

The relatedness effect on judgments of learning: A closer look at the contribution of processing fluency

Monika Undorf · Edgar Erdfelder

Published online: 12 November 2014
© Psychonomic Society, Inc. 2014

Abstract The cue-utilization view to judgments of learning (JOLs) assumes that both ease of processing during study and people's beliefs about memory may contribute to people's predictions on the likelihood of remembering recently studied information. However, a recent study (Mueller, Tauber, & Dunlosky, *Psychonomic Bulletin & Review*, 20(2), 378–384, 2013; Experiment 3) found that processing fluency does not contribute to the effect of pair relatedness on JOLs, that is, to higher JOLs for related paired associates as compared to unrelated paired associates. We investigated whether this finding primarily depends on specific aspects of the paired associates employed or on the measure of processing fluency used in the previous study. In our first two experiments, participants therefore studied lists with (a) uniformly high associative strengths versus (b) a wide range of associative strengths. Results showed that processing disfluency—operationalized as number of trials to acquisition in Experiment 1 and as self-paced study time in Experiment 2—partially mediated the effect of relatedness on JOLs for both types of lists. Finally, in Experiment 3, the contribution of processing fluency to the relatedness effect increased with study-test experience. Unlike Mueller et al., we thus found that processing fluency contributes to the relatedness effect on JOLs. These findings are consistent with the assumption that ease of processing is an important basis for JOLs.

Keywords Metamemory · Judgments of learning · Relatedness · Fluency · Beliefs

The question of how people know that they know has been addressed by a large body of research on metacognition. A

general conclusion from this work is that people infer their own cognitive processes from various cues and heuristics. In the domain of judgments of learning (JOLs), that is, people's predictions on the likelihood that they will remember recently studied information at a later time, this idea has been advanced by Koriat's (1997) cue-utilization approach. According to the cue-utilization theory, JOLs may be based on different types of cues. *Intrinsic cues* refer to characteristics that disclose the a priori difficulty of to-be-learned information. *Extrinsic cues* refer to study conditions and to the learner's encoding operations. Both intrinsic and extrinsic cues may affect JOLs either directly through the deliberate application of a rule or a belief about memory, or indirectly through their effect on *mnemonic cues*. Mnemonic cues such as the fluency of encoding during study derive from people's current processing of items and are assumed to give rise directly to a subjective feeling of mastery (Koriat, 1997).

As an example, consider one of the largest and most robust effects in the literature on JOLs: the effect of pair relatedness on JOLs. It is well documented that JOLs are much higher for related paired associates than for unrelated paired associates (for a review, see Mueller, Tauber, & Dunlosky, 2013). Following the cue-utilization approach, it has been suggested that this relatedness effect relies on processing fluency as well as on people's a priori theories about memory. With respect to intrinsic cues, it has been argued that related pairs receive higher JOLs than unrelated pairs, because people deliberately apply the belief that memory performance is better for related pairs than for unrelated pairs (e.g., Mueller et al., 2013; Soderstrom & McCabe, 2011). With respect to mnemonic cues, it has been proposed that related pairs are more fluently encoded than unrelated pairs and thus evoke higher experiences of knowing (e.g., Mueller et al., 2013; Soderstrom & McCabe, 2011).

A large number of studies have examined the cue-utilization approach (Koriat, 1997). Although some studies

M. Undorf (✉) · E. Erdfelder
Department of Psychology, School of Social Sciences, University of
Mannheim, Schloss, Ehrenhof-Ost, 68131 Mannheim, Germany
e-mail: undorf@uni-mannheim.de

suggested that some revision or additional assumptions may be required (e.g., Dunlosky & Matvey, 2001; Jang & Nelson, 2005; Kimball & Metcalfe, 2003), most findings were consistent with its predictions (e.g., Castel, 2008; Fraundorf & Benjamin, 2014; Koriat, Bjork, Sheffer, & Bar, 2004; Kornell, Rhodes, Castel, & Tauber, 2011; Soderstrom & McCabe, 2011). Most important for present purposes, a growing body of evidence suggests that mnemonic cues do indeed affect JOLs (e.g., Benjamin, Bjork, & Schwartz, 1998; Besken & Mulligan, 2013, 2014; Castel, McCabe, & Roediger, 2007; Koriat & Ma'ayan, 2005; Matvey, Dunlosky, & Guttentag, 2001; Susser et al., 2013; Undorf & Erdfelder, 2011, 2013). There is also some evidence that beliefs about memory may influence JOLs (for a review, see Bjork, Dunlosky, & Kornell, 2013). For example, JOLs were affected by the beliefs that forgetting occurs over time (e.g., Koriat et al., 2004) and that studying results in learning (e.g., Kornell & Bjork, 2009; Kornell et al., 2011). Following training, JOLs accurately reflected the belief of a memory advantage for low frequency words in recognition tests (Benjamin, 2003) and the belief that specific targets seem very obvious when presented along with the cue at study but are hard to recall at test (Koriat & Bjork, 2005, 2006). However, all previous studies also showed that ignoring or discounting metacognitive beliefs tends to be the rule rather than the exception (e.g., Ariel, Hines, & Hertzog, 2014; Kornell & Bjork, 2009; Kornell et al., 2011) and that beliefs must be activated to be incorporated into metacognitive judgments (e.g., Ariel et al., 2014; Koriat et al., 2004). In sum, there is evidence that JOLs are based on mnemonic cues such as processing fluency and, to a lesser extent, on people's beliefs about memory (Bjork et al., 2013).

A recent line of research has, however, challenged the idea that JOLs rely on mnemonic cues. From a series of studies, Mueller et al. (2013) concluded that “people's beliefs largely – if not entirely – mediate the substantial effect of pair relatedness on JOLs” (Mueller et al., 2013, p. 383). What is the evidence for this conclusion? Mueller et al. (2013) found relatedness effects not only with JOLs that were made immediately after each pair had been studied, but also with pre-study JOLs, which were elicited prior to studying each pair and thus cannot rely on processing fluency (Experiment 1). A second experiment revealed that pairs for which perceptual fluency was disrupted by presentation in alternating case (e.g., *tOoTh*) gave rise to a relatedness effect of about the same size as did normal lowercase pairs. In Experiment 3, controlling statistically for processing fluency as measured by lexical decision latencies did not significantly reduce the correlation between relatedness and JOLs.

Several aspects of Mueller et al.'s (2013) experiments do, however, suggest that their conclusion about the contribution of processing fluency to the relatedness effect on JOLs may be somewhat premature. As reported by the authors, JOL differences between related and unrelated pairs were reduced with

pre-study JOLs compared to immediate JOLs in their first experiment. This indicates that processing fluency made a significant, albeit small, contribution to the relatedness effect in Experiment 1. Moreover, the contribution of processing fluency to the relatedness effect is probably underestimated in this study, because the only information available with pre-study JOLs, but not with immediate JOLs, was whether pairs were related or unrelated. This may have led to a particularly pronounced influence of beliefs on pre-study JOLs. Experiment 2 by Mueller et al. (2013) clearly shows that the ease with which pairs are perceived—that is, perceptual fluency—does not contribute to the relatedness effect on JOLs. Conceptual fluency, however, supposedly remained unaffected by presenting words in alternating case and thus may have caused the relatedness effect (cf. Mueller et al., 2013).

Hence, Mueller et al.'s (2013) Experiment 3 alone supports the idea that the relatedness effect on JOLs is entirely mediated by beliefs. However, it remains an open question whether findings from this experiment depend critically (a) on the specific set of paired associates or (b) on the specific measure of processing fluency used. Mueller et al. (2013) employed a carefully constructed list of pairs from Rhodes and Castel (2008). Compared with other JOL studies (e.g., Castel et al., 2007; Connor, Dunlosky, & Hertzog, 1997; Hertzog, Sinclair, & Dunlosky, 2010; Koriat & Bjork, 2005), the average association between members of related pairs was strong and related pairs were easily distinguishable from unrelated pairs. This may have increased people's reliance on beliefs about relatedness. To investigate this possibility, two different study lists are used in the first two experiments presented below: The high association list is comparable to the list used by Mueller et al. (2013), whereas the wide range list consisted of pairs with a wider range and a lower mean of associative strengths.

The measure of processing fluency used by Mueller et al. (2013) was response time in a lexical decision task. In this task, each cue was presented in isolation for 1 second, after which the target—a related or an unrelated word or a non-word—appeared on the screen. Both cue and target remained on the screen until participants decided whether the target was a word or a non-word. Because lexical decision latencies have served as measures of processing fluency in few JOL studies so far (Mueller, Dunlosky, Tauber, & Rhodes, 2014),¹ we use two established measures of processing fluency to address its contribution to the relatedness effect in the experiments reported

¹ Lexical decision latencies have been used in two studies on feeling-of-knowing judgments and tip-of-the-tongue feelings (Connor, Balota, & Neely, 1992; Yaniv & Meyer, 1987). Results showed that lexical decisions were faster for items that induced high feeling-of-knowing judgments and tip-of-the-tongue feelings. However, unlike in the study by Mueller et al. (2013), decision latencies were obtained in a separate phase of the experiment rather than included in the judgment phase in these studies.

below: (1) the number of trials to acquisition in Experiment 1 (e.g., Hoffmann-Biencourt, Lockl, Schneider, Ackerman, & Koriat, 2010; Koriat, Ackerman, Lockl, & Schneider, 2009; Koriat, 2008) and (2) self-paced study time (e.g., Castel et al., 2007; Koriat, 2008; Koriat & Ackerman, 2010; Koriat, Ma'ayan, & Nussinson, 2006; Miele, Finn, & Molden, 2011; Undorf & Erdfelder, 2011, 2013) in Experiments 2 and 3.

If Mueller et al.'s (2013, Experiment 3) finding that processing fluency does not contribute to the relatedness effect crucially depends on using lists with uniformly strong associations between the members of related pairs, their null effect should replicate with high association lists, whereas processing fluency should contribute to the relatedness effect with wide range lists. Moreover, if Mueller et al.'s (2013) results crucially depend on their measure of processing fluency, the relatedness effect should be mediated by processing fluency when using standard fluency measures such as the number of trials to acquisition or self-paced study time. Both predictions were tested in Experiments 1 and 2.

Experiment 1

All participants studied and recalled both a high association list and a wide range list in Experiment 1. The order of list presentation was counterbalanced across participants. The number of trials required for a pair to be correctly recalled was used as a measure of processing fluency.

Method

Participants Participants were 42 University of Mannheim undergraduates. Three participants did not finish the experiment because of time constraints. Their data were discarded from all analyses. The remaining participants were randomly assigned to studying the wide range list in either the first ($n = 20$) or the second half of the experiment ($n = 19$).

Materials Target words consisted of 120 German nouns with a mean log frequency of 1.63 ($SD = 0.76$) and a mean number of letters of 5.52 ($SD = 1.38$). Sixty targets were paired with unrelated cues and 60 with related cues. Two sets of related pairs were constructed by pairing each target with two different cues. Cue words from the two sets were equated for log frequency ($M = 1.02$, $SD = 0.61$) and number of letters ($M = 5.67$, $SD = 1.20$). Association values for related pairs ranged between .02 and .75 ($M = .16$, $SD = .17$) in the wide range set and between .41 and .75 ($M = .55$, $SD = .09$) in the high association set (Melinger & Weber, 2006).

We constructed four study lists of 30 unrelated and 30 related pairs. Two of the lists contained related pairs from the wide range set and from the high association set,

respectively. Study lists were comparable with respect to associative strength, word frequency, and number of letters. Two apparently related and two apparently unrelated buffer pairs were placed at the beginning of each list and served as primacy buffers.

Procedure All participants were presented with one wide range list and with one high association list. The experiment consisted of two parts, each containing several study-test-sequences, a JOL phase, a filler task, and a final cued recall test. The two parts were merely replications of each other with a new study list. During each study phase, 64 pairs were presented one by one for 2 s each. A self-paced cued recall test, in which the cues were presented alone and participants were asked to type in the targets, occurred immediately after the last pair. Correctly recalled items were removed from subsequent study-test sequences. Participants were informed of this procedure prior to the second study phase. After having correctly recalled each pair, all pairs were presented once again for 2 s and participants made a self-paced JOL in which they estimated the probability of recalling the target in a final test on a percentage scale (0 % to 100 %). This test was preceded by a 7-min filler task consisting of addition problems. Responses that were very similar to the target word (e.g., *teeth* instead of *tooth*) were scored as correct responses in the final test.

The order of pairs was randomly determined for each participant except that buffer pairs were always presented first, and that pairs presented in the first half of the study phase were part of the first half of the test phase. Half of the participants were presented with a wide range list and half with a high association list in the first half of the experiment. Each of the two wide range lists and each of the two high association lists was presented to one half of the participants in each order condition.

Results and discussion

Means (and standard deviations) of recall performance, number of trials to acquisition, and JOLs can be found in Table 1. Data were submitted to mixed three-way ANOVAs with list order (high association list first, wide range list first) as a between-participants factor and with relatedness (related pairs, unrelated pairs) and list type (high association list, wide range list) as within-participant factors.² Recall performance was higher for related pairs than for unrelated pairs, $F(1, 37) = 32.25$, $p < 0.001$, $\eta_p^2 = .47$. No other effects were significant.

² For this analysis and for all following analyses, unrelated pairs were coded as 0 and related pairs were coded as 1. The pattern of results is identical, however, when actual associative strengths for related pairs are included in the analyses. The same is true for Experiments 2 and 3.

Table 1 Basic descriptive statistics for Experiment 1 and Experiment 2

	High association list				Wide range list			
	Studied first		Studied last		Studied first		Studied last	
	Rel.	Unrel.	Rel.	Unrel.	Rel.	Unrel.	Rel.	Unrel.
Experiment 1								
Recall	96.84 (6.89)	84.56 (18.23)	92.50 (21.95)	74.17 (27.65)	95.33 (7.75)	82.17 (22.30)	93.68 (12.81)	83.86 (23.15)
TTA	1.41 (0.38)	2.36 (0.83)	1.39 (0.37)	2.45 (1.39)	1.69 (0.52)	2.96 (1.60)	1.50 (0.50)	2.15 (0.73)
JOL	83.90 (12.75)	50.54 (22.03)	70.07 (18.95)	48.53 (19.97)	64.36 (15.10)	42.56 (18.03)	78.41 (16.31)	56.76 (19.84)
Experiment 2								
Recall	89.07 (13.62)	45.19 (28.68)	94.17 (7.25)	51.50 (22.57)	88.17 (9.94)	55.17 (22.57)	75.00 (17.69)	43.15 (26.85)
ST	3.84 (2.03)	6.86 (3.57)	2.59 (1.01)	5.45 (1.97)	4.58 (1.93)	7.53 (2.55)	3.03 (1.63)	5.63 (3.19)
JOL	78.81 (12.74)	38.89 (15.39)	78.12 (9.50)	45.55 (20.45)	71.77 (11.79)	47.34 (17.30)	70.95 (18.27)	36.31 (14.25)

Values represent the mean (and standard deviation) of means for recall performance (in percentage correct) and JOLs and the mean (and standard deviation) of medians for number of trials to acquisition and study time. Rel. = related pairs, Unrel. = unrelated pairs, TTA = number of trials to acquisition, JOL = judgments of learning, ST = study time

Mean number of trials to acquisition was lower for related pairs than for unrelated pairs, $F(1, 37) = 50.77$, $p < 0.001$, $\eta_p^2 = .58$, and was lower for the high association list than for the wide range list, $F(1, 37) = 4.45$, $p = 0.042$, $\eta_p^2 = .11$. Significant interactions between list type and list order, $F(1, 37) = 8.43$, $p = .006$, $\eta_p^2 = .19$, and between relatedness, list type, and list order, $F(1, 37) = 5.88$, $p = .020$, $\eta_p^2 = .14$, showed that the influence of relatedness on number of trials to acquisition was most pronounced with the wide range list in the wide range list first condition. No other effects were significant.

JOLs were significantly higher for related pairs than for unrelated pairs, $F(1, 37) = 130.75$, $p < .001$, $\eta_p^2 = .78$. A significant effect of list order showed that JOLs were higher in the high association list first condition, $F(1, 37) = 4.89$, $p = .033$, $\eta_p^2 = .12$. Furthermore, significant interactions between relatedness and list type, $F(1, 37) = 6.99$, $p = .012$, $\eta_p^2 = .16$, and between relatedness, list type, and list order, $F(1, 37) = 7.63$, $p = .009$, $\eta_p^2 = .17$, showed that the relatedness effect was most pronounced for the high association list in the high association list first condition. No other effects were significant.

In order to examine whether the relatedness effect on JOLs was mediated by number of trials to acquisition, we first conducted multilevel regression analyses (cf. Kenny, Korchmaros, & Bolger, 2003; Krull & MacKinnon, 2001) using the R packages lme4 and lmerTest (Bates, Maechler, & Bolker, 2014; Kuznetsova, Brockhoff, & Christensen, 2014; R Core Team, 2013). Two mixed linear models (level 1: items, level 2: participants) with participants as random effects and with relatedness and number of trials to acquisition as fixed effects were fitted separately for the two list types in each order condition. Number of trials to acquisition was regressed on relatedness in the

first model and JOLs were regressed on number of trials to acquisition and relatedness in the second model. Figure 1 shows the direct effects of number of trials to acquisition on JOLs (Panel a) and of relatedness on number of trials to acquisition and JOLs (Panels b and c, respectively). It can be seen that all effects were significant for each list in both list order conditions. Importantly, this is also true for the direct effects of relatedness on JOLs.

Mediation analyses were carried out using the R package mediation (Tingley, Yamamoto, Keele, & Imai, 2013; see also Imai, Keele, & Tingley, 2010). Thus, indirect effects of relatedness on JOLs mediated by trials to acquisition and their 95 % confidence intervals were estimated using Tingley et al.'s (2013) nonparametric bootstrapping procedure with 2,000 bootstrap samples. For the high association list, the indirect effects of relatedness on JOLs were 7.94, 95 % CI [6.57, 9.39], $p < .001$, when studied first and 2.50, [1.56, 3.46], $p < .001$, when studied last. The respective values for the wide range list were 5.69, [4.46, 6.96], $p < .001$, when studied first and 4.51, [3.49, 5.73], $p < .001$, when studied last. The proportions of the total effect of relatedness on JOLs mediated by number of trials to acquisition were 0.24, [0.20, 0.28], $p < .001$, and 0.12, [0.07, 0.16], $p < .001$, for the high association list studied first and last, respectively, and 0.26, [0.20, 0.32], $p < .001$, and 0.21, [0.16, 0.26], $p < .001$, for the wide range list studied first and last, respectively.

For purposes of comparability with Mueller et al.'s (2013) results, we also compared zero-order correlations between relatedness and JOLs with partial correlations between relatedness and JOLs while controlling for number of trials to acquisition. As shown in the Appendix, these analyses led to

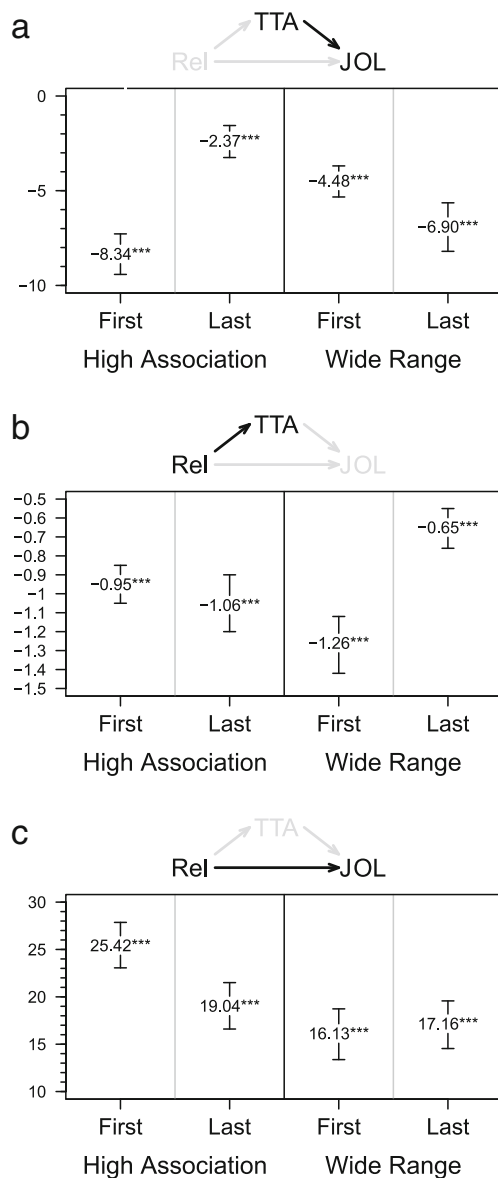


Fig. 1 Unstandardized regression coefficients for the direct effects of number of trials to acquisition (TTA) on judgments of learning (JOLs) (a) and of relatedness on TTA (b) and JOLs (c) in Experiment 1. Coefficients are presented separately for high association lists studied first and last and for wide range lists studied first and last. Error bars represent 95 % confidence intervals. *** $p < .001$

the same conclusions as mediation analyses. It should be noted, however, that comparing zero-order and partial correlations to examine mediation effects is problematic (e.g., Cheung & Lau, 2007; MacKinnon, Lockwood, Hoffman, West, & Sheets, 2002).

In sum, mediation analyses revealed that both the direct effects of relatedness on JOLs and its indirect effects on JOLs mediated by number of trials to acquisition were significant for both lists, regardless of list order. This suggests that processing fluency as measured by number of trials to acquisition partially mediates the

relatedness effect on JOLs. Unexpectedly, the proportion mediated was significantly smaller for the high association list studied last than for the other conditions.

Experiment 2

Results from Experiment 1 showed that processing fluency contributed to the relatedness effect on JOLs with high association lists as well as with wide range lists. Experiment 2 was designed to test whether this pattern of results would also be obtained when processing fluency is operationalized as self-paced study time rather than as number of trials to acquisition.

Method

Participants Participants were 38 undergraduates (26 female). They were randomly assigned to studying the wide range list in either the first ($n = 20$) or the second half of the experiment ($n = 18$).

Materials and procedure Materials were the same as in Experiment 1. The procedure was identical to that of Experiment 1 with the following exceptions. In each half of the experiment, participants were presented with a single study phase. Participants were told to study the pairs for a cued recall test and to choose their study times so that they could recall the second word in the test phase while spending as little study time as possible. Immediately after clicking an onscreen button to indicate that they had finished studying a pair, participants made a self-paced JOL. As soon as the JOL was made, the next item was presented for study. After completing the study phase, participants solved easy mathematical problems as a filler task (90 s).

Results and discussion

Means (and standard deviations) of recall performance, JOLs, and study time are presented in Table 1. A mixed three-way ANOVA revealed that recall performance was higher for related pairs than for unrelated pairs, $F(1, 36) = 136.27$, $p < 0.001$, $\eta_p^2 = .79$, and higher for the high association list than for the wide range list, $F(1, 36) = 6.75$, $p = 0.014$, $\eta_p^2 = .16$. A significant interaction between list type and relatedness revealed that the relatedness effect on recall performance was more pronounced for the high association list than for the wide range list, $F(1, 36) = 14.06$, $p = 0.001$, $\eta_p^2 = .28$. No other effects were significant.

Mean self-paced study time was lower for related pairs than for unrelated pairs, $F(1, 36) = 80.82$, $p < 0.001$, $\eta_p^2 = .69$. A significant interaction between list type and list order indicated that study time was higher for the list that was studied first,

$F(1, 36) = 8.43, p = .006, \eta_p^2 = .19$. No other effects were significant. However, a marginally significant main effect of list type was also observed, $F(1, 36) = 3.93, p = 0.055, \eta_p^2 = .10$.

JOLs were significantly higher for related pairs than for unrelated pairs, $F(1, 36) = 143.50, p < .001, \eta_p^2 = .80$. A significant effect of list type showed that the high association list evoked higher JOLs than the wide range list, $F(1, 36) = 10.47, p = .003, \eta_p^2 = .23$. Furthermore, a significant interaction between relatedness and list type, $F(1, 36) = 6.99, p = .012, \eta_p^2 = .16$, showed that the relatedness effect was more pronounced for the high association list than for the wide range list. No other effects were significant.

Because linear mixed models rest on the assumption of normality, logarithmic transformation of study times was used to move their distribution closer to normality. Figure 2 shows that the direct effects of study time on JOLs (Panel a) and of relatedness on study time and JOLs (Panels b and c, respectively) were significant for both lists regardless of list order. The only exception was the effect of study time on JOLs for the high association list when studied last. As in Experiment 1, direct effects of relatedness on JOLs were significant in all conditions.

Bootstrapped mediation analyses were conducted in the same way as in Experiment 1 and revealed estimates for the indirect effect of relatedness on JOLs mediated by study time of 3.27, 95 % CI [1.91, 4.69], $p < .001$, and 1.11, [−0.32, 2.59], $p = .121$, for the high association list when studied first and last, respectively. The respective estimates for the wide range list were 3.69, [2.55, 4.88], $p < .001$, and 2.91, [1.55, 4.30], $p < .001$. The proportions of the total effect of relatedness on JOLs mediated by study time were 0.08, [0.05, 0.12], $p < .001$, and 0.03, [−0.01, 0.08], $p = .121$, for the high association list studied first and last, respectively, and 0.15, [0.10, 0.20], $p < .001$, and 0.08, [0.04, 0.12], $p < .001$, for the wide range list studied first and last, respectively. Estimates for the indirect effect of relatedness on JOLs mediated by study time and for the proportion mediated were thus significant for the high association list when studied first and for the wide range list in both list order conditions, but not for the high association list studied first the (see Appendix for zero-order correlations between relatedness and JOLs and partial correlations between relatedness and JOLs while controlling for study time).

This pattern of results confirms findings from Experiment 1 in showing that processing fluency as measured by self-paced study time partially mediates the relatedness effect on JOLs. Compared to Experiment 1, the proportion of the relatedness effect mediated by processing

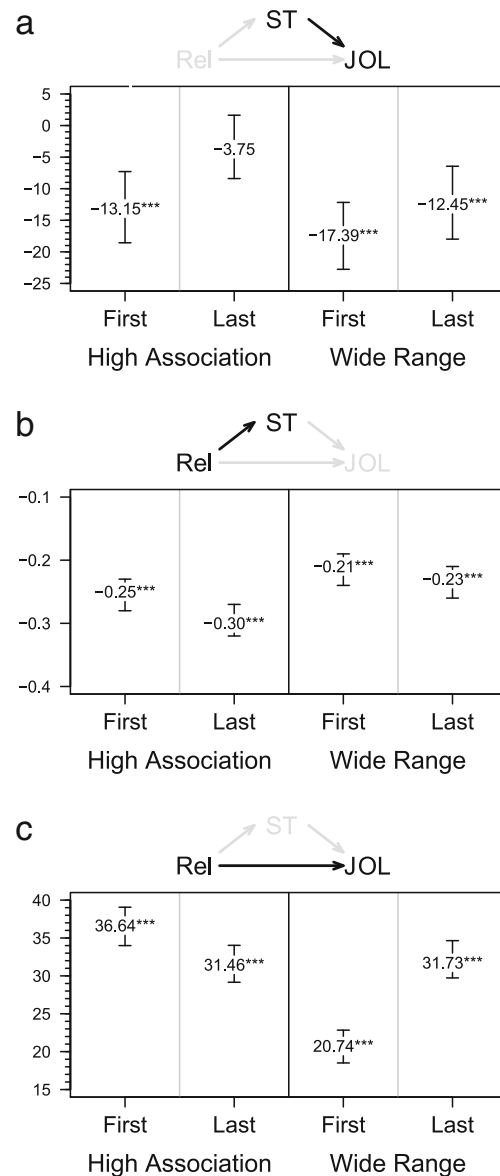


Fig. 2 Unstandardized regression coefficients for the direct effects of self-paced study time (ST) on JOLs (a) and of relatedness on ST (b) and judgments of learning (JOLs) (c) in Experiment 2. Coefficients are presented separately for high association lists studied first and last and for wide range lists studied first and last. Error bars represent 95 % confidence intervals. *** $p < .001$

fluency was somewhat smaller and did not even reach significance for the high association list when studied last.

Experiment 3

Results from the first two experiments showed that processing fluency as measured by number of trials to acquisition and by self-paced study time contributed to the relatedness effect on JOLs both with high association lists and with wide range

lists. This raises the question of whether the contribution of processing fluency to the relatedness effect on JOLs increases with repeated study-test practice. The cue-utilization approach (Koriat, 1997) predicts that with increased practice, JOLs should shift from reliance on a priori theories to reliance on mnemonic cues such as processing fluency. Across repeated presentations, JOLs are assumed to become more sensitive to inter-item differences within the classes of related and unrelated pairs and, thus, more accurate in predicting actual recall performance (see also Ariel & Dunlosky, 2011; Jang, Wallsten, & Huber, 2012; Serra & Ariel, 2014). This may also result in a decrease of the relatedness effect with repeated study-test practice. Consistent with this idea, Koriat (1997, Experiment 2) found that study-test practice increased the predictive accuracy of JOLs for recall and likewise decreased but did not eliminate the correlation between relatedness and JOLs.

Experiment 3 was designed to test whether the contribution of processing fluency to the relatedness effect would increase with repeated practice studying the same materials, as predicted by the cue-utilization approach (Koriat, 1997). Therefore, people were given four study-test trials on the same list of paired associates.

Method

Participants Participants were 36 undergraduates (30 female). Data from 14 participants were discarded from all analyses because they gave JOLs of 100 (13 participants) or 90 (one participant) for all pairs in one or more presentations (two, five, and 13 participants in Presentations 2, 3, and 4, respectively). This left us with a total of 23 participants.

Materials and procedure All participants studied one wide range list from Experiment 1. The procedure was identical to that of Experiment 2 except that participants completed four study-test cycles. Prior to the first study phase, participants were told that they would study and recall the same list of paired associates four times. In each study-test cycle, participants made immediate JOLs regarding cued recall in the next test.

Results and discussion

Table 2 presents means (and standard deviations) of recall performance, study time, and JOLs. A repeated measures ANOVA with relatedness (related pairs, unrelated pairs) and presentation (1, 2, 3, 4) as within-participant factors showed that recall performance was greater for related pairs than for unrelated pairs, $F(1, 21) = 32.56, p < 0.001, \eta_p^2 = .61$, and that recall performance increased with presentation, $F(3, 63) = 120.83, p < 0.001, \eta_p^2 = .85$. A significant interaction showed that the effect of relatedness on recall performance decreased

with presentation, $F(3, 63) = 49.84, p < 0.001, \eta_p^2 = .70$. Planned comparisons confirmed that recall performance was reliably higher for related pairs than for unrelated pairs across all presentations, $t(21) = 11.08, p < 0.001, d = 2.36$; $t(21) = 4.45, p < 0.001, d = 0.95$; $t(21) = 2.70, p = 0.013, d = 0.58$; and $t(21) = 2.39, p = 0.026, d = 0.51$, for Presentations 1, 2, 3, and 4, respectively.

Study time was lower for related pairs than for unrelated pairs, $F(1, 21) = 18.42, p < 0.001, \eta_p^2 = .47$, and decreased with presentation, $F(3, 63) = 23.11, p < 0.001, \eta_p^2 = .52$. A significant interaction revealed that the effect of relatedness on study time decreased with presentation, $F(3, 63) = 4.75, p = 0.005, \eta_p^2 = .19$. Planned comparisons revealed that study time was reliably lower for related pairs than for unrelated pairs in all but the last presentation, $t(21) = 2.69, p = 0.014, d = 0.57$; $t(21) = 4.53, p < 0.001, d = 0.97$; $t(21) = 3.09, p = 0.006, d = 0.65$; and $t(21) = 1.65, p = 0.114, d = 0.35$, for Presentations 1, 2, 3, and 4, respectively.

JOLs were higher for related than for unrelated pairs, $F(1, 21) = 64.10, p < 0.001, \eta_p^2 = .75$, and increased with presentation, $F(3, 63) = 195.69, p < 0.001, \eta_p^2 = .90$. A significant interaction revealed that the relatedness effect on JOLs decreased with presentation, $F(3, 63) = 18.41, p < 0.001, \eta_p^2 = .47$. Planned comparisons confirmed that JOLs were reliably higher for related pairs than for unrelated pairs across all presentations, $t(21) = 11.11, p < 0.001, d = 2.37$; $t(21) = 8.43, p < 0.001, d = 1.80$; $t(21) = 4.40, p < 0.001, d = 0.94$; and $t(21) = 2.82, p = 0.010, d = 0.60$, for Presentations 1, 2, 3, and 4, respectively.

As shown in Fig. 3, direct effects of log-transformed study time on JOLs (Panel a) and of relatedness on log-transformed study time and JOLs (Panels b and c, respectively) were significant across all presentations. It can be seen that the effect of relatedness on both study time and JOLs decreased with presentation, whereas the effect of study time on JOLs increased with presentation.

Mediation analyses revealed that the indirect effect of relatedness on JOLs mediated by study time was significant across all presentations, 0.87, 95 % CI [0.17, 1.63], $p = .018$; 4.35, [3.32, 5.40], $p < .001$; 3.16, [2.28, 4.07], $p < .001$; and 0.96, [0.26, 1.70], $p = .011$, for Presentations 1, 2, 3, and 4, respectively. The same was true for the proportion of the total effect mediated by study time, 0.03, [0.01, 0.05], $p = .018$; 0.17, [0.13, 0.21], $p < .001$; 0.19, [0.14, 0.25], $p < .001$; 0.10, [0.03, 0.17], $p = .011$, for Presentations 1, 2, 3, and 4, respectively. It can be seen that the proportion mediated increased from Presentation 1 to Presentation 2 and remained about the same for the following presentations (see Appendix for zero-order correlations between relatedness and JOLs and partial correlations between relatedness and JOLs while controlling for study time). As predicted by the cue-utilization approach (Koriat, 1997), study-test practice thus decreased the size of the relatedness effect on

Table 2 Basic descriptive statistics for Experiment 3

	Presentation							
	1		2		3		4	
	Rel.	Unrel.	Rel.	Unrel.	Rel.	Unrel.	Rel.	Unrel.
Recall	81.52 (15.52)	43.18 (23.96)	93.79 (9.22)	73.33 (27.02)	95.61 (6.54)	85.45 (21.96)	97.58 (5.06)	90.30 (16.96)
ST	3.83 (2.48)	6.17 (5.25)	2.07 (1.03)	3.07 (1.69)	1.46 (0.70)	1.90 (1.00)	1.29 (0.63)	1.36 (0.63)
JOL	66.83 (13.23)	34.00 (14.11)	80.09 (12.86)	53.99 (20.69)	88.51 (11.24)	72.16 (20.68)	93.73 (7.77)	83.58 (18.39)

Values represent the mean (and standard deviation) of means for recall performance (in percentage correct) and JOLs and the mean (and standard deviation) of medians for study time. Rel. = related pairs, Unrel. = unrelated pairs, ST = study time, JOL = judgments of learning

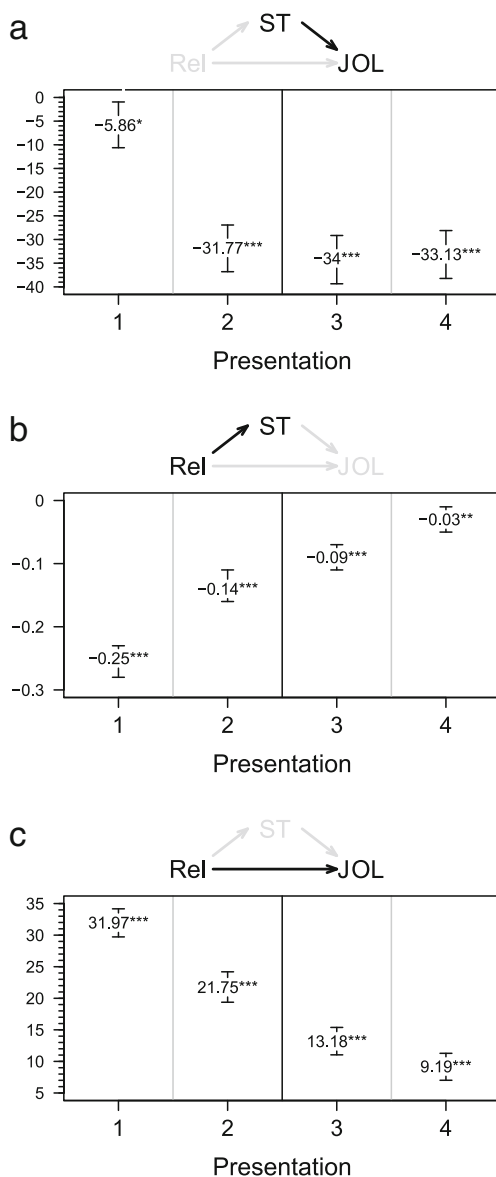


Fig. 3 Unstandardized regression coefficients for the direct effects of self-paced study time (ST) on JOLs (**a**) and of relatedness on ST (**b**) and judgments of learning (JOLs) (**c**) in Experiment 3, presented separately for Presentations 1 to 4. Error bars represent 95 % confidence intervals. *** $p < .001$ ** $p < .01$ * $p < .05$

JOLs, but increased the contribution of processing fluency to the relatedness effect in Experiment 3.

General discussion

The goal of the present experiments was to investigate the contribution of processing fluency to the effect of pair relatedness on JOLs. In three experiments, we found that processing fluency contributed to the relatedness effect. First, Experiment 1 showed that processing fluency as measured by number of trials to acquisition (e.g., Hoffmann-Biencourt et al., 2010; Koriat, 2008) contributed to the effect of relatedness on JOLs both for lists with uniformly strong associations and for lists with a wide range of associative strengths. Mediation analyses revealed that the indirect effect of relatedness on JOLs mediated by number of trials to acquisition explained up to 26 % of the total effect of relatedness on JOLs. The size of this indirect effect was not affected by whether lists were studied in the first or in the second half of the experiment. Converging evidence for these findings was obtained in Experiment 2, which used the same paired associates as in Experiment 1 but a different measure of processing fluency. Experiment 2 revealed that self-paced study time (e.g., Castel et al., 2007; Koriat, 2008; Miele et al., 2011; Undorf & Erdfelder, 2011, 2013) significantly mediated the effect of relatedness on JOLs for wide range lists irrespective of list order, and for high association lists when studied first. Up to 15 % of the relatedness effect was mediated by self-paced study time; the proportion mediated was thus somewhat lower than in Experiment 1. Finally, Experiment 3 demonstrated that repeated practice studying the same pairs increased the contribution of processing fluency to the relatedness effect (e.g., Koriat, 1997). The proportion of the relatedness effect mediated by self-paced study time rose from only 3 % in Presentation 1 to a maximum of 19 % in Presentation 3. Inspection of direct effects revealed that this increase relied on stronger effects of study time on JOLs with repeated study-test experience.

Our findings are consistent with the cue-utilization approach to JOLs (Koriat, 1997). Most importantly, all three experiments showed that the relatedness effect on JOLs is mediated by processing fluency. This is all the more remarkable considering that the methods of eliciting JOLs differed substantially between experiments. Whereas JOLs were made during study in Experiments 2 and 3, JOLs were made in a separate phase following criterion recall in Experiment 1. Thus, our findings provide strong support for the idea that JOLs usually rely on mnemonic cues. Additionally, Experiment 3 provides evidence for the claim that JOLs shift from reliance on beliefs to reliance on mnemonic cues with increased study-test practice: Mediation analyses revealed that study-test practice significantly increased the impact of processing fluency (as indexed by self-paced study time) on JOLs.

The current data are compatible with Experiment 1 by Mueller et al. (2013) in which a more pronounced relatedness effect was found with pre-study JOLs than with immediate JOLs. Assuming that the relatedness effect is mainly driven by conceptual fluency, our results are also consistent with Mueller et al.'s (2013, Experiment 2) finding that perceptual fluency does not contribute to the relatedness effect on JOLs.

However, our findings appear to be at odds with Experiment 3 by Mueller et al. (2013), which showed that controlling statistically for processing fluency as measured by lexical decision latencies did not significantly reduce the correlation between relatedness and JOLs. The first two experiments reported here tested two possible explanations for this finding. One was that processing fluency does not contribute to the relatedness effect for study lists with uniformly strong associations between the members of related pairs, but contributes to the relatedness effect for lists with varying degrees of association. This account, however, conflicts with our finding that the proportion of the relatedness effect mediated by processing fluency was significant for both high association lists and wide range lists in Experiments 1 and 2. The second explanation was that processing fluency, when measured by lexical decision latencies, does not mediate the relatedness effect, whereas commonly used measures of processing fluency do. In line with this idea, both measures of processing fluency used in our studies, that is, number of trials to acquisition (Experiment 1) and self-paced study time (Experiment 2 and Experiment 3), significantly mediated the relatedness effect.

Currently, it is unclear what accounts for the discrepancy between the results for different indicators of processing fluency. Nevertheless, it seems safe to conclude that there are systematic differences between different measures of processing fluency. Further support for this conclusion comes from our observation that self-paced study time contributed less to the relatedness effect than number of trials to acquisition. Considering that previous studies found identical results with

different measures of encoding fluency in general (for a review, see Undorf & Erdfelder, 2011), and with lexical decision time and self-paced study time in particular (Mueller et al., 2014), this is an interesting new finding that deserves scrutiny in future research.

What are the implications of our studies for the contribution of beliefs to the relatedness effect on JOLs? One might argue that any direct effect of relatedness on JOLs reflects the contribution of metacognitive beliefs to the relatedness effect. However, this interpretation is based on the assumption that the contributions of processing fluency are fully captured by indirect effects. Contrary to this view, the present findings suggest that no specific measure captures all aspects of processing fluency. It therefore seems more likely that the direct effect of relatedness on JOLs reflects both beliefs about memory and aspects of processing fluency not grasped by the respective fluency measure. Yet another possibility would be that the direct effect of relatedness on JOLs reflects some currently unknown third factor in addition to beliefs and a remaining portion of processing fluency. To our knowledge, however, no factor beyond processing fluency and beliefs has been proposed or found to influence JOLs so far. It therefore is reasonable to assume that direct effects of relatedness on JOLs reflect a combination of beliefs and some remaining portion of processing fluency (for a similar rationale, see Besken & Mulligan, 2014). Considering that the proportion of the relatedness effect mediated by processing fluency did not exceed 26 % of the total effect in any of our experiments and that metacognitive beliefs have been shown to be involved in the relatedness effect (Mueller et al., 2013), it is likely that beliefs contributed to the relatedness effect in the current studies as well.

Taken together, these results suggest that both processing fluency and beliefs contributed to the relatedness effect on JOLs and that the impact of beliefs on the relatedness effect was reduced with repeated study-test trials on the same list of items.

Concluding comments

Our study demonstrated that processing fluency partially mediates the relatedness effect on JOLs. These results thus do not support the idea that the relatedness effect on JOLs is entirely mediated by beliefs (Mueller et al., 2013). They are, however, also inconsistent with the claim that JOLs are based exclusively on processing fluency (Koriat et al., 2004). Rather, our findings indicate that (a) beliefs and processing fluency both contribute to JOLs and that (b) their relative contributions may change with practice. This conclusion is in close accordance with the cue-utilization approach to JOLs (Koriat, 1997).

Author note Manuscript preparation was supported in part by a grant from the Deutsche Forschungsgemeinschaft (UN 345/1-1) to Monika Undorf.

Appendix

Table 3 shows mean within-person correlations (Pearson r) between JOLs and relatedness, JOLs and number of trials to acquisition, and JOLs and relatedness while controlling for number of trials to acquisition obtained in Experiment 1, and mean within-person correlations between JOLs and relatedness, JOLs and self-paced study time, and JOLs and relatedness while controlling for study time obtained in Experiment 2. Following Mueller et al. (2013), zero-order correlations between JOLs and relatedness and partial correlations between JOLs and relatedness while controlling for processing fluency were submitted to mixed ANOVAs with list order (high association list first, wide range list first) as a between-participants factor and with list type (high association list, wide range list) and type of correlation (zero-order correlation, partial correlation) as within-participant factors.

In Experiment 1, a large main effect of type of correlation revealed that zero-order correlations were higher than partial correlations, $F(1, 37) = 137.31$, $p < .001$, $\eta_p^2 = .79$, and a significant main effect of list type indicated that correlations were higher for the high association list than for the wide

range list, $F(1, 37) = 5.18$, $p = .029$, $\eta_p^2 = .12$. No other effects were significant.

The same pattern of results was found for Experiment 2. Two significant main effects indicated that zero-order correlations were higher than partial correlations, $F(1, 36) = 26.94$, $p < .001$, $\eta_p^2 = .43$, and that correlations were higher for the high association list than for the wide range list, $F(1, 36) = 9.97$, $p = .003$, $\eta_p^2 = .22$. No other effects were significant.

Table 4 presents correlations between JOLs and relatedness, JOLs and self-paced study time, and JOLs and relatedness while controlling for study time in Experiment 3. A mixed ANOVA revealed that zero-order correlations between JOLs and relatedness were higher than partial correlations between JOLs and relatedness while controlling for study time, $F(1, 21) = 16.42$, $p = .001$, $\eta_p^2 = .44$, and a main effect of presentation showed that correlations decreased with presentation, $F(3, 63) = 34.42$, $p < .001$, $\eta_p^2 = .62$. The interaction was also significant, $F(3, 63) = 3.36$, $p = .024$, $\eta_p^2 = .14$. Planned comparisons indicated that zero-order correlations were higher than partial correlations in all but the last presentation, $t(21) = 2.52$, $p = 0.020$, $d = 0.54$; $t(21) = 3.49$, $p = 0.002$, $d = 0.74$; $t(21) = 2.27$, $p = 0.034$, $d = 0.48$; $t(21) = 0.65$, $p = 0.524$, $d = 0.14$, for Presentations 1, 2, 3, and 4, respectively.

Table 3 Correlations between JOLs and relatedness, JOLs and number of trials to acquisition, and JOLs and relatedness controlling for number of trials to acquisition in Experiment 1, and between JOLs and

relatedness, JOLs and study time, and JOLs and relatedness controlling for study time in Experiment 2

	High association list				Wide range list			
	Studied first		Studied last		Studied first		Studied last	
	M	SD	M	SD	M	SD	M	SD
Experiment 1								
$r_{\text{Rel JOL}}$.64***	.25	.48***	.23	.45***	.24	.51***	.18
$r_{\text{TTA JOL}}$	-.55***	.18	-.41***	.18	-.44***	.18	-.49***	.17
$r_{\text{Rel JOL.TTA}}$.54***	.25	.39***	.24	.34***	.24	.41***	.21
Experiment 2								
$r_{\text{Rel JOL}}$.74***	.23	.66***	.34	.56***	.30	.66***	.23
$r_{\text{ST JOL}}$	-.47***	.26	-.36***	.28	-.38***	.27	-.37***	.29
$r_{\text{Rel JOL.ST}}$.67***	.23	.62***	.31	.48***	.29	.57***	.25

Rel = relatedness, JOL = judgments of learning, TTA = number of trials to acquisition, ST = study time

*** $p < .001$

Table 4 Correlations between JOLs and relatedness, JOLs and study time, and JOLs and relatedness controlling for study time in Experiment 3

	Presentation							
	1		2		3		4	
	M	SD	M	SD	M	SD	M	SD
$r_{\text{Rel JOL}}$	0.68***	0.15	0.51***	0.16	0.36***	0.23	0.20***	0.23
$r_{\text{ST JOL}}$	-0.27***	0.19	-0.48***	0.26	-0.40***	0.26	-0.27***	0.28
$r_{\text{Rel JOL.ST}}$	0.65***	0.16	0.44***	0.19	0.33***	0.22	0.19***	0.21

Rel = relatedness, JOL = judgments of learning, ST = study time

*** $p < .001$

In sum, we found significant decreases in the relatedness-JOL correlation when statistically controlling for effects of processing fluency as measured by number of trials to acquisition (Experiments 1) or self-paced study time (Experiments 2 and 3). These results are consistent with findings from bootstrapped mediation analyses (see main text for details) but inconsistent with results from Mueller et al. (2013, Experiment 3) who used lexical decision times as indicators of fluency.

References

- Ariel, R., & Dunlosky, J. (2011). The sensitivity of judgment-of-learning resolution to past test performance, new learning, and forgetting. *Memory & Cognition*, 39(1), 171–184. doi:10.3758/s13421-010-0002-y
- Ariel, R., Hines, J. C., & Hertzog, C. (2014). Test framing generates a stability bias for predictions of learning by causing people to discount their learning beliefs. *Journal of Memory and Language*, 75, 181–198. doi:10.1016/j.jml.2014.06.003
- Bates, D., Maechler, M., & Bolker, B. (2014). lme4: Linear mixed-effects models using Eigen and R syntax (R package version 1.1-6) [Software]. Retrieved from <http://cran.r-project.org/package=lme4>
- Benjamin, A. S. (2003). Predicting and postdicting the effects of word frequency on memory. *Memory & Cognition*, 31(2), 297–305. doi:10.3758/BF03194388
- Benjamin, A. S., Bjork, R. A., & Schwartz, B. L. (1998). The mismeasure of memory: When retrieval fluency is misleading as a metamnemonic index. *Journal of Experimental Psychology: General*, 127(1), 55–68. doi:10.1037/0096-3445.127.1.55
- Besken, M., & Mulligan, N. W. (2013). Easily perceived, easily remembered? Perceptual interference produces a double dissociation between metamemory and memory performance. *Memory & Cognition*, 41(6), 897–903. doi:10.3758/s13421-013-0307-8
- Besken, M., & Mulligan, N. W. (2014). Perceptual fluency, auditory generation, and metamemory: Analyzing the perceptual fluency hypothesis in the auditory modality. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40(2), 429–440. doi:10.1037/a0034407
- Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology*, 64, 417–444. doi:10.1146/annurev-psych-113011-143823
- Castel, A. D. (2008). Metacognition and learning about primacy and recency effects in free recall: The utilization of intrinsic and extrinsic cues when making judgments of learning. *Memory & Cognition*, 36(2), 429–437. doi:10.3758/MC.36.2.429
- Castel, A. D., McCabe, D. P., & Roediger, H. L. (2007). Illusions of competence and overestimation of associative memory for identical items: Evidence from judgments of learning. *Psychonomic Bulletin & Review*, 14(1), 107–111. doi:10.3758/BF03194036
- Cheung, G. W., & Lau, R. S. (2007). Testing mediation and suppression effects of latent variables: Bootstrapping with structural equation models. *Organizational Research Methods*, 11(2), 296–325. doi:10.1177/1094428107300343
- Connor, L. T., Balota, D. A., & Neely, J. H. (1992). On the relation between feeling of knowing and lexical decision: Persistent sub-threshold activation or topic familiarity? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18(3), 544–554. doi:10.1037/0278-7393.18.3.544
- Connor, L. T., Dunlosky, J., & Hertzog, C. (1997). Age-related differences in absolute but not relative metamemory accuracy. *Psychology and Aging*, 12(1), 50–71. doi:10.1037/0882-7974.12.1.50
- 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. doi:10.1037/0278-7393.27.5.1180
- Fraundorf, S. H., & Benjamin, A. S. (2014). Knowing the crowd within: Metacognitive limits on combining multiple judgments. *Journal of Memory and Language*, 71(1), 17–38. doi:10.1016/j.jml.2013.10.002

- Hertzog, C., Sinclair, S. M., & Dunlosky, J. (2010). Age differences in the monitoring of learning: Cross-sectional evidence of spared resolution across the adult life span. *Developmental Psychology*, 46(4), 939–948. doi:10.1037/a0019812
- Hoffmann-Biencourt, A., Lockl, K., Schneider, W., Ackerman, R., & Koriat, A. (2010). Self-paced study time as a cue for recall predictions across school age. *British Journal of Developmental Psychology*, 28(4), 767–784. doi:10.1348/026151009X479042
- Imai, K., Keele, L., & Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological Methods*, 15(4), 309–334. doi:10.1037/a0020761
- Jang, Y., & Nelson, T. O. (2005). How many dimensions underlie judgments of learning and recall? Evidence from state-trace methodology. *Journal of Experimental Psychology: General*, 134(3), 308–326. doi:10.1037/0096-3445.134.3.308
- Jang, Y., Wallsten, T. S., & Huber, D. E. (2012). A stochastic detection and retrieval model for the study of metacognition. *Psychological Review*, 119(1), 186–200. doi:10.1037/a0025960
- Kenny, D. A., Korchmaros, J. D., & Bolger, N. (2003). Lower level mediation in multilevel models. *Psychological Methods*, 8(2), 115–128. doi:10.1037/1082-989X.8.2.115
- Kimball, D. R., & Metcalfe, J. (2003). Delaying judgments of learning affects memory, not metamemory. *Memory & Cognition*, 31(6), 918–929. doi:10.3758/BF03196445
- Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General*, 126(4), 349–370. doi:10.1037/0096-3445.126.4.349
- Koriat, A. (2008). Easy comes, easy goes? The link between learning and remembering and its exploitation in metacognition. *Memory & Cognition*, 36(2), 416–428. doi:10.3758/MC.36.2.416
- Koriat, A., & Ackerman, R. (2010). Metacognition and mindreading: Judgments of learning for self and other during self-paced study. *Consciousness and Cognition*, 19(1), 251–264. doi:10.1016/j.concog.2009.12.010
- Koriat, A., Ackerman, R., Lockl, K., & Schneider, W. (2009). The easily learned, easily remembered heuristic in children. *Cognitive Development*, 24(2), 169–182. doi:10.1016/j.cogdev.2009.01.001
- 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. doi:10.1037/0278-7393.31.2.187
- Koriat, A., & Bjork, R. A. (2006). Mending metacognitive illusions: A comparison of mnemonic-based and theory-based procedures. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(5), 1133–1345. doi:10.1037/0278-7393.32.5.1133
- Koriat, A., Bjork, R. A., Sheffer, L., & Bar, S. K. (2004). Predicting one's own forgetting: The role of experience-based and theory-based processes. *Journal of Experimental Psychology: General*, 133(4), 643–656. doi:10.1037/0096-3445.133.4.643
- Koriat, A., & Ma'ayan, H. (2005). The effects of encoding fluency and retrieval fluency on judgments of learning. *Journal of Memory and Language*, 52(4), 478–492. doi:10.1016/j.jml.2005.01.001
- Koriat, A., Ma'ayan, H., & Nussinson, R. (2006). The intricate relationships between monitoring and control in metacognition: Lessons for the cause-and-effect relation between subjective experience and behavior. *Journal of Experimental Psychology: General*, 135(1), 36–69. doi:10.1037/0096-3445.135.1.36
- Kornell, N., & Bjork, R. A. (2009). A stability bias in human memory: Overestimating remembering and underestimating learning. *Journal of Experimental Psychology: General*, 138(4), 449–468. doi:10.1037/a0017350
- Kornell, N., Rhodes, M. G., Castel, A. D., & Tauber, S. K. (2011). The ease-of-processing heuristic and the stability bias: Dissociating memory, memory beliefs, and memory judgments. *Psychological Science*, 22(6), 787–794. doi:10.1177/0956797611407929
- Krull, J. L., & MacKinnon, D. P. (2001). Multilevel modeling of individual and group level mediated effects. *Multivariate Behavioral Research*, 36(2), 249–277. doi:10.1207/S15327906MBR3602_06
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2014). lmerTest: Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package) (R package version 2.0-6) [Software]. Retrieved from <http://cran.r-project.org/package=lmerTest>
- MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods*, 7(1), 83–104. doi:10.1037/1082-989X.7.1.83
- Matvey, G., Dunlosky, J., & Guttentag, R. (2001). Fluency of retrieval at study affects judgments of learning (JOLs): An analytic or nonanalytic basis for JOLs? *Memory & Cognition*, 29(2), 222–233. doi:10.3758/BF03194916
- Melinger, A., & Weber, A. (2006). Database of noun associations for German. Retrieved from www.coli.uni-saarland.de/projects/nag/
- Miele, D. B., Finn, B., & Molden, D. C. (2011). Does easily learned mean easily remembered? It depends on your beliefs about intelligence. *Psychological Science*, 22(3), 320–324. doi:10.1177/0956797610397954
- Mueller, M. L., Dunlosky, J., Tauber, S. K., & Rhodes, M. G. (2014). The font-size effect on judgments of learning: Does it exemplify fluency effects or reflect people's beliefs about memory? *Journal of Memory and Language*, 70, 1–12. doi:10.1016/j.jml.2013.09.007
- Mueller, M. L., Tauber, S. K., & Dunlosky, J. (2013). Contributions of beliefs and processing fluency to the effect of relatedness on judgments of learning. *Psychonomic Bulletin & Review*, 20(2), 378–384. doi:10.3758/s13423-012-0343-6
- R Core Team. (2013). R: A language and environment for statistical computing [Software]. Vienna, Austria: R Foundation for Statistical Computing.
- 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. doi:10.1037/a0013684
- Serra, M. J., & Ariel, R. (2014). People use the memory for past-test heuristic as an explicit cue for judgments of learning. *Memory & Cognition*. doi:10.3758/s13421-014-0431-0
- Soderstrom, N. C., & McCabe, D. P. (2011). The interplay between value and relatedness as bases for metacognitive monitoring and control: Evidence for agenda-based monitoring. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(5), 1236–1242. doi:10.1037/a0023548
- Susser, J. A., Mulligan, N. W., & Besken, M. (2013). The effects of list composition and perceptual fluency on judgments of learning (JOLs). *Memory & Cognition*, 41(7), 1000–1011. doi:10.3758/s13421-013-0323-8
- Tingley, D., Yamamoto, T., Keele, L., & Imai, K. (2013). Mediation: R package for causal mediation analysis (R package version 4.4) [Software]. Retrieved from <http://cran.r-project.org/package=mediation>
- Undorf, M., & Erdfelder, E. (2011). Judgments of learning reflect encoding fluency: Conclusive evidence for the ease-of-processing hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(5), 1264–1269. doi:10.1037/a0023719
- Undorf, M., & Erdfelder, E. (2013). Separation of encoding fluency and item difficulty effects on judgements of learning. *The Quarterly Journal of Experimental Psychology*, 66(10), 2060–2072. doi:10.1080/17470218.2013.777751
- Yaniv, I., & Meyer, D. E. (1987). Activation and metacognition of inaccessible stored information: Potential bases for incubation effects in problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13(2), 187–205. doi:10.1037/0278-7393.13.2.187