vs. Mediated vs. Unrelated) mixed measures ANOVA. Overall, this analysis yielded a significant effect of Encoding Group, *F*(1, 118) = 14.20, *MSE* = 731.18, *η*p2 = .11, such that mean cued-recall rates were higher for participants in the JOL than the no-JOL group (45.29 vs. 34.56, respectively). Additionally, a significant effect of pair type was found, *F*(2, 236) = 778.00, *MSE* = 111.03, *η*p2 = .87, in which recall was highest for forward associates (69.47), followed by mediated associates (33.22), and unrelated pairs (17.08). Post-hoc testing confirmed that all comparisons between pair types differed significantly, *t*s ≥ 6.75, *d*s ≥ 0.74. Importantly, a significant Encoding Group × Pair Type interaction was detected, *F*(2, 236) = 27.07, *MSE* = 111.03, *η*p2 = .19. Starting with forward associates, a robust reactivity effect was detected, such that recall rates for participants making JOLs greatly exceeded participants in the no-JOL group (79.94 vs. 59.00; *t*(118) = 6.73, *SEM* = 3.16, *d* = 1.23). Critically, this positive reactivity pattern extended to mediated associates, as making JOLs similarly facilitated recall of this pair type (38.55 vs. 27.89; *t*(118) = 2.82, *SEM* = 3.82, *d* = 0.51). However, no reactivity was observed on unrelated pairs, as cued-recall rates did not differ between participants in the JOL and no-JOL groups (17.39 vs. 16.67; *t*(118) < 1, *SEM* = 2.81, *p* = .83, *p*BIC = .88).

**Discussion**

Experiment 1 tested the cue-strengthening and relational encoding accounts of JOL reactivity by assessing whether reactivity patterns observed on forward paired associates would extend to mediated associates in cued recall. Based on cue-strengthening account, providing JOLs would be expected to benefit forward but not mediated associates, as this account requires the presence of salient relatedness cues at encoding. A relational encoding account, however, predicts positive reactivity on mediated associates, as based on this account, the additional relational encoding afforded by JOLs should strengthen indirect relations between the cue and target. Overall, we replicated previous research showing that JOLs produce positive reactivity on forward associates but are not reactive on unrelated pairs (e.g., Maxwell & Huff, 2022; Soderstom et al., 2015; Rivers et al., 2021). Importantly, positive reactivity extended to mediated associates, suggesting that the mere presence of cue-target associations, even if these associations are indirect, are sufficient for JOLs to trigger positive reactivity. Thus, our findings in Experiment 1 support the notion that JOLs encourage process of cue-target relations, leading to a memorial benefit for related but not unrelated cue-target pairs.

Because positive reactivity on forward associates extended to mediated associates, Experiment 2 tested whether this pattern would occur when participants were tested via recognition. We selected this test type because Myers et al. (2020; Experiments 3 and 4) showed that reactivity patterns observed with cued-recall testing extended to recognition testing. Thus, our use of recognition testing in Experiment 2 provided a further test of JOL reactivity effects on recognition memory while additionally testing whether reactivity effects observed on mediated associates in the previous experiment would replicate.

**Experiment 2: Mediated Associates and Recognition Testing**

Experiment 2 closely followed the design of Experiment 1 with the exception that participants were tested via recognition rather than via cued-recall testing. We selected this test type given that Myers et al. (2020) found that JOL reactivity effects on cued-recall testing extended to this test type. Thus, our use of recognition testing provided an opportunity to replicate this pattern. Because Myers et al. reported that JOL reactivity patterns observed with cued-recall testing extended to recognition, testing we expected that forward pairs would again demonstrate a positive reactivity pattern, such that providing JOLs at encoding would improve correct recognition relative to the no-JOL group. Additionally, based on the previous experiment, we anticipated that any positive reactivity patterns observed on forward pairs extend to mediated pairs, though again, this effect was expected to be smaller for mediated pairs versus forward pairs. Finally, consistent with previous research using cued-recall or recognition testing (e.g., Janes et al., 2018; Maxwell & Huff, 2022; Myers et al., 2020; Soderstrom et al., 2015), we expected that no reactivity would occur on unrelated pairs.

**Method**

**Participants**

One hundred and thirty-three participants completed Experiment 2. All participants were undergraduate students recruited from either the University of Southern Mississippi (*n* = 77) or Midwestern State University (*n* = 56) who completed the study in exchange for partial course credit. Consistent with Experiment 1, participants were randomly assigned to either the JOL or no-JOL encoding groups. Data screening followed the same procedure outlined in Experiment 1, and data from eight participants were excluded from the final analyses. Thus, our final sample consisted of 125 participants (JOL group *n* = 62, no-JOL group *n* = 63). Our final sample was based on Experiment 1, and a sensitivity analysis conducted with *G\*Power 3.1* confirmed that our sample had sufficient power to detect small main effects and interactions or larger (*d*s ≥ 0.22). All participants were native English speakers.

**Materials and Procedure**

Experiment 2 used the same materials and followed the same general procedure as Experiment 1 with the following exceptions. First, while Experiment 2 used the same word pair study lists, lists were randomly selected to serve as either studied items or distractors (i.e., control items). Thus, unlike Experiment 1, participants only completed one study-test block with study items from the other study-test block serving as distractors. Next, the cued-recall test was replaced with a 90-item old/new recognition test. Following the design of Myers et al. (2020), this test contained all 45 previously studied target items and 45 non-studied distractor items which were presented in a randomized order. Distractors consisted of all 45 target items from the non-studied list. Participants were instructed to indicate whether the presented target item had been previously studied (“old”) or was not presented at encoding (“new”). The recognition test was self-paced, and participants pressed the ENTER key to advance after making their selection. Counterbalanced versions of the study were created which swapped studied items and distractor items. All other aspects of Experiment 2, including our use of self-paced, online testing, were identical to Experiment 1. The full experiment took approximately 20 min to complete.

**Results**

**Analysis of Hits and False Alarms**

Figure 2 (top panel) plots mean hits as functions of encoding group and pair type. For completeness, all comparisons are reported in Table A4. A 2 (Encoding Group: JOL vs. No-JOL) × 3 (Pair Type: Forward vs. Mediated vs. Unrelated) mixed measures ANOVA was used to test for potential JOL reactivity effects. This analysis yielded a significant main effect of Encoding Group, *F*(1, 123) = 22.78, *MSE* = .06, *η*p2 = .16. Collapsed across pair types, hit rates in the JOL group exceeded the no-JOL control (.79 vs. .66, respectively). Additionally, a significant effect of Pair Type emerged, *F*(2, 246) = 16.84, *MSE* = .01, *η*p2 = .12. Across encoding groups, hits were greatest for mediated associates (.77), followed by forward associates (.72), and unrelated pairs (.69). All comparisons differed significantly (*t*s ≥ 2.03, *d*s ≥ .27), except for the comparison between forward associates and unrelated pairs, *t*(248) = 1.60, *SEM* = .02, *p* = .11, *p*bic = .85. The Encoding Group × Pair Type interaction, however, was not reliable, *F*(2, 246) = 1.20, *MSE* = .01, *p* = .30, *p*bic = .99, suggesting no difference in reactivity patterns across pair types. A series of planned post-hoc comparisons confirmed this finding. Hits in the JOL group exceeded hits in the no-JOL group on forward associates (.80 vs. .65; *t*(123) = 4.77, *SEM* = .03, *d* = 0.88), mediated associates (.83 vs. .71; *t*(123) = 4.20, *SEM* = .03, *d* = 0.71), and unrelated pairs (.74 vs. .64; *t*(123) = 3.11, *SEM* = .03, *d* = 0.55). Thus, when participants were tested via recognition, all pair types benefited from the requirement to provide JOLs, regardless of relatedness.

False alarms (i.e., false recognition of lure items) were significantly lower for participants in the JOL group vs the no-JOL conditions (.17 vs .30; *t*(123) = 5.30, *SEM* = .02, *d* = 0.95). However, because lure items were not presented in pairs, they were not yoked to a specific pair direction and could not be compared.

**Signal Detection**

Following the design of Myers et al. (2020), we tested for differences in discriminability (*d′*) and response criterion (*c*) between encoding groups. These indices were computed in *R* using the *psycho* package (Makowski, 2018). Overall, mean *d′* was significantly higher for participants in the JOL group versus the no-JOL group (1.86 vs. 1.03, respectively; *t*(123) = 7.43, *SEM* = .11, *d* = 2.36). However, *c* did not differ between encoding groups, (.06 vs. .07; *t*(123) < 1, *SEM* = .06, *p* = .91, *p*bic = .92).

**Discussion**

Experiment 2 tested whether positive reactivity on mediated associates observed in Experiment 1 would extend to recognition testing. In doing so, we also tested whether previously reported JOL reactivity patterns reported with recognition testing (e.g., Myers et al., 2020) would replicate within this context. Indeed, providing JOLs at study increased hit rates and reduced false alarms on related pairs relative to silent reading, regardless of whether pairs were forward or mediated associates—a pattern that lead to an increase in discriminability in our signal-detection analysis. Additionally, signal detection analyses revealed that making JOLs improved discriminability but not response criterion, suggesting providing JOLs at encoding increased memory discriminability for targets but did not alter participants’ response bias. Thus, our findings on related pairs are consistent with Experiment 1 and provide greater evidence for a relational encoding account of reactivity. Regarding unrelated pairs, positive reactivity was found for this pair type as well. This pattern is inconsistent with both Myers et al. (2020), who showed no reactivity for unrelated pairs on recognition tests and Halamish (2018), who reported positive reactivity for recognition of unrelated pairs, though a related pair type was not included. Thus, both the presence and direction of JOL reactivity effects on recognition remain unclear. Given these discrepancies, Experiment 3 sought to replicate whether the positive reactivity on unrelated pairs observed in the previous experiment would replicate using a new sample.

**Experiment 3: Recognition Testing Replication**

Experiment 3 further investigated whether JOLs would produce positive reactivity on unrelated pairs when participants were tested via recognition. This experiment was designed as a direct replication of Experiment 2. Thus, participants again studied mixed lists of forward, mediated, and unrelated cue-target pairs and either made JOLs at encoding or silently read each pair. Following encoding, participants were again tested via recognition. Overall, we anticipated that making JOLs would improve hit rates on forward and mediated pairs. Furthermore, based on our findings in Experiment 2, we predicted that JOLs would be similarly reactive on unrelated pairs.

**Method**

**Participants**

We recruited 129 participants from Prolific (www.prolific.co) who completed Experiment 3 online at a rate of $4.00/half hour. To be eligible for participation, participants were required to be native English speakers and to have achieved at least a high school degree or equivalent. Data screening followed the same criteria used in the previous experiments, and data from six participants were omitted. As a result, our final sample contained data from 123 participants (JOL group *n* = 61; no-JOL group *n* = 62). A sensitivity analysis conducted with *G\*Power 3.1* confirmed that this sample had sufficient power to detect small main effects and interactions (*d*s ≥ 0.22).

**Materials and Procedure**

All Experiment 3 materials and procedures were identical to Experiment 2.

**Results**

**Analysis of Hits and False Alarms**

Figure 2 (middle panel) displays mean hit rates and false alarms as functions of pair type and encoding group, and all comparisons are reported in Table A4. To test for reactivity effects, hit rates were analyzed using a 2 (Encoding Group: JOL vs. No-JOL) × 3 (Pair Type: Forward vs. Mediated vs. Unrelated) mixed ANOVA. Overall, this analysis yielded a significant main effect of Encoding Group, *F*(1, 121) = 10.77, *MSE* = .06, *η*p2 = .08, as hits in the JOL group exceeded the no-JOL group (.72 vs. .64). Next, a significant effect of Pair Type was detected, *F*(2, 242) = 12.89, *MSE* = .02, *η*p2 = .10. Collapsed across encoding groups, mean hits were highest for mediated associates (.71), followed by forward associates (.69) and unrelated pairs (.63). All comparisons differed (*t*s ≥ 2.53, *d*s ≥ 0.33), except for the comparison between forward and mediated associates which was not reliable, *t* < 1, *SEM* = .02, *p* = .45, *p*bic = .91. Finally, consistent with Experiment 2, the Encoding Group × Pair Type interaction was not reliable, *F*(2, 224) < 1, *MSE* = .01, *p* = .88, *p*bic = .99. However, given our research question, we again separately assessed changes in hit rates between encoding groups as a function of pair type. Starting with forward associates, a positive reactivity effect emerged, as hit rates in the JOL group exceeded the no-JOL group (.74 vs. .65; *t*(121) = 2.60, *SEM* = .03, *d* = 0.51). This pattern similarly extended to mediated associates (.76 vs. .66; *t*(121) = 3.07, *SEM* = .03, *d* = 0.59) and, importantly, unrelated pairs (.67 vs. .60; *t*(121) = 2.41, *SEM* = .03, *d* = 0.39). Thus, the requirement to make JOLs at encoding benefited correct recognition of all pair types, regardless of relatedness.

Finally, like the previous experiment, false alarm rates were lower for participants in the JOL group versus the no-JOL group (.17 vs. .24; ; *t*(121) = 2.97, *SEM* = .02, *d* = 0.54). Taken together, making JOLs improved overall recognition accuracy relative to silent reading.

**Signal Detection**

We similarly assessed changes in discriminability and response criterion as a function of encoding group. Overall, *d′* was greater for participants in the JOL group relative to the no-JOL group (1.65 vs. 1.12; *t*(121) = 5.06, *SEM* = .11, *d* = 1.58). No differences in *c* were detected between the JOL and no-JOL groups (.20 vs. .20; *t*(121) < 1, *SEM* = .06, *p* = .95, *p*bic = .92).

**Discussion**

Consistent with our findings in Experiment 2, making JOLs again improved hit rates across all pair types, regardless of relatedness. [CUE-STRENGTHENING/RELATIONAL SENTENCE HERE] However, given these findings depart from previous research showing that recognition testing demonstrates the traditional reactivity pattern reported with cued-recall testing (i.e., positive reactivity on related pairs, no reactivity on unrelated pairs; Myers et al., 2020), Experiment 4 tested whether these discrepancies in reactivity patterns on unrelated pairs emerged due to differences in items, including our use of mediated associates in the previous experiments. As such, Experiment 4 was designed to provide a closer replication of Myers et al.’s reactivity patterns on recognition testing by only comparing hit rates on forward associates and unrelated pairs.

**Experiment 4: Replication of Myers et al. (2020)**

The goal of Experiment 4 was to provide a closer replication of Myers et al.’s (2020) JOL reactivity patterns by omitting mediated pairs given the study lists used in Myers et al.’s experiments included only forward paired associates and unrelated pairs. In doing so, this provided an additional test of positive reactivity effects on unrelated pairs while also providing an additional opportunity to replicate reactivity effects observed on forward associates in recognition. Based on our findings in Experiments 2 and 3, we anticipated that making JOLs would again produce positive reactivity on both pair types, leading to increased hit rates relative to participants in the no-JOL control group. Finally, because our findings in the previous experiments suggest that making JOLs improves recognition accuracy, we additionally anticipated that a decrease in false alarms for participants in the JOL group vs. the no-JOL group.

**Method**

**Participants**

We recruited 125 participants via Prolific to complete Experiment 4. Like Experiment 3, participants completed the study online and were compensated at a rate of $4.00/half hour. To be eligible for participation, participants were required to meet the same criteria outlined in Experiment 3. Data screening followed the same process used in the previous experiments, which lead to the exclusion of five participants. As such, our final sample consisted of data from 120 participants (JOL group *n* = 61; no-JOL group *n* = 59). A sensitivity test performed with *G\*Power 3.1* confirmed that our final sample had sufficient power to detect small main effects and interactions (*d*s ≥ 0.25).

**Materials and Procedure**

Experiment 4 used the same materials as Experiments 2 and 3 with the following modifications. First, the 15 mediated associates were removed from each of the two study lists. This resulted in each list containing 30 cue-target word pairs (15 forward paired associates and 15 unrelated pairs). Next, the recognition test was reduced from 90 to 60 items to accommodate the removal of mediated associates from the study lists. Like Experiments 2 and 3, this test presented participants with each of the previously studied target items as well as the 30 target items from the non-studied list, which served as distractors. All other aspects of the materials and procedure were identical to Experiments 2 and 3. The total experiment took approximately 15 minutes to complete.

**Results**

Figure 2 (bottom panel) plots mean hits as a function of encoding group and pair type, and all comparisons are available in Table A4. Beginning with hits, a 2 (Encoding Group: JOL vs. No-JOL) × 2 (Pair Type: Forward vs. Unrelated) mixed ANOVA yielded a significant effect of Encoding Group, *F*(1, 118) = 9.28, *MSE* = .06, *η*p2 = .07, as hits for JOL participants exceeded the no-JOL group (.73 vs. .64). Next, a marginal effect of Pair Type was detected, as collapsed across encoding groups, mean hit rates for forward associates numerically exceeded hits for unrelated pairs XX vs. XX; *F*(1, 118) = 3.38, *MSE* = .02, *p* = .07, *p*bic = .67, *η*p2 = .03). Consistent with Experiments 2 and 3, the Encoding Group × Pair Type interaction was not reliable, *F*(1, 118) < 1, *MSE* = .02, *p* = .37, *p*bic = .90, indicating that any reactivity effects did not differ between pair types. To assess reactivity effects on each pair type, we conducted a set of post-hoc *t*-tests, which separately compared hits between the JOL and no-JOL groups on forward and unrelated pairs. Providing JOLs produced positive reactivity on forward pairs, as hits in the JOL group exceeded the no-JOL group on this pair type (.75 vs. .65, *t*(118) = 2.95, *SEM* = .04, *d* = 0.52), a pattern that extended to unrelated pairs (.71 vs. .63, *t*(118) = 2.41, *SEM* = .03, *d* = 0.42). Finally, false alarms were marginally lower for participants making JOLs relative to the no-JOL group (.20 vs. .25, *t*(118) = 1.82, *SEM* = .02, *p* = .07, *p*bic = .67, *d* = 0.40). Thus, the requirement to provide JOLs at encoding again improved correct recognition of all pair types, regardless of relatedness.

**Signal Detection**

Consistent with Experiments 2 and 3, we tested for changes in discriminability and response criterion between encoding groups. Starting with discriminability, mean *d′* for participants in the JOL exceeded participants in the no-JOL group (1.49 vs. 1.09; *t*(118) = 3.83, *SEM* = .11, *d* = 0.69). Regarding response criterion, mean *c* did not differ between the JOL and no-JOL groups (.10 vs. .17; *t*(118) = 1.13, *SEM* = .06, *p* = .26, *p*bic = .85).

**Discussion**

Experiment 4 tested whether positive reactivity on unrelated pairs observed in Experiments 2 and 3 would replicate in the absence of mediated associates. In doing so, Experiment 4 provided an additionaltest of JOL reactivity effects on unrelated pairs, as this design more closely matched Myers et al.’s (2020) experiments assessing JOL reactivity effects with recognition testing. Reactivity patterns observed in the previous experiments were observed in Experiment 4. Thus, making JOLs at encoding benefited recognition of all pair types, regardless of relatedness.

**General Discussion**

Previous work investigating JOL reactivity on cue-target pairs has revealed a consistent pattern on cued-recall tests: Making immediate JOLs at encoding generally improves recall of related but not unrelated pairs. In the present study, we tested the cue-strengthening and relational accounts of reactivity by investigating whether positive reactivity patterns observed on related cue-target pairs (i.e., Janes et al., 2018; Maxwell & Huff, 2022; Soderstrom et al., 2015) would extend to mediated associates. Because mediated associates are not explicitly related, we reasoned that obvious relatedness cues for this pair type would be unavailable at encoding. Thus, like unrelated pairs, any strengthening of relatedness cues that occurs on forward associates would be unable to occur on mediated associates. A cue-strengthening account therefore predicts no reactivity on this pair type. However, if JOL reactivity instead occurs via an associative process, positive reactivity would still be expected to occur on mediated associates, given the underlying relations between cue and target that are inherent to mediated but not unrelated pairs. Thus, our use of mediated associates directly tested the cue-strengthening account’s requirement that JOLs strengthen intrinsic relatedness cues at encoding.

To test these accounts, Experiment 1 first assessed changes in cue-recall performance on forward and mediated paired associates and unrelated pairs between JOL and no-JOL groups of participants. Experiments 2 and 3 then tested whether reactivity on mediated pairs extended to recognition testing. Across experiments, a consistent pattern emerged: Making JOLs produced positive reactivity on forward and mediated pairs, suggesting that the requirement to make JOLs caused participants to engage in relational encoding for these pair types. For unrelated pairs, however, discrepancy was observed. When participants completed a cued-recall test, JOLs were non-reactive, a finding consistent with the broader literature on JOL reactivity (e.g., Double et al, 2017; Janes et al., 2018; Maxwell & Huff, 2022; Soderstrom et al., 2015; etc.). However, contrary to findings reported by Myers et al. (2020), positive reactivity emerged for unrelated pairs when recognition testing was used. This finding was additionally replicated in Experiment 4, which omitted mediated pairs and provided a closer replication of Myers et al.’s design. Finally, a series of signal detection analyses conducted across Experiments 2-4 provided further evidence that making JOLs modified recognition memory, as discriminability was consistently higher for participants making JOLs relative to participants in the no-JOL group. Taken together, making JOLs consistently modified memory for related cue-target pairs, though reactivity patterns on unrelated pairs differed as a function of test type. Thus, it is likely that differences in reactivity on unrelated pairs reflect cued-recall and recognition tests emphasizing different cues at retrieval.

**JOL Reactivity on Mediated Associates**

Our finding that positive reactivity extended to mediated pairs within the context of cued-recall testing is consistent with a relational processing account of JOL reactivity (i.e., Halamish & Undorf, 2023; Maxwell & Huff, 2022). Unlike forward pairs, which contain obvious relatedness cues, the relations between concepts in mediated pairs are not readily apparent at encoding. Cue-strengthening is therefore not likely to occur on this pair type, given the lack of visible relatedness cues. However, because JOLs produce positive reactivity on mediated pairs within this context, positive reactivity observed on related cue-target pairs likely reflects a relational encoding process. Thus, instead of only strengthening discernable relatedness cues used to inform JOLs, we propose that providing JOLs additionally strengthens pre-existing cue-target associations. While these associations are automatically activated when participants encounter related cue-target pairs at encoding (Hutchison, 2003), the additional relational processing afforded by JOLs at encoding results in these associations being strengthened to a greater degree versus silent reading. As a result, providing JOLs benefits memory for related cue-target pairs, regardless of associative direction (i.e., forward vs. backward pairs) or whether pairs are direct or mediated associates, so long as memory is tested using a format in which cue-target relations are beneficial to retrieval (i.e., cued-recall testing).

Previous research is consistent with a relational account of JOL reactivity. For example, Maxwell and Huff (2022) showed that positive reactivity on forward pairs readily extended to backward pairs. Unlike forward pairs, intrinsic relatedness cues for backward pairs are generally unavailable at test and, furthermore, the cue item is a poor predictor of the target (i.e., *card* *– credit* at encoding vs. *card – ?* at test; see Koriat & Bjork, 2005). Similarly, Maxwell & Huff (in press) replicated these findings on backward pairs while also demonstrating that reactivity on backward pairs additionally occurs in the absence of a forward pair comparison as well as in pure lists containing no unrelated pairs. Finally, Halamish and Undorf (2023) found that while identical cue-target pairs incur similar benefits as related pairs, JOLs also improved relatedness judgments of related cue-target pairs (i.e., judging whether a previously presented cue had been paired with a related or unrelated target), providing further evidence that JOLs differentially affect processing of related cue-target relations. Considered alongside the present study, pattern emerges in which JOLs consistently benefit cued recall of related pairs, regardless of pair direction or type of association. Thus, positive reactivity on related pairs reflects a relational encoding process, though further research is needed to test the degree to which associations and cue-strengthening contribute to reactivity.

**Recognition Testing and Unrelated Pairs**

While the primary goal of this study was to test the cue-strengthening and relational accounts of reactivity, our use of cued-recall and recognition testing additionally allowed us to assess potential differences in reactivity based on test type. We initially elected to use recognition testing in Experiments 2, as Myers et al. (2020) demonstrated that reactivity patterns observed with cued-recall testing extended to this test type. However, in Experiment 2, we observed a divergent pattern of reactivity, such that in addition to benefitted forward and mediated associates, making JOLs also benefited recognition of unrelated pairs. Experiments 3 and 4 tested the reliability of this pattern, with each replicating this finding. Thus, contrary to Myers et al., who reported positive reactivity on related but not unrelated pairs when recognition testing was used, making JOLs benefited all pair types when recognition testing was used. Importantly, the classic reactivity pattern reported in the literature (i.e., positive reactivity on related pairs, no reactivity on unrelated pairs) was observed in Experiment 1 when cued-recall testing was used. Thus, these discrepancies in reactivity based on test type likely reflect differences in cues that benefit recollection versus recognition memory.

An obvious explanation is that cued-recall and recognition tests rely on different processes at retrieval. For example, to successfully complete a cued-recall test, participants must recollect the correct target item from memory based on the provided cue. Thus, because this test type is recollection based, participants must rely on specific cues or characteristics of the stimuli to successfully retrieve them. Therefore, additional encoding of cue-target associations afforded by JOLs may be particularly effective at improving recollection processes between the cue and target, benefiting pairs already containing pre-existing relations (i.e., forward pairs and mediated paired associates). Thus, JOLs improve memory for related but not unrelated pairs when memory is assessed via cued-recall testing. However, recognition testing benefits from familiarity-based cues rather than intrinsic cues such as relatedness, which are more beneficial for recollection (Koriat & Goldsmith, 1996; Yonelinas, 2002). Because JOLs produce positive reactivity on all pair types when recognition testing is used, it is likely that JOLs enhance familiarity for studied all studies items, regardless of relatedness. Our findings in Experiments 2 and 3 support this notion, as in addition to improving hits for studied items, false alarms for distractors were reduced for JOL participants. Thus, unrelated pairs demonstrate a memorial benefit, but only when on tests that are more sensitive to familiarity-based retrieval processes.

Taken together, differences in reactivity patterns between cued-recall and recognition testing provide further evidence that JOL reactivity effects are strongly contingent upon test type. For example, when participants study related cue-target pairs, providing JOLs strengthens cue-target associations as well as other salient information which can affect later memory, including perceived relatedness. Importantly, other cues such as familiarity are simultaneously strengthened across all pair types. Thus, in addition to strengthened cue-target associations via relational encoding, traditional cue-strengthening likely also occurs. However, whether strengthened cue-target associations or strengthened intrinsic cues ultimately influence memory is dictated by the type of test being used, with recollection-based tests relying heavily on associations and recognition-based tests relying primarily on familiarity. Thus, while familiarity cues are strengthened for all pair types, unrelated pairs are only benefited via recognition testing in which this cue is beneficial.

To test the associative nature of JOL reactivity, future studies may wish to explore whether JOLs are reactive in other associative tasks beyond recall of cue-target pairs. For example, if JOL reactivity indeed reflects strengthening of cue-target associations, making JOLs may similarly facilitate repetition priming of related but not unrelated cue-target pairs relative to silent reading. Furthermore, given the salience of cue-target associations, future research may wish to test the cue-strengthening account using situations in which relatedness cues are not available, such as having participants make JOLs on rhyming and non-rhyming word pairs that are semantically unrelated. Ultimately, more work will be needed to fully understand the complex interplay between the relational and cue-strengthening accounts of JOL reactivity.

**Conclusion**

In recent years, the reactive effects of immediate JOLs on cue-target pairs have been increasingly documented. While several accounts have been proposed to explain reactivity, Soderstrom et al.’s (2015) cue-strengthening account has received considerable support. However, because this account requires the presence of discernable relatedness cues at encoding, it does not explain positive reactivity on mediated associates. In the present study, we show that mediated associates demonstrate reactivity patterns mirroring forward associates when using cued-recall (Experiment 1) and recognition testing (Experiments 2 and 3). Importantly, in Experiments 2-4, we show that JOLs produce positive reactivity on unrelated cue-target pairs, a novel finding. Thus, our findings suggest that JOL reactivity reflects a combination of cue-strengthening (i.e., perceived relatedness, familiarity, etc.) and strengthened cue-target associations. The present study therefore adds to a growing body of evidence (e.g., Halamish & Undorf, 2023; Maxwell & Huff, 2022) indicating that JOL reactivity on cued-recall reflects a relational process, rather than solely being reliant on cue-strengthening.

**Open Practices Statement**

The data for all experiments have been made available at https://osf.io/mfbnz/. None of the experiments were preregistered.

**References**

Balota, D. A. & Lorch, R. F. (1986). Depth of automatic spreading activation: Mediated priming effects in pronunciation but not in lexical decision. Journal of Experimental Psychology: Learning, Memory, and Cognition, 12(3), 336–345.

Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B., & Treiman, R. (2007). The English lexicon project. *Behavior Research Methods, 39*(3), 445–459.

Brysbaert, M. & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods, 41*, 977–990.

Chang, M. & Brainerd, C. J. (2023). Changed-goal or cue-strengthening? Examining the reactivity of judgments of learning with the dual-retrieval model. *Metacognition and Learning 18*, 183–217.

Double, K. S., Birney, D. P., & Walker, S. A. (2018). A meta-analysis and systematic review of reactivity to judgments of learning. *Memory, 26*(6), 741–750.

Dunlosky, J. & Matvey, G. (2001). Empirical analysis of the intrinsic–extrinsic distinction of judgments of learning (JOLs): Effects of relatedness and serial position on JOLs. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(5), 1180–1191

Dunlosky, J. & Nelson, T. O. (1994). Does the sensitivity of judgments of learning (JOLs) to the effects of various study activities depend on when the JOLs occur? *Journal of Memory and Language, 33*(4), 545-565.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data (Rev. ed.)*. Cambridge, MA: Bradford Books/ MIT Press.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G\*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191.

Garcia, M. & Kornell, N. (2015). Collector [Computer software]. Retrieved April 3rd, 2020 from https://github.com/gikeymarica/Collector

Halamish, V. (2018). Can very small font size enhance memory? *Memory & Cognition, 46*, 979-993.

Halamish, V. & Undorf, M. (2023). Why do judgments of learning modify memory? Evidence from identical pairs and relatedness judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 49*(4), 547–556,

Hanczakowski, M., Zawadzka, K., Pasek, T., & Higham, P. A. (2013). Calibration of metacognitive judgments: Insights from the underconfidence-with-practice effect. *Journal of Memory and Language, 69*, 429–444.

Huff, M. J. & Hutchison, K. A. (2011). The effects of mediated word lists on false recall and recognition. *Memory & Cognition, 39*, 941–953

Hutchison, K. A. (2003). Is semantic priming due to association strength or feature overlap? A microanalytic review. *Psychonomic Bulletin & Review, 10*(4), 785-813.

Janes, J. L., Rivers, M. L., & Dunlosky, J. (2018). The influence of making judgments of learning on memory performance: Positive, negative, or both? *Psychonomic Bulletin & Review, 25*(6), 2356–2364.

Jones, L. L. (2010). Pure mediated priming: A retrospective semantic matching model. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*(1), 135–146.

Koriat, A. (1997). Monitoring one’s own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experiment Psychology: General, 126*(4), 349–370.

Koriat, A. & Bjork, R. A. (2005). Illusions of competence in monitoring one’s knowledge during study. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(2), 187–194.

Koriat, A. & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review, 103*(3), 490–517.

Maki, W. S. (2007). Judgments of associative memory. *Cognitive Psychology, 54*(4), 319–353.

Makowski, D. (2018). The *psycho* package: An efficient and publishing-oriented workflow for psychological science. *Journal of Open Source Software*, *3*(22), 470.

Masson, M. E. J. (2011). A tutorial on a practical Bayesian alternative to null-hypothesis significance testing. *Behavior Research Methods, 43*, 679–690.

Maxwell, N. P. & Huff, M. J. (2021). The deceptive nature of associative word pairs: The effects of associative direction on judgments of learning. *Psychological Research, 85*(4), 1757-1775.

Maxwell, N. P. & Huff, M. J. (2022). Reactivity from judgments of learning is not only due to memory forecasting: Evidence from associative memory and frequency judgments. *Metacognition and Learning, 17*, 589-625.

Maxwell, N. P. & Huff, M. J. (in press). Is discriminability a requirement for reactivity? Comparing the effects of mixed vs. pure list presentations on judgment of learning reactivity*. Memory & Cognition*, 1-16.

Maxwell, N. P., Huff, M. J., & Buchanan, E. M. (2022). The *lrd* package: An *R* package and Shiny application for processing lexical data. *Behavior Research Methods, 54*, 2001-2024.

Mitchum, A. L., Kelley, C. M., & Fox, M. C. (2016). When asking the question changes the ultimate answer: Metamemory judgments change memory. *Journal of Experimental Psychology: General, 145*(2), 200–219.

Mueller, M. L., Tauber, S. K., & Dunlosky, J. (2013). Contributions of beliefs and processing fluence to the effect of relatedness on judgments of learning. *Psychonomic Bulletin & Review, 20*, 378–384.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments, & Computers, 36*(3), 402–407.

Nelson, T. O. & Dunlosky, J. (1991). When people’s judgments of learning (JOLs) are extremely accurate at predicting subsequent recall: The “Delayed-JOL Effect.” *Psychological Science, 2*(4), 267-270.

Nelson, T. O. & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In: *The psychology of learning and motivation*, ed. G. Bower. American Psychologist.

Rhodes, M. G. (2016). Judgments of learning. In J. Dunlosky & S. K. Tauber (Eds.), *The Oxford Handbook of Metamemory* (pp. 65–80). Oxford University Press.

Rhodes, M. G., & Castel, A. D. (2008). Memory predictions are influenced by perceptual information: Evidence for metacognitive illusions. *Journal of Experimental Psychology: General, 137*(4), 615–625.

Rivers, M. L., Janes, J. L., & Dunlosky, J. (2021). Investigating memory reactivity with a within-participant manipulation of judgments of learning: Support for the cue-strengthening hypothesis. *Memory, 29*(10), 1342–1353.

Soderstrom, N. C., Clark, C. T., Halamish, V., & Bjork, E. L. (2015). Judgments of learning as memory modifiers. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*, 553–558.

Wagenmakers, E. (2007). A practical solution to the pervasive problems of *p* values. *Psychonomic Bulletin & Review, 14*, 779–804.

Valentine, K. D. & Buchanan, E. M. JAM-boree: An application of observation oriented modeling to judgments of associative memory. *Journal of Cognitive Psychology, 25*(4), 400–422.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language, 46*(3), 441–517.



*Figure 1:* Mean percent recall as functions of pair type and encoding group. Bars indicate 95% *CI*s.

A picture containing text, diagram, screenshot, rectangle

Description automatically generated

*Figure 2:* Mean proportion of “old” responses in Experiment 2 (top), Experiment 3 (middle), and Experiment 4 (bottom). “New” columns indicate “old” responses to lure items. Bars indicate 95% *CI*s.

**Appendix**

Table A1

*Summary Statistics for Cue and Target Concreteness, Length, and Frequency as a function of pair type.*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Pair Type | Position | Variable | *M* | *SD* |
| Forward | Cue | Concreteness | 5.24 | 0.92 |
|  |  | Length | 5.20 | 1.44 |
|  |  | Frequency | 2.51 | 0.59 |
|  | Target | Concreteness | 5.44 | 0.95 |
|  |  | Length | 5.03 | 1.28 |
|  |  | Frequency | 3.53 | 0.63 |
| Mediated | Cue | Concreteness | 5.82 | 0.87 |
|  |  | Length | 4.97 | 1.30 |
|  |  | Frequency | 3.35 | 0.54 |
|  | Target | Concreteness | 5.52 | 0.95 |
|  |  | Length | 5.03 | 1.10 |
|  |  | Frequency | 3.13 | 0.58 |
| Unrelated | Cue | Concreteness | 4.97 | 1.24 |
|  |  | Length | 5.10 | 1.56 |
|  |  | Frequency | 3.22 | 0.82 |
|  | Target | Concreteness | 5.16 | 1.00 |
|  |  | Length | 5.17 | 1.58 |
|  |  | Frequency | 3.05 | 0.78 |

*Note*: Frequency ratings were derived from SUBLTEX (Brysbaert & New, 2009). Concreteness ratings were derived from the English Lexicon Project (Balota et al., 2007). Values are collapsed across study lists. The full stimuli set has been made available at https://osf.io/mfbnz/.

Table A2

*Associative Strength Summary Statistics for Forward Associates in each Study List*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| List | *M* | *SD* | *Min.* | *Max.* |
| List 1 | 0.445 | 0.234 | 0.141 | 0.808 |
| List 2 | 0.448 | 0.211 | 0.101 | 0.808 |

*Note:* Cells reflect FAS (forward associative strength) values derived from the University of South Florida Free Association Norms (Nelson et al., 2004).

Table A3

*Comparisons of Mean Recall Percentages for each Encoding Group as a function of Pair Type in Experiment 1*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Encoding Task | Pair Type | *M* | *± 95% CI* | F | M |
| JOL | Forward | 79.94 | 3.09 |  |  |
|  | Mediated | 38.56 | 5.21 | 2.45\* |  |
|  | Unrelated | 17.39 | 3.89 | 4.50\* | 1.17\* |
| No-JOL | Forward | 59.00 | 5.26 |  |  |
|  | Mediated | 27.89 | 5.29 | 1.59\* |  |
|  | Unrelated | 16.78 | 3.82 | 2.32\* | 0.61\* |

*Note:* The two left-most columns denote Cohen’s *d* effect sizes for post-hoc comparisons. \* = *p* < .05. F = Forward associates; M = Mediated associates.

Table A4

*Comparisons of Mean Hit Rates for each Encoding Group as a function of Pair Type in Experiments 2-4*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Experiment | Encoding Task | Pair Type | *M* | *± 95% CI* | F | M |
| Ex. 2 | JOL | Forward | .80 | .05 |  |  |
|  |  | Mediated | .83 | .03 | 0.27 |  |
|  |  | Unrelated | .74 | .04 | 0.38\* | 0.64\* |
|  | No-JOL | Forward | .65 | .05 |  |  |
|  |  | Mediated | .71 | .04 | 0.29 |  |
|  |  | Unrelated | .64 | .05 | 0.08 | 0.37\* |
| Ex. 3 | JOL | Forward | .74 | .04 |  |  |
|  |  | Mediated | .76 | .04 | 0.14 |  |
|  |  | Unrelated | .67 | .04 | 0.35 | 0.49\* |
|  | No-JOL | Forward | .65 | .05 |  |  |
|  |  | Mediated | .66 | .04 | 0.06 |  |
|  |  | Unrelated | .60 | .04 | 0.31 | 0.39\* |
| Ex. 4 | JOL | Forward | .75 | .04 |  |  |
|  |  | Unrelated | .71 | .04 | 0.25 |  |
|  | No-Jol | Forward | .65 | .06 |  |  |
|  |  | Unrelated | .63 | .05 | 0.09 |  |

*Note:* The two left-most columns denote Cohen’s *d* effect sizes for post-hoc comparisons. \* = *p* < .05. F = Forward associates; M = Mediated associates.