Cognition 222 (2022) 105011

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

"You don't know what this means to me" – Uncovering idiosyncratic influences on metamemory judgments

Monika Undorf [*], Sofia Navarro-Báez, Arndt Bröder

Department of Psychology, School of Social Sciences, University of Mannheim, Germany

ARTICLE INFO

Keywords:
Metamemory
Lens model
Idiosyncratic information
Cue integration
Judgments of learning

ABSTRACT

Studies of the mind often focus on general effects on cognitive processes, whereas influences of idiosyncratic interactions between participants and items evade experimental control or assessment. For instance, assessments of one's own learning and memory processes—metamemory judgments—are attributed to people's reliance on commonly shared characteristics of study materials (e.g., word frequency) or learning conditions (e.g., number of study opportunities). By contrast, few studies have investigated how idiosyncratic information such as the personal significance of items affects memory and metamemory. We propose that hitherto elusive idiosyncratic influences on metamemory can be measured by the *C* component of Egon Brunswik's (1952) lens model. In two experiments, we made randomly chosen items personally significant (Experiment 1) or assessed the personal significance of items (Experiment 2). Personal significance increased both metamemory judgments and memory performance. Including personal significance as a predictor in the lens model reduced *C*, whereas including familiarity from a previous encounter did not. Hence, at least part of the lens model's *C* parameter captures idiosyncratic influences on metamemory. The *C* parameter may serve as a useful tool for future research.

1. Introduction

In studies of the mind, interactions between items and participants are often treated as error variance (Hintzman, 1980; Raaijmakers, 2003; Rouder & Lu, 2005). This is not because participant-item interactions are theoretically uninteresting or minor sources of variability, but because they are extremely difficult to manipulate and assess (Curran & Hintzman, 1995; Flexser, 1981; Hintzman, 1980). In research on memory or metamemory—the ability to assess one's own learning and memory—participant-item interactions arise when the actual or perceived memorability of specific items vary across participants. This might be due to idiosyncratic encoding strategies, individual differences in pre-experimental familiarity, or personal significance of items (Hintzman, 1980). For instance, the study word *knot* might be personally significant and particularly memorable for an avid sailor but unremarkable for other participants.

2. The importance of idiosyncratic influences on memory and metamemory

Researchers have long recognized the importance of idiosyncratic influences on memory (e.g., Brown, 1976; Brown, Lewis, & Monk, 1977; Flexser, 1981; Glanzer & Adams, 1990; Hintzman, 1980) and metamemory (e.g., Daniels, Toth, & Hertzog, 2009; Koriat, 1997; Koriat, Ma'ayan, & Nussinson, 2006; Koriat, Undorf, Newman, & Schwarz, 2020; Murayama, Sakaki, Yan, & Smith, 2014; Nelson & Narens, 1990; Tullis & Fraundorf, 2017). For instance, in recognition memory tests, people are particularly confident in correct rejections of subjectively memorable distractor items such as their hometown or the first name of a close friend (Brown et al., 1977). Because of this, subjective memorability has been proposed to underlie inverse relations of hits and false alarms in recognition memory (mirror effect, Glanzer & Adams, 1990; but see Wixted, 1992). Similarly, Wahlheim, Maddox, and Jacoby (2014) spaced item repetitions across two study lists and found that between-list repetitions improved memory performance only when the presentation of an item on the second study list reminded the specific learner of the item's presentation on the first study list (i.e., when the learner indicated that they had previously studied the item). In metamemory research, idiosyncratic influences on people's predictions of their future memory performance (judgments of learning, JOLs) are considered responsible for improved accuracy of JOLs with repeated study-test trials (Koriat, 1997). Also, idiosyncratic influences on

* Corresponding author at: Department of Psychology, School of Social Sciences, University of Mannheim, A 5, 6, C 203, 68159 Mannheim, Germany.
E-mail address: undorf@uni-mannheim.de (M. Undorf).

https://doi.org/10.1016/j.cognition.2021.105011
Received 8 June 2021; Received in revised form 23 December 2021; Accepted 28 December 2021
Available online 7 February 2022
0010-0277/

metamemory are assumed to underlie accuracy advantages of metamemory judgments for oneself over metamemory judgments for others (Koriat, 1993; Koriat & Ackerman, 2010; Nelson & Narens, 1990; Tullis & Fraundorf, 2017). As a final example, when responding to personality items (e.g., "I enjoy intellectual challenges"), idiosyncratic influences probably contribute to stronger relations between people's confidence in their responses and the consistency of their own responses across repeated item administrations than between people's confidence and the likelihood that others make the same response (Koriat et al., 2020).

Despite the hypothesized or empirically demonstrated importance of idiosyncratic processes, the vast majority of metamemory research has left these processes unaddressed. This is probably due to the lack of a method for assessing idiosyncratic influences on metamemory judgments in standard metamemory paradigms (i.e., without repeated study-test phases or item administrations; without metamemory judgments for self and others). Unfortunately, failing to take idiosyncratic participant-item interactions into account may produce invalid conclusions concerning the basis and accuracy of metamemory.

People are known to base metamemory judgments such as JOLs on commonly shared intrinsic or extrinsic cues pertaining to the stimuli and to the learning conditions (Koriat, 1997; for a review, see Rhodes, 2016). Examples include word frequency, valence, arousal, font size, and the number of study opportunities (e.g., Undorf, Söllner, & Bröder, 2018). Such cues can affect JOLs through theory-based and experience-based processes. Theory-based processes involve deliberate applications of beliefs about memory (Koriat, 1997). For instance, learners may become aware of differences in word frequency and draw explicitly on their belief that high frequency words are more memorable than low frequency words (Mendes, Luna, & Albuquerque, 2019). In contrast, experience-based processes rely on mnemonic cues such as the fluency of processing items during study (Koriat, 1997). Experience-based processes have the phenomenal quality of direct and unexplained intuitions without awareness of their basis. We propose that idiosyncratic information, such as, for example, the personal significance of specific stimuli, might affect JOLs through theory-based processes, experience-based processes, or both.

Addressing the contribution of idiosyncratic information to metamemory may extend our understanding of the basis of metamemory and qualify conclusions concerning its accuracy. For instance, someone who assigns high JOLs to *beetroot*, *rake*, and *stewpot* might not fail to rely on the valid commonly shared cues word frequency, arousal, and valence, but might supplement these cues by idiosyncratic information: her late aunt who was a passionate gardener and enjoyed beetroot casserole. Considering idiosyncratic information may therefore help to explain why JOLs reveal above-chance resolution (i.e., distinguish between items people will and will not remember) even when people fail to base their JOLs on valid commonly shared cues or base their JOLs on invalid commonly shared cues (for metamemory illusions, see Bjork, Dunlosky, & Kornell, 2013; Undorf, 2020). For instance, in a study by Undorf and Zimdahl (2019), JOLs dramatically overestimated the beneficial effects of large font sizes for memory but revealed reliable resolution; mean gamma correlation between JOLs and recall of 0.24, $t(234) = 13.96$, $p < .001$, $d = 0.91$.

3. Measuring idiosyncratic influences on memory and metamemory

In experiments on memory and metamemory, researchers often aggregate dependent variables across items, participants, or both to derive summary measures such as recall probabilities or mean JOLs (Murayama et al., 2014; Rouder, Lu, Morey, Sun, & Speckman, 2008). Aggregation, however, makes it impossible to examine idiosyncratic influences on memory and metamemory, because these are bound to average out (see, e.g., Nelson, Leonesio, Landwehr, & Narens, 1986). Consequently, mixed-effects models might appear as an ideal tool for examining idiosyncratic influences on memory and metamemory. While it is true that mixed-effects models do not require aggregation, are run on unaveraged data, and can simultaneously assess variation at the level of items and participants (Baayen, Davidson, & Bates, 2008; Barr, Levy, Scheepers, & Tily, 2013; Murayama et al., 2014), there is often no straightforward way for assessing idiosyncratic influences on memory and metamemory. The reason for this is that mixed-effects models need two or more responses from a participant to an item in order to distinguish variability in the reactions of specific participants to specific items from residual error (Baayen et al., 2008; Barr et al., 2013).[1] Taken together, assessing idiosyncratic influences on memory and metamemory in standard designs requires a novel approach, because statistical techniques previously used in the literature are incapable of assessing idiosyncratic influences.

3.1. Brunswik's lens model as a potential measurement tool

Bröder and Undorf (2019) hypothesized that the *C* component of Egon Brunswik's (1952) *lens model* might be a potential indicator of idiosyncratic influences on metamemory judgments. This speculation was based on the authors' accidental finding that a lens model analysis of JOL data yielded high values of the *C* component that exceeded those usually found in judgment studies (see below for more information).

The lens model is a classical analysis tool in judgment research and can be applied to any task in which people predict a distal variable in the environment (a criterion) on the basis of probabilistic cues (Hammond, 1955; Karelaia & Hogarth, 2008). When applied to JOLs (see Fig. 1), people predict their memory performance (the criterion) using cues that are available at study (Dunlosky & Tauber, 2014; Koriat et al., 2006; Robey, Dougherty, & Buttaccio, 2017). As mentioned above, the notion that JOLs are based on probabilistic cues is generally agreed on by metamemory researchers (Dunlosky & Metcalfe, 2009; Koriat, 1997; Rhodes, 2016). For example, in a study by Undorf et al. (2018), study words varied in emotionality, concreteness, font size, and frequency of study presentations. These attributes of the stimuli and the learning conditions may affect the to-be-predicted criterion (memory performance) and may inform people's JOLs. Brunswik's (1952) lens model methodology jointly analyses the weighting of cues by the judge (in JOL studies: the learner) in their judgments (in JOL studies: JOLs) as well as the objective weights with which the cues actually affect the criterion (in JOL studies: memory performance).

A lens model analysis consists of computing two multiple linear regressions for each judge across a set of items. One regression analysis is used to determine optimal regression weights of cues for predicting the participant's memory performance (*cue validities*). The second regression analysis determines *cue utilization*, that is, the weight the participant gives to each cue in their JOLs. In addition, the Pearson correlation r_a between JOLs and recall is computed as a measure of accuracy. It is typically referred to as *achievement* in judgment studies and is conceptually equivalent to the resolution of JOLs.

A statistical framework developed by Hammond, Hursch, and Todd (1964) and Tucker (1964) allows decomposing the achievement r_a into four components: (1) the linear predictability of each participant's memory performance on the basis of cues (R_{REC}), (2) the consistency with which the participant weighs the cues when making JOLs (R_{JOL}), (3) the matching index *G* reflecting how closely the participant's cue weights correspond to the cues' effects on memory, and, finally, (4) the *C* component indicating systematic covariation between judgments and recall that cannot be attributed to a linear combination of the cues included in the lens model. R_{REC} and R_{JOL} are the multiple correlation coefficients for the regression predicting recall and the regression

[1] Notably, replicating item-participant combinations is often impossible in studies on memory and metamemory because actual and perceived memorability of study materials increase with repeated presentations, confounding residual error with learning effects.

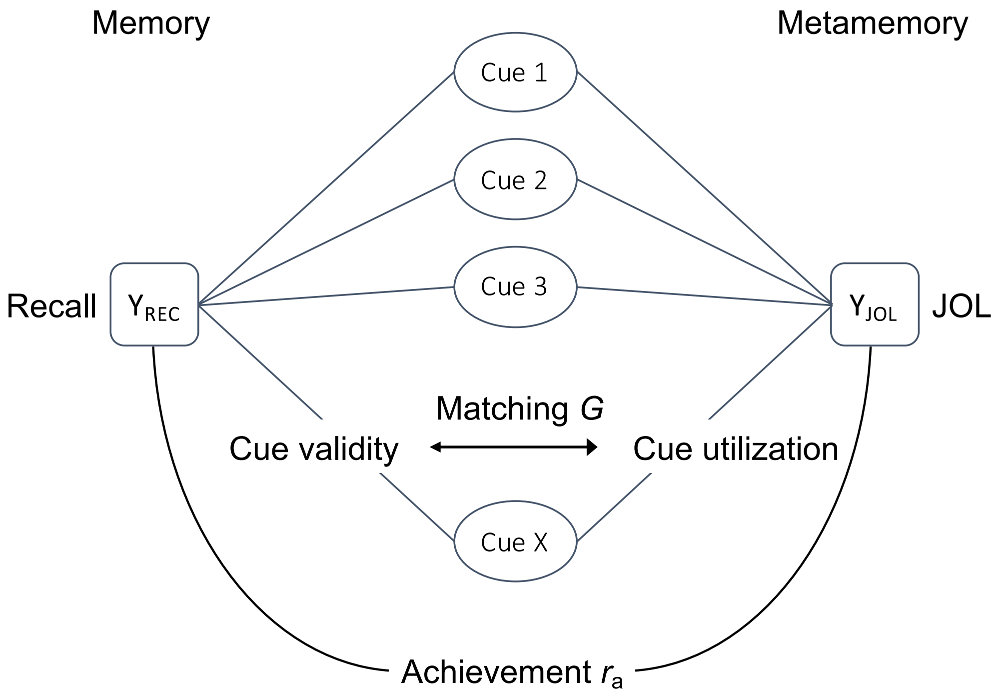

Fig. 1. Schematic representation of the lens model.
Note. Recall is modeled as a linear combination of the cues whose weights are called cue validities. JOLs are modeled as a linear combination of the cues, with weights called cue utilizations. The correlation r_a between JOLs and recall, called achievement, is a function of the match between cue weights (G), linear predictability of the environment (R_{Rec}), consistency of the JOLs (R_{JOL}), and a nonlinear component C.

predicting JOLs, respectively. The matching parameter G, which measures the appropriate weighting of cues by the judge, is technically given by the correlation between the predicted values of the regression models. The C component is the correlation between the regression residuals. Hence, C reflects the portion of achievement r_a that is not predictable from the cues but nevertheless systematic. Tucker's (1964) famous lens model equation (Eq. (1)) formulated the decomposition of achievement into the lens model parameters, adapted here to JOL data:

$$r_a = GR_{REC}R_{JOL} + C\sqrt{\left(1 - R_{REC}^2\right)\left(1 - R_{JOL}^2\right)} \quad (1)$$

Hence, a lens model analysis of a participant's JOLs and memory performance provides four parameters that characterize their achievement or, put differently, the accuracy of their JOLs.

3.2. Idiosyncratic influences and the lens model's C component

For typical applications in judgment research, the lens model's C component is often small and not reliably different from zero. For instance, a meta-analysis of 204 lens model studies revealed a mean C parameter of 0.04 (Karelaia & Hogarth, 2008). Consequently, C is typically ignored, and judgments are characterized as following from linear cue combinations (Brehmer, 1994). Surprisingly, when applying the lens model to data from metamemory experiments, Bröder and Undorf (2019) found that C had a mean of 0.19 and was significantly different from zero in each experiment and condition (for similar findings, see Mendes & Undorf, 2021). This difference between metamemory studies and typical judgment studies requires explanation.

In standard lens model studies, participants predict a criterion, such as a diagnosis, on the basis of presented cue profiles, such as vignettes (see Karelaia & Hogarth, 2008, for examples). In these studies, all relevant information is coded in the cues and no additional information is available. In contrast, the way in which participants process the natural stimuli in studies on memory and metamemory may depend on their autobiographical or semantic memories and personalities. Bröder and Undorf (2019) therefore conjectured that the relatively high C parameter in metamemory experiments might reflect idiosyncratic influences. These idiosyncratic influences do *not* depend on commonly shared item characteristics (like concreteness or emotionality) but represent the meaning of a specific item for a specific person. Consistent with this interpretation, pre-study JOLs made on the basis of explicit information about concreteness and emotionality *before* studying the respective item exhibited a C parameter near zero (see Bröder & Undorf, 2019, Fig. 6). It is important to note that this was an accidental finding from a study focusing on a different research question (see Undorf & Bröder, 2020). Moreover, because soliciting pre-study JOLs deprived people not only of idiosyncratic information but also of all commonly shared cues except for concreteness and emotionality, this preliminary finding is no more than suggestive. Thus, the proposal that the lens model's C parameter measures idiosyncratic influences on metamemory judgments so far is post hoc and speculative.

The current study directly tested the hypothesis that the lens model's C parameter measures idiosyncratic influences on metamemory judgments. In two experiments, participants studied, judged, and were tested on words, some of which we knew to be personally significant. In Experiment 1, we induced personal significance by having participants write a story using synonyms of some study words prior to the JOL experiment. In Experiment 2, personally significant study words matched pictures that participants had previously selected as characteristic of themselves. We expected that personal significance would increase JOLs and recall, consistent with research showing superior memory and higher metamemory judgments for information related to the self (self-reference effect; Boduroglu, Pehlivanoglu, Tekcan, & Kapucu, 2015; Rogers, Kuiper, & Kirker, 1977; Symons & Johnson, 1997). Based on our hypothesis that the lens model's C parameter measures idiosyncratic influences, we predicted that including experimentally-controlled personal significance as a predictor in the lens model would reduce the C parameter.

4. Experiment 1

In Experiment 1, we made a subset of study words personally significant by having participants write a story about synonyms of these words.

4.1. Method

4.1.1. Design and materials

The study list consisted of 50 German 5–9 letter nouns. Words were chosen on the basis of Võ et al. (2009) and were of neutral valence ($M = 0.43$, $SD = 1.04$), moderate arousal ($M = 2.74$, $SD = 0.55$), and moderate concreteness ($M = 4.19$, $SD = 2.23$). Four additional words served as primacy buffers and were not included in the analysis. We divided the study list into five sublists of ten study words. Sublists were comparable with respect to means and variance of valence, arousal, and concreteness. For each sublist, we created a synonym list that comprised synonyms of all ten words. Synonyms rather than the target words were used in the story-writing task to prevent that potential effects of personal significance were due to increased familiarity.

4.1.2. Participants

We aimed at $N = 50$ for counterbalancing of the five synonym lists. This sample size provides a statistical power of $(1 - \beta) = 0.93$ to detect medium-sized main effects ($f = 0.25$, $d = 0.50$) with $\alpha = 0.05$ in repeated-measures ANOVAs and *t*-tests for dependent samples (Faul, Erdfelder, Lang, & Buchner, 2007). Participants were 50 University of Mannheim undergraduates.

4.1.3. Procedure

The experiment consisted of a story-writing task, a study phase, and a free recall test. Participants learned that the study consisted of two unrelated tasks. The first task supposedly addressed creativity. Participants had 10 min to write a creative story that had to include all ten words from one synonym list. Synonym lists were counterbalanced such that all words were made personally significant equally often. After the story-writing task, participants studied 54 words for a memory test and estimated the probability of recalling each word at test. At study, words were presented for 3 s each and in a new random order for each participant. Immediately after each word, the JOL prompt “Chance of recall (0%-100%)?” appeared, and participants typed in any whole number from 0 to 100 (self-paced). A 100-ms blank screen preceded the next study trial. Following a 3-min numerical filler task, participants had 5 min to write down as many studied words as they could remember.

4.2. Results

The mean number of intrusions in the recall test was 1.10 ($SD = 1.43$) and the mean number of story word intrusions was 0.26 ($SD = 0.53$, 16.33% of intrusions). Fig. 2 presents mean JOLs and mean recall performance for items with and without personal significance. JOLs were submitted to a one-way ANOVA with item type (personal significance: yes vs. no) as the within-subjects factor. A significant main effect revealed higher JOLs for personally significant items, $F(1, 49) = 33.30$, $p < .001$, $\eta_p^2 = 0.40$. A similar ANOVA on recall performance revealed better memory performance for personally significant items, $F(1, 49) = 29.38$, $p < .001$, $\eta_p^2 = 0.37$.

As can be seen in Table 1, JOL resolution (as measured by gamma correlations) was reliable and did not vary with item type, $t(49) = 1.92$, $p = .061$, $d = 0.27$. JOLs overestimated memory performance for control items but not for personally significant items, resulting in better calibration for personally significant items, $t(49) = 3.04$, $p = .004$, $d = 0.43$.

We applied two lens models to each participant's data. The basic model included the item characteristics valence, arousal, concreteness,

Table 1
Means (and Standard Deviations) of resolution (Gamma) and calibration (Bias) in experiments 1 and 2.

Experiment and item type	Gamma	Bias
Experiment 1		
No personal significance	0.38*** (0.34)	8.94*** (15.18)
Personal significance	0.54*** (0.44)	1.40 (17.88)
Overall	0.47*** (0.24)	7.43*** (14.11)
Experiment 2		
Not previously encountered	0.23*** (0.43)	4.27 (18.91)
Previously encountered	0.32*** (0.53)	−6.12 (24.55)
Personal significance	0.12 (0.48)	−10.22** (23.97)
Overall	0.31*** (0.30)	−0.71 (18.82)

Note. Gamma = Gamma correlation between JOLs and recall; bias = signed difference between mean JOLs and mean recall performance.
Asterisks refer to one-sample *t*-tests against 0.
** $p < .010$.
*** $p < .001$.

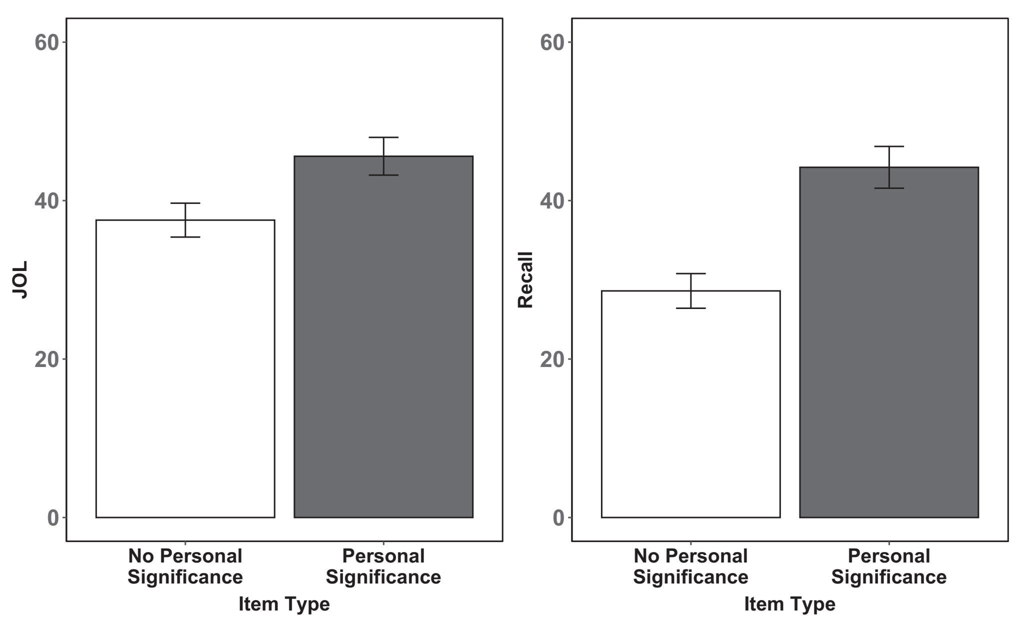

Fig. 2. Mean Judgments of Learning (JOL) and Percentage of Correctly Recalled Words (Recall) in Experiment 1.
Note. Error bars represent one standard error of the mean.

frequency, and number of letters as predictors. The full model additionally included a dummy variable coding whether items had experimentally-controlled personal significance (1 = yes, 0 = no). Table 2 presents descriptive statistics for all parameters from both lens models. Both models yielded significantly positive *C* parameters, $t \geq 10.36$, $p < .001$, $d \geq 1.48$. Most importantly, a paired *t*-test showed that including personal significance in the lens model reliably reduced *C*, $t(49) = 3.24$, $p = .001$, $d = 0.46$ (see Fig. 3). We also found that including personal significance in the lens model increased the linear predictability of memory performance (R_{REC}) and the consistency of cue use in JOLs (R_{JOL}). The matching index *G* was numerically but not reliably higher when the lens model included personal significance.

4.3. Discussion

Making words personally significant increased JOLs and memory performance. Including experimentally-controlled personal significance as a predictor in the lens model reduced the *C* parameter, suggesting that this parameter reflects idiosyncratic influences on JOLs. However, although participants wrote stories about *synonyms* rather than the study words, one objection might be that some of the variance reflected in *C* is due to increased familiarity. We addressed this issue in Experiment 2. Another objection might be that *C* was still substantial when including personal significance in the lens model. We do not, however, claim that idiosyncratic influences on JOLs are restricted to experimentally-controlled personal significance. Rather, we suspect that various idiosyncratic features affect JOLs and memory and therefore contribute to *C* in addition to experimentally-controlled personal significance. At the same time, part of *C* may reflect influences other than idiosyncratic processes (see also General Discussion).

In the lens models used in Experiment 1, we included the item characteristics valence, arousal, concreteness, word frequency, and number of letters as predictors of JOLs and memory performance. We did so to minimize the chances that the lens model's *C* parameter reflected contributions of omitted item characteristics to JOLs and memory performance. We do not mean to suggest, however, that participants used each of these characteristics when making JOLs. Based on prior research, it is plausible that valence, arousal, concreteness, and word frequency affected JOLs and memory performance, even though not necessarily to the same degree or in the same direction (Fiacconi & Dollois, 2020; Hourihan, Fraundorf, & Benjamin, 2017; Mendes & Undorf, 2021; Undorf et al., 2018; Undorf & Bröder, 2020; Witherby & Tauber, 2017). In contrast, there is no reason to expect that JOLs and memory performance depended on the number of letters.

5. Experiment 2

Experiment 2 aimed to disentangle personal significance and increased familiarity due to a recent encounter. Prior to the JOL experiment, participants selected images that characterized themselves in an allegedly unrelated personality study (e.g., a hardworking participant might select a bee, an athletic participant might select a sneaker). We expected that study words matching selected images were personally significant, whereas words that matched images participants had seen but not selected were only familiar. Also, including whether items were familiarized in the lens model should reduce *C* not at all or not as much as including experimentally-controlled personal significance.

5.1. Method

5.1.1. Design and materials

Stimuli were 80 German 4–8 letter nouns and 80 images that matched these words. Words were of neutral valence ($M = 0.75$, $SD = 0.58$), moderate arousal ($M = 2.21$, $SD = 0.35$), and high concreteness ($M = 5.80$, $SD = 0.38$; normed values taken from Võ et al., 2009). Four additional words served as primacy buffers and were not included in the analysis. To ensure that words with and without experimentally-controlled personal significance were similar in valence, arousal, and concreteness, we constructed ten sublists of eight items that were similar in these characteristics.

5.1.2. Participants

Power calculations were identical to those reported in Experiment 1. Participants were 50 University of Mannheim undergraduates.

5.1.3. Procedure

The procedure was identical to that of Experiment 1 with the following exceptions. The story-writing task was replaced with a personality questionnaire (BFI-10, Rammstedt & John, 2007) and an image-selection task presented as an image-based personality test. The image-selection task consisted of ten trials (order randomly determined for each participant). On each trial, participants saw five randomly chosen images from one sublist and selected the image that best characterized themselves. Each participant's study list included 10 words with personal significance (matching selected images) and 40 words without personal significance, 10 of which were previously encountered (matching images that were presented but not selected) and 30 of which were not previously encountered (matching images that were not presented).

5.2. Results

The mean number of intrusions in the recall test was 1.40 ($SD = 1.77$) and the mean number of intrusions that matched images from the image-selection task was 0.56 ($SD = 0.73$, 48.88% of intrusions). Fig. 4 presents mean JOLs and mean recall performance for not previously encountered, previously encountered, and personally significant items. A one-way ANOVA on JOLs with item type (not previously encountered, previously encountered, personally significant) as within-subjects factor

Table 2
Means and medians of lens model parameters in experiments 1 and 2.

Experiment and model	Predictors	Lens model parameters				
		Achievement r_a	Matching *G*	R_{JOL}	R_{REC}	*C*
Experiment 1						
Basic	Item characteristics	0.35/0.35	0.53/0.60	0.21_a/0.19	0.16_a/0.15	0.31_a /0.31
Full	+ Personal significance (y/n)	0.35/0.35	0.56/0.69	0.26_b/0.25	0.21_b /0.20	0.28_b /0.28
Experiment 2						
Basic	Item characteristics	0.25/0.26	0.13_a/0.12	0.14_a/0.14	0.10_a/0.09	0.25_a/0.24
Familiarity	+ Previously encountered (y/n)	0.25/0.26	0.16_a/0.10	0.15_a/0.16	0.14_b/0.11	0.25_a/0.25
Full	+ Personal significance (y/n)	0.25/0.26	0.37_b/0.46	0.27_b/0.24	0.21_c/0.18	0.20_b/0.20

Note. Achievement r_a = Pearson correlation between JOLs and recall; matching G = matching index; R_{JOL} and R_{REC} = multiple correlations between cues and JOLs or recall, respectively; C = nonlinear parameter. Within each experiment and column, means with different subscripts are different at $p < .05$ (Experiment 1: *t*-tests, Experiment 2: Tukey's HSD tests).

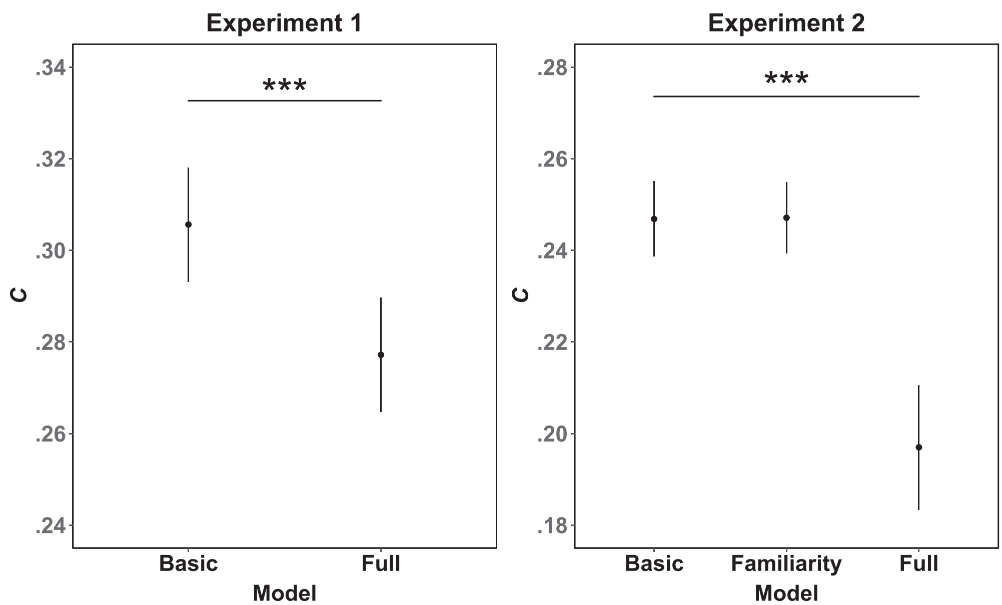

Fig. 3. *C* Parameter from Brunswik's (1952) Lens Model in Experiments 1 and 2.
Note. The basic model included item characteristics as predictors. The familiarity model additionally included as a predictor whether items were previously encountered (y/n, Experiment 2 only) and the full model additionally included experimentally-controlled personal significance as a predictor (y/n).

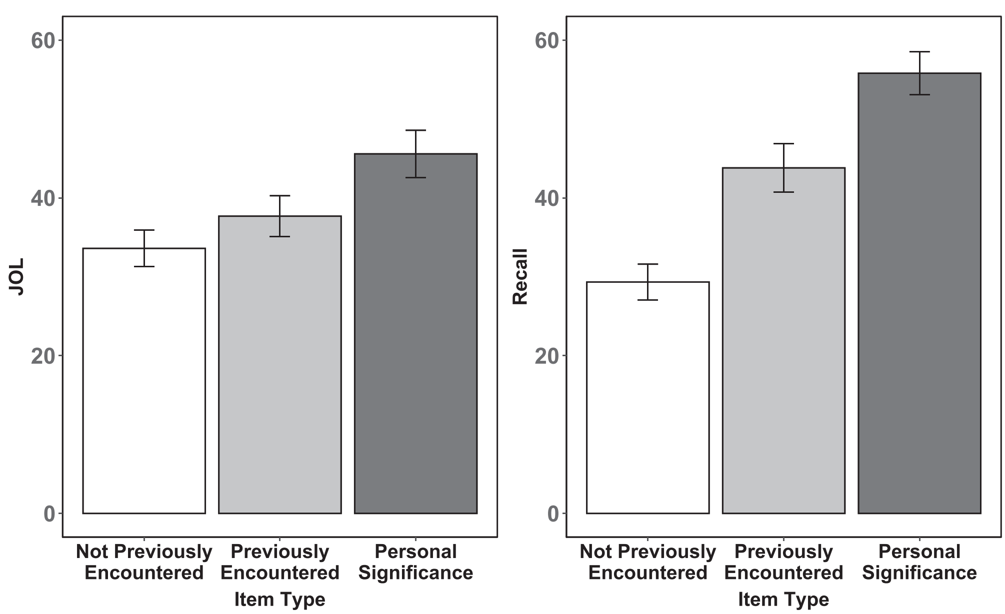

Fig. 4. Mean Judgments of Learning (JOL) and Percentage of Correctly Recalled Words (Recall) in Experiment 2.
Note. Error bars represent one standard error of the mean.

revealed that JOLs varied with item type, $F(1.48, 72.34) = 41.92$, $p < .001$, $\eta_p^2 = 0.46$ (Greenhouse-Geisser corrected). Pairwise comparisons showed significant differences among all three item types, $t \geq 4.63$, $p < .001$, $d \geq 0.66$. A similar ANOVA on recall performance indicated that memory performance varied with item type, $F(1.57, 76.80) = 40.99$, $p < .001$, $\eta_p^2 = 0.46$. Pairwise comparisons showed significant differences among all three item types, $t \geq 3.32$, $p \leq 0.002$, $d \geq 0.47$.

JOL resolution (see Table 1) was reliable (except for personally significant items) and did not vary with item type, $F(2.10, 49.14) = 2.10$, $p = .131$, $\eta_p^2 = 0.04$. JOLs underestimated memory performance for personally significant items but were well-calibrated for the other item types. Calibration varied with item type, $F(1.55, 76.05) = 13.78$, $p < .001$, $\eta_p^2 = 0.22$, with significant differences between not previously encountered items and the other item types, $t \geq 4.20$, $p < .001$, $d \geq 0.60$, but no differences between previously encountered and personally significant items, $t(49) = 1.16$, $p = .25$, $d = 0.17$.

We applied three lens models to each participant's data (see Table 2). As in Experiment 1, the basic model included the word characteristics valence, arousal, concreteness, frequency, and number of letters as predictors. The familiarity model additionally included a dummy variable coding whether items were previously encountered (1 = yes, 0 = no). The full model additionally included a dummy variable coding whether items were personally significant (1 = yes, 0 = no). All three models yielded significant *C* parameters, $t \geq 10.13$, $p < .001$, $d \geq 1.45$, that differed across models, $F(1.25, 61.14) = 32.33$, $p < .001$, $\eta_p^2 = 0.40$ (see Fig. 3). One-tailed pairwise comparisons confirmed that the *C* parameter from the basic model was similar to the C parameter from the familiarity model, $t(49) = 0.07$, $p = .53$, $d = 0.01$, but significantly higher than C from the full model, $t(49) = 5.82$, $p < .001$, $d = 0.83$. C from the familiarity model was significantly higher than C from the full model, $t(49) = 6.00$, $p < .001$, $d = 0.86$ (one-tailed). The linear predictability of memory performance (R_{REC}) was lowest in the basic model, intermediate in the familiarity model and highest in the full model. In contrast, the consistency of cue use in JOLs (R_{JOL}) and the matching index G were reliably higher in the full model than in the familiarity or basic model but similar in the familiarity and basic model.

5.3. Discussion

Selecting images that characterize oneself presumably was a stronger manipulation of personal significance than that used in Experiment 1. We maintained experimental control by comparing words that matched selected images with words that matched other presented and non-presented images. Experimentally-controlled personal significance increased JOLs and memory, and more so than familiarity from the image-selection task. Including personal significance as a predictor in the lens model reduced the C parameter, whereas including previous encounters did not affect C. These results again indicate that the C parameter reflects idiosyncratic influences on metamemory.

6. General discussion

Despite being recognized as important, interactions between items and participants are rarely addressed in psychological research, because they are extremely difficult to assess (Curran & Hintzman, 1995; Flexser, 1981; Hintzman, 1980). This is also true for the domains of memory and metamemory: While researchers have hypothesized and found that item-participant interactions play an important role (Brown et al., 1977; Koriat, 1997; Koriat et al., 2006; Murayama et al., 2014; Nelson & Narens, 1990; Tullis & Fraundorf, 2017; Wahlheim et al., 2014), the majority of studies have left such idiosyncratic influences unaddressed. The reason for this is that there has been no method for assessing idiosyncratic influences in standard paradigms. The present study examined whether the C component of Egon Brunswik's (1952) lens model may serve to measure idiosyncratic item-participant interactions in memory and metamemory.

In two experiments, we investigated the effects of experimentally-controlled personal significance on metamemory and memory. Results showed that personally significant words yielded higher JOLs and memory performance than control words. Including experimentally-controlled personal significance as a predictor in the lens model reduced the model's C parameter. These findings demonstrate that people base metamemory judgments on idiosyncratic information such as personal significance, indicating that learners retrieve specific past events during encoding (Benjamin & Tullis, 2010; Wahlheim et al., 2014) and, more generally, that the self plays an important role in memory and metamemory (Boduroglu et al., 2015; Rogers et al., 1977; Symons & Johnson, 1997). The current findings also reveal that the lens model's C parameter reflects idiosyncratic influences. The latter conclusion holds despite our finding that lens models including personal significance still yielded substantial C parameters, showing that the personal significance induced or assessed by our experimental manipulations is only part of the systematic covariance between JOLs and recall reflected in C. The persistence of a high C parameter despite including experimentally-controlled personal significance as a predictor is not surprising, but was to be expected because various idiosyncratic influences were operative *in addition* to experimentally-controlled significance (e.g., *knot* is personally significant for a sailor even when this word serves as a control item in the experiment). Of course, our data cannot (and are not intended to) prove that the remaining part of the C parameter *exclusively* mirrors idiosyncratic influences. Part of C may stem from unmodeled predictors or nonlinear relationships between cues and recall or JOLs, respectively. Nevertheless, our results clearly show that experimentally-controlled idiosyncratic influences are captured in the C parameter, and this makes it plausible that naturally occurring idiosyncratic influences will be reflected in C as well.

In addition to reducing the C parameter, including experimentally-controlled personal significance as a predictor in the lens model increased the linear predictability of memory performance, the consistency of cue use in JOLs, and the matching index G, although not significantly in Experiment 1. These findings show that considering idiosyncratic information is critical to our understanding of the basis of metamemory judgments.

The present finding that the C parameter is substantial in size, reliably different from zero, and affected by experimental manipulations is new to the lens model literature. Judgment researchers typically ignore C, because it is small and insignificant (Karelaia & Hogarth, 2008). Idiosyncratic information probably does not affect judgments in standard lens model studies, where all participants are provided with the same cues and can base their judgments only on these cues.

When using the lens model to assess idiosyncratic influences on metamemory judgments and memory, it is advisable to include all available commonly shared characteristics of the study materials, the learning conditions, and the learners as predictors. This helps minimize the extent to which the lens model's C parameter reflects contributions of omitted item characteristics to JOLs and memory performance. Of course, including some characteristic of the items, learning conditions, or learners as a predictor in the lens model does not imply that people base their JOLs on this piece of information or that it affects people's memory performance. What is more, even if one or more predictors included in the lens model have significant effects on JOLs, it is possible that people rely on a single unified feeling of ease or distinctiveness rather than base their JOLs on each predictor (but see Undorf & Bröder, 2020).

As mentioned in the Introduction, statistical approaches that require aggregation of dependent variables (e.g., ANOVAs) are incapable of examining naturally occurring idiosyncratic influences in standard designs, because these are likely to cancel each other. The situation is more complex for mixed-effects models that do not require aggregation. In standard memory and metamemory designs, mixed-effects models cannot distinguish participant-item interactions from residual error, because each participant responds only once to each item (Baayen et al., 2008; Barr et al., 2013). Consequently, idiosyncratic influences should reveal themselves in significant random effects. When including various commonly shared characteristics of the study materials, the learning conditions, and the learners as fixed effects in the mixed-effects model, random effects may provide a reasonably pure measure of idiosyncratic influences. Even then, however, using the lens model for measuring idiosyncratic influences has two crucial advantages. First, while mixed-effects models can assess idiosyncratic influences on memory and metamemory only separately, lens model parameters reflect the joint effects of idiosyncratic influences on metacognitive accuracy. Second, unlike random effects from mixed-effects models, the lens model's C parameter assesses only systematic covariation between JOLs and memory performance and is therefore unaffected by unreliability in memory performance, inconsistency in people's cue utilization in JOLs, and missing knowledge about optimal weights of commonly shared cues.

After decades of insightful inquiries into the impact of *commonly shared* characteristics of items and situations on metamemory (e.g., Koriat, 1997), the present study opened up possibilities for future research on the impact of *idiosyncratic* influences using Brunswik's (1952) lens model and its often neglected C parameter. Future research might profitably explore the personal and situational factors driving idiosyncratic processes in memory and metamemory. Also, an important question for future research is whether idiosyncratic influences affect metamemory judgments through theory-based processes, experience-based processes, or both.

Author contribution

Monika Undorf: Conceptualization, Methodology, Software, Formal analysis, Data Curation, Writing - Original Draft, Writing - Review & Editing, Funding acquisition; Sofia Navarro-Báez: Methodology, Software, Formal analysis, Data Curation, Writing - Original Draft, Writing - Review & Editing; Arndt Bröder: Conceptualization, Methodology, Writing - Original Draft, Writing - Review & Editing, Funding acquisition.

Declarations of interest

None

Supplementary material

All materials, data, and code are publically available at https://osf.io/pd8m2/.

Acknowledgments

This work was supported by grants UN 345/2-1 (Monika Undorf) and BR 2130/14-1 (Arndt Bröder) from the Deutsche Forschungsgemeinschaft and by a Margarete von Wrangell fellowship from the state of Baden-Württemberg (Monika Undorf). We thank Franziska Schäfer and Anna Fee Wefelmeier for help with data collection. We thank David Izydorczyk and Sophie Scharf for helpful comments.

References

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Benjamin, A. S., & Tullis, J. G. (2010). What makes distributed practice effective? *Cognitive Psychology, 61*(3), 228–247. https://doi.org/10.1016/j.cogpsych.2010.05.004

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology, 64*, 417–444. https://doi.org/10.1146/annurev-psych-113011-143823

Boduroglu, A., Pehlivanoglu, D., Tekcan, A.İ., & Kapucu, A. (2015). Effects of self-referencing on feeling-of-knowing accuracy and recollective experience. *Memory, 23*(5), 736–747. https://doi.org/10.1080/09658211.2014.925927

Brehmer, B. (1994). The psychology of linear judgement models. *Acta Psychologica, 87*(2–3), 137–154. https://doi.org/10.1016/0001-6918(94)90048-5

Bröder, A., & Undorf, M. (2019). Metamemory viewed through the judgment lens. *Acta Psychologica, 197*, 153–165. https://doi.org/10.1016/j.actpsy.2019.04.011

Brown, J. (1976). An analysis of recognition and recall and of problems in their comparison. In J. Brown (Ed.), *Recall and recognition* (pp. 1–35). Wiley.

Brown, J., Lewis, V. J., & Monk, A. F. (1977). Memorability, word frequency and negative recognition. *The Quarterly Journal of Experimental Psychology, 29*(3), 461–473. https://doi.org/10.1080/14640747708400622

Brunswik, E. (1952). *The conceptual framework of psychology*. University of Chicago Press.

Curran, T., & Hintzman, D. L. (1995). Violations of the independence assumption in process dissociation. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 21*(3), 531–547. https://doi.org/10.1037/0278-7393.21.3.531

Daniels, K. A., Toth, J. P., & Hertzog, C. (2009). Aging and recollection in the accuracy of judgments of learning. *Psychology and Aging, 24*(2), 494–500. https://doi.org/10.1037/a0015269

Dunlosky, J., & Metcalfe, J. (2009). *Metacognition*. Sage.

Dunlosky, J., & Tauber, S. K. (2014). Understanding people's metacognitive judgments: An isomechanism framework and its implications for applied and theoretical research. In T. J. Perfect, & D. S. Lindsay (Eds.), *The sage handbook of applied memory* (pp. 444–464). Sage.

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. https://doi.org/10.3758/BF03193146

Fiacconi, C. M., & Dollois, M. A. (2020). Does word frequency influence judgments of learning (JOLs)? A meta-analytic review. *Canadian Journal of Experimental Psychology, 74*(4), 346–353. https://doi.org/10.1037/cep0000206

Flexser, A. J. (1981). Homogenizing the 2 × 2 contingency table: A method for removing dependencies due to subject and item differences. *Psychological Review, 88*(4), 327–339. https://doi.org/10.1037/0033-295X.88.4.327

Glanzer, M., & Adams, J. K. (1990). The mirror effect in recognition memory: Data and theory. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 16*(1), 5–16. https://doi.org/10.1037/0278-7393.16.1.5

Hammond, K. R. (1955). Probabilistic functioning and the clinical method. *Psychological Review, 62*(4), 255–262. https://doi.org/10.1037/h0046845

Hammond, K. R., Hursch, C. J., & Todd, F. J. (1964). Analyzing the components of clinical inference. *Psychological Review, 71*(6), 438–456. https://doi.org/10.1037/h0040736

Hintzman, D. L. (1980). Simpson's paradox and the analysis of memory retrieval. *Psychological Review, 87*(4), 398–410. https://doi.org/10.1037/0033-295X.87.4.398

Hourihan, K. L., Fraundorf, S. H., & Benjamin, A. S. (2017). The influences of valence and arousal on judgments of learning and on recall. *Memory and Cognition, 45*(1), 121–136. https://doi.org/10.3758/s13421-016-0646-3

Karelaia, N., & Hogarth, R. M. (2008). Determinants of linear judgment: A meta-analysis of lens model studies. *Psychological Bulletin, 134*(3), 404–426. https://doi.org/10.1037/0033-2909.134.3.404

Koriat, A. (1993). How do we know that we know? The accessibility model of the feeling of knowing. *Psychological Review, 100*(4), 609–639. https://doi.org/10.1037/0033-295X.100.4.609

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. https://doi.org/10.1037/0096-3445.126.4.349

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. https://doi.org/10.1016/j.concog.2009.12.010

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. https://doi.org/10.1037/0096-3445.135.1.36

Koriat, A., Undorf, M., Newman, E. J., & Schwarz, N. (2020). Subjective confidence in the response to personality questions: Some insight into the construction of people's responses to test items. *Frontiers in Psychology, 11*. https://doi.org/10.3389/fpsyg.2020.01250

Mendes, P. S., Luna, K., & Albuquerque, P. B. (2019). Word frequency effects on judgments of learning: More than just beliefs. *The Journal of General Psychology*. https://doi.org/10.1080/00221309.2019.1706073

Mendes, P. S., & Undorf, M. (2021). On the pervasive effect of word frequency in metamemory. *The Quarterly Journal of Experimental Psychology, 174702182110533*. https://doi.org/10.1177/17470218211053329

Murayama, K., Sakaki, M., Yan, V. X., & Smith, G. M. (2014). Type I error inflation in the traditional by-participant analysis to metamemory accuracy: A generalized mixed-effects model perspective. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 40*(5), 1287–1306. https://doi.org/10.1037/a0036914

Nelson, T. O., Leonesio, R. J., Landwehr, R. S., & Narens, L. (1986). A comparison of three predictors of an individual's memory performance: The individual's feeling of knowing versus the normative feeling of knowing versus base-rate item difficulty. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 12*, 279–287.

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. Bower (Ed.), *Vol. 26. The psychology of learning and motivation* (pp. 125–173). Academic Press.

Raaijmakers, J. G. W. (2003). A further look at the "language-as-fixed-effect fallacy". *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale, 57*(3), 141–151. https://doi.org/10.1037/h0087421

Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the big five inventory in English and German. *Journal of Research in Personality, 41*(1), 203–212. https://doi.org/10.1016/j.jrp.2006.02.001

Rhodes, M. G. (2016). Judgments of learning: Methods, data, and theory. In J. Dunlosky, & S. K. Tauber (Eds.), *The Oxford handbook of metamemory* (pp. 65–80). Oxford University Press.

Robey, A. M., Dougherty, M. R., & Buttaccio, D. R. (2017). Making retrospective confidence judgments improves learners' ability to decide what not to study. *Psychological Science, 28*(11), 1683–1693. https://doi.org/10.1177/0956797617718800

Rogers, T. B., Kuiper, N. A., & Kirker, W. S. (1977). Self-reference and the encoding of personal information. *Journal of Personality and Social Psychology, 35*(9), 677–688. https://doi.org/10.1037/0022-3514.35.9.677

Rouder, J. N., & Lu, J. (2005). An introduction to Bayesian hierarchical models with an application in the theory of signal detection. *Psychonomic Bulletin & Review, 12*, 573–604. https://doi.org/10.3758/BF03196750

Rouder, J. N., Lu, J., Morey, R. D., Sun, D., & Speckman, P. L. (2008). A hierarchical process-dissociation model. *Journal of Experimental Psychology. General, 137*(2), 370–389. https://doi.org/10.1037/0096-3445.137.2.370

Symons, C. S., & Johnson, B. T. (1997). The self-reference effect in memory: A meta-analysis. *Psychological Bulletin, 121*(3), 371–394. https://doi.org/10.1037/0033-2909.121.3.371

Tucker, L. R. (1964). A suggested alternative formulation in the developments by Hursch, Hammond, and Hursch, and by Hammond, Hursch, and Todd. *Psychological Review, 71*(6), 528–530. https://doi.org/10.1037/h0047061

Tullis, J. G., & Fraundorf, S. H. (2017). Predicting others' memory performance: The accuracy and bases of social metacognition. *Journal of Memory and Language, 95*, 124–137. https://doi.org/10.1016/j.jml.2017.03.003

Undorf, M. (2020). Fluency illusions in metamemory. In A. M. Cleary, & B. L. Schwartz (Eds.), *Memory quirks: The study of odd phenomena in memory* (pp. 150–174). Routledge.

Undorf, M., & Bröder, A. (2020). Cue integration in metamemory judgements is strategic. *The Quarterly Journal of Experimental Psychology, 73*(4), 629–642. https://doi.org/10.1177/1747021819882308

Undorf, M., Söllner, A., & Bröder, A. (2018). Simultaneous utilization of multiple cues in judgments of learning. *Memory and Cognition, 46*(4), 507–519. https://doi.org/10.3758/s13421-017-0780-6

Undorf, M., & Zimdahl, M. F. (2019). Metamemory and memory for a wide range of font sizes: What is the contribution of perceptual fluency? *Journal of Experimental Psychology. Learning, Memory, and Cognition, 45*(1), 97–109. https://doi.org/10.1037/xlm0000571

Võ, M. L.-H., Conrad, M., Kuchinke, L., Urton, K., Hofmann, M. J., & Jacobs, A. M. (2009). The Berlin affective word list reloaded (BAWL-R). *Behavior Research Methods, 41*(2), 534–538. https://doi.org/10.3758/BRM.41.2.534

Wahlheim, C. N., Maddox, G. B., & Jacoby, L. L. (2014). The role of reminding in the effects of spaced repetitions on cued recall: Sufficient but not necessary. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 40*(1), 94–105. https://doi.org/10.1037/a0034055

Witherby, A. E., & Tauber, S. K. (2017). The concreteness effect on judgments of learning: Evaluating the contributions of fluency and beliefs. *Memory and Cognition, 45*(4), 639–650. https://doi.org/10.3758/s13421-016-0681-0

Wixted, J. T. (1992). Subjective memorability and the mirror effect. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 18*(4), 681–690. https://doi.org/10.1037/0278-7393.18.4.681