

Processing Fluency Mediates the Influence of Perceptual Information on Monitoring Learning of Educationally Relevant Materials

B. Hunter Ball, Kathleen N. Klein, and Gene A. Brewer
Arizona State University

Previous research has shown that perceptual characteristics of stimuli inaccurately bias assessments of perceived memorability. However, little research has investigated how perceptual information using real-world study materials affects study time allocation and assessments of future memory performance. In the current study, participants studied a series of terms and their corresponding definitions that varied on perceptual dimensions commonly used in educational material. When participants were allowed to control their own study time, font bolding (Experiment 1) and font size (Experiment 2), but not borders surrounding the text (Experiment 3), influenced judgments of learning despite having no effect on actual memory performance. Items that were processed more easily (as evidenced by study duration) consistently resulted in metacognitive monitoring biases, suggesting that encoding fluency may lead to inaccurate beliefs about one's own learning and future memory performance in educational settings.

Keywords: metamemory, judgments of learning, fluency, monitoring, control

Metamemory refers to the processes by which people monitor and control the contents of memory and make judgments about their memories during acquisition, retention, and retrieval of information (Dunlosky & Metcalfe, 2009; Metcalfe & Dunlosky, 2008). These processes are important when considering how an individual regulates study habits (Mazzoni & Comoldi, 1993; Mazzoni, Cornoldi, & Marchitelli, 1990; Metcalfe & Finn, in press; Pressley & Ghatala, 1990; Thiede, Anderson, & Theriault, 2003; Thiede & Dunlosky, 1999; Winne & Hadwin, 1998), such as deciding what information is important to study, what strategies are most beneficial for learning, and when to terminate study. When metamemory processes are implemented successfully, studying can be enhanced. However, metamemory often exhibits regular departures from accuracy thereby reducing the effectiveness of learning (Serra & Metcalfe, 2009). Previous research suggests that people use a variety of cues when assessing how well information is learned and predicting how well it will be remembered (Koriat, 1997), and that perceptual information may be one particular cue that influences perceived memorability (Rhodes & Castel, 2008, 2009). Given the ubiquity of salient perceptual features in textbooks and other educational material, the current study sought to examine whether metamemory processes can be inappropriately biased by when studying ecologically valid and educationally relevant materials.

Declarative knowledge about one's own cognition (e.g., knowledge of one's own abilities or how a specific task operates; Dunlosky & Metcalfe, 2009) informs monitoring and control processes. Monitoring involves assessing the activity or progress of a cognitive task (e.g., how well the information is learned; Hart, 1965), whereas control involves regulating a cognitive task through action (e.g., making a decision to terminate an activity or selecting a new strategy; Miller, Galanter, & Pribram, 1960). Control decisions usually result from the feedback obtained through monitoring a cognitive task (Nelson & Narens, 1994). For example, when studying for an upcoming test, one might attempt to memorize the labeled components of several diagrams displayed in a textbook. One would monitor this cognitive task by making judgments about how well one has learned a diagram. Based on those judgments of learning one might decide to terminate studying of that diagram and move on to the next diagram, demonstrating control of cognition (e.g., Serra & Dunlosky, 2010). Thus, metamemory monitoring informs the control of decisions to terminate study, which should affect future test performance depending on how well information is learned (cf. Nelson & Leonesio, 1988).

Metamemory monitoring and control processes are often studied by investigating judgments of learning (JOLs) and study duration, respectively, during information acquisition. JOLs involve making confidence judgments (e.g., 0%–100%) of one's ability to recall study information on a test following learning (Dunlosky & Metcalfe, 2009), whereas study duration is generally assessed by measuring the amount of time spent learning an item before making the decision to terminate study (Koriat, 2008). Koriat's (1997) cue-utilization framework assumes that JOLs are made on the basis of rules or heuristics to make an assessment of recall probability through inference (e.g., Begg, Duft, Lalonde, Melnick, & Sanvito, 1989) and that rather than monitoring the strength of the memory trace of an item directly (e.g., Hart, 1967), participants use cues that are generally predictive of future memory perfor-

This article was published Online First October 27, 2014.

B. Hunter Ball, Kathleen N. Klein, and Gene A. Brewer, Department of Psychology, Arizona State University.

Portions of this research were presented in Kathleen N. Klein's Honors Thesis at Arizona State University. We would like to extend a special acknowledgment to Holly O'Rourke and Amanda Baraldi for their assistance in constructing and interpreting the multilevel mediation analyses presented in this article.

Correspondence concerning this article should be addressed to Gene A. Brewer, Department of Psychology, Arizona State University, Tempe, AZ 85287-1104. E-mail: gene.brewer@asu.edu

mance. Theory-based (information-based) judgments are made from analytic inferences about future memory performance based off of deliberate rules or theories retrieved from memory while assessing learning (e.g., “I should remember” a particular item because of its saliency). Mnemonic-based (experience-based) judgments are made from implicit inferences about future memory performance based on subjective experience while performing the task (e.g., encoding fluency; *Koriat & Bjork, 2006*). Considerable research has investigated the types of cues utilized to inform JOLs and study decisions and suggests that learners are generally accurate in their assessments of future memory performance (e.g., *Arbuckle & Cuddy, 1969; King, Zechmeister, & Shaughnessy, 1980; Leonesio & Nelson, 1990*). However, recent evidence has accumulated suggesting that perceptual information (e.g., font size, auditory volume) may be one type of cue that serves to inappropriately bias metamemory processes (e.g., *Kornell, Rhodes, Castel, & Tauber, 2011; McDonough & Gallo, 2012; Mueller, Dunlosky, Tauber, & Rhodes, 2014; Rhodes & Castel, 2008, 2009; Yue, Castel, & Bjork, 2013*).

A study by *Rhodes and Castel (2008)* manipulated perceptual features of stimuli by varying the font size of words (48- vs. 18-pt font) during encoding. Across several experiments JOLs were influenced by perceptual information such that words presented in a large font were perceived as more memorable than words presented in a small font despite showing no differences in final recall performance for the two class of stimuli (see also *Kornell et al., 2011; Mueller et al., 2014*). Biases in monitoring persisted after repeated study-test trials, when more diagnostic information could be used for judgments (i.e., item relatedness), when participants were told font size was unrelated to future memory performance, and when judgments of forgetting were made instead of JOLs. Similarly, *Rhodes and Castel (2009)* found that items presented in a loud auditory volume were given higher JOLs than items presented in a quiet volume even though there were no differences in recall performance between the two item types. Furthermore, items regarded as less memorable (quiet items) were more likely to be selected for restudy. These findings suggest that perceptual information may serve to inappropriately bias metamemory monitoring and control processes.

The dominant explanation for the perceptual effect on JOLs is that large (loud) and small (quiet) items differ in the ease with which they are processed. Processing ease, or *fluency*, refers to speed or probability with which information is perceived, encoded, or retrieved (*Undorf & Erdfelder, 2011*). Considerable research has demonstrated that information that is more easily processed as indicated by faster encoding latencies engender higher JOLs (e.g., *Hertzog, Dunlosky, Robinson, & Kidder, 2003; Koriat, 2008; Koriat, Ma'ayan, & Nussinson, 2006; Undorf & Erdfelder, 2011, 2013*). Thus, it is suggested that participants may misinterpret the ease of initial processing of large (loud) items as diagnostic of future ease of retrieval. Accordingly, when stimuli are made dysfluent by alternating between lowercase and uppercase letters (e.g., PiAnO), the difference in JOLs for large and small items is eliminated (*Rhodes & Castel, 2008*). However, because previous research examining perceptual effects on JOLs has not directly examined processing ease by measuring encoding latencies, it is possible that cues other than fluency may influence JOLs.

A recent study by *Mueller et al. (2014)* examined the relationship between processing fluency and JOLs for large and small

items by measuring study duration when study was self-paced. Consistent with previous research, large items engendered higher JOLs than small items. It is important, however, that there were no differences in study duration between the two items. Furthermore, even when the use of fluency as a cue was eliminated by having participants make JOLs prior to seeing the item, JOLs were higher when participants were informed that the study item would be large versus small. When asked why large items were predicted to be better remembered, participants most often claimed that they believed large items to be more distinct or draw more attention, rather than being easier to learn or read. Therefore, it was suggested that rather than because of processing fluency, the effect of font size on JOLs may reflect a form of analytical processing in which beliefs or theories of how the manipulation should influence memory are used to infer the likelihood of remembering. That is, learners may believe that perceptual saliency is a diagnostic cue of memorability and therefore give higher JOLs to more salient (e.g., large or loud) items. Thus, one of the primary motivations of the current study was to arbitrate between the two hypotheses that have been proposed to explain the perceptual effect on JOLs. In addition, we were also interested in how perceptual effects may influence JOLs in more educationally relevant materials.

Metacognitive Illusions and Educational Materials

Inaccuracies in metamemory judgments are often attributable in part to illusions of competence (or, foresight bias) when learners do not appreciate the important differences between information that will be present at encoding and information that will be present at retrieval (*Koriat & Bjork, 2005*). The persistence of inflated JOLs for items presented in large font in the *Rhodes and Castel (2008)* study suggests that learners fail to discount perceptual information at encoding that should not aid in performance at test, even after explicit warnings that font size is not related to future memory performance. Such findings are particularly troublesome when considering the study habits of students who may not be privy to such knowledge when studying for exams by reading through book chapters or PowerPoint lectures. Textbooks are often filled with perceptually salient information used to guide the reader toward key terms or figures that they should be knowledgeable about. One can find an assortment of different text fonts, sizes, bolding, colors, and other perceptual features in textbooks, often all on the same page. While this is undoubtedly a useful method to direct a reader's attention to important information, little is known about how perceptual information of educationally relevant material affects learners' perceived memorability of the to-be-remembered information. Currently, the results are inconclusive as to whether perceptual information can negatively influence performance.

Miele and Molden (2010) conducted an individual differences examination of perceptions of text fluency and memory. They had participants study texts presented in one of two formats (italicized vs. standard) and gave them a subsequent memory test for the passage. A subgroup of participants consistently regarded the easy-to-read texts as easier to comprehend despite little impact on their recall. In a different study, *Rawson and Dunlosky (2002)* had participants read a text in which words were intact or partially presented. Participants in this study predicted better comprehension and performance for intact items. However, memory for the

two item classes did not differ. Alternatively, a study by [Diemand-Yauman, Oppenheimer, and Vaughan \(2011\)](#) found that text presented in dysfluent typeface (e.g., Monotype Corsiva) was remembered better than text presented in normal typeface. That is, information that was processed more fluently was remembered worse. The memory advantage for dysfluent information was even extended to actual classroom environments. Thus, it is possible that the degree to which a student is overconfident in how well they have learned material may result in less effortful processing ([Alter, Oppenheimer, Epley, & Eyre, 2007](#)) or premature termination of study ([Koriat et al., 2006](#)), thus resulting in poor test performance.

Current Study

The purpose of the current study was to investigate whether or not metamemory illusions occur when perceptual information is manipulated in educationally relevant material. Although recent evidence has accumulated suggesting that perceptual metacognitive illusions may occur using common memory research materials as study items (i.e., random lists of unrelated nouns), there is little research examining how monitoring processes are affected by perceptual information when utilizing real-world learning materials. Furthermore, although processing fluency has been implicated as an important mechanism underlying the perceptual influence on JOLs, the majority of the extant research study duration was determined by the experimenter thereby eliminating the possibility for participants to use control processes to decide when to terminate study (e.g., [Kornell et al., 2011](#); [Rhodes & Castel, 2008](#); but see [Mueller et al., 2014](#)). Thus, we were also interested in examining whether processing fluency during self-regulated learning mediates the influence of perceptual information on JOLs should monitoring biases occur, or whether additional cues (e.g., beliefs) are utilized when assessing learning. It is important to note that although processing fluency can be characterized in multiple ways (see [Alter & Oppenheimer, 2009](#); [Kelley & Rhodes, 2002](#)), in the current study we define processing fluency as the speed with which information is learned as evidenced by study duration.

In the current study, participants studied a series of terms and their corresponding definitions taken from a psychology textbook. In Experiments 1, 2, and 3, we manipulated font bolding, font size, and borders surrounding text, respectively, all of which are commonly used in educational material. In Experiment 1, half of the participants received 7 s to study each item whereas the other half had unlimited time to allocate toward studying each item. In Experiments 2 and 3, participants were given unlimited study time. Because conceptual processing is the primary determinant of effective encoding ([Craik & Lockhart, 1972](#)), we expect perceptual information to have little impact on subsequent memory performance. However, perceptual information may nevertheless influence processing fluency, thereby resulting in differences in JOLs. Such findings would demonstrate that salient perceptual information provided in real-world learning materials may lead to inappropriate biases in metacognitive monitoring processes engaged during study.

Experiment 1A

In Experiment 1A, half of the terms (but not definitions) were bolded, whereas the other half of the terms were presented in

regular font. Furthermore, half of the participants received 7 s to study each item, whereas the other half had unlimited time to allocate toward studying each item (see the [Appendix](#) for examples of materials used in all experiments). It should be noted that this is a much less salient perceptual manipulation than has been used previously. It was hypothesized that perceptual information should have little influence on memory, but that participants in the unlimited time condition would recall more words than participants in the timed condition because of the ability to allocate their study time for each item and study more difficult items longer. If perceptual information influences beliefs, bolded items should engender higher JOLs than unbolded items regardless of whether encoding is self-paced or timed. Furthermore, there should be no differences in study duration when study is self-paced. However, if perceptual information influences processing fluency, higher JOLs should be found only when encoding is self-paced and should result in faster encoding latencies for bolded than unbolded items.

Method

Participants. Seventy-four Arizona State University undergraduate students enrolled in an introductory psychology course participated in the study. Students received course credit for their participation in the study. Thirty-five and 39 participants were randomly assigned to the unlimited and fixed time (7 s) study condition, respectively.¹ Participants were tested individually or in groups of up to 8 individuals.

Materials. Materials consisted of 40 terms and corresponding definitions from *The Science of Psychology: An Appreciative View*, an introductory psychology textbook ([King, 2010](#)). The terms and corresponding definitions were presented individually in random order. For each participant, the 40 items were randomly selected to be presented in bolded 24-pt Arial font (20 items) or unbolded 24-pt Arial font (20 items). The difference between the bolded and unbolded terms was not particularly salient (see the [Appendix](#)). All definitions were presented in regular (unbolded) 24-pt Arial font. All study and test stimuli were in black font and centered on a white background.

Experimental procedure. Participants were informed that the computer would present terms and corresponding definitions for them to study and that they would later complete a cued recall memory test. The instructions also explained that terminology would be presented in bolded or unbolded form, but all information was equally important to memorize. Participants in the timed condition were told that each term and corresponding definition would be presented for study on the computer screen for a fixed amount of time. Participants in the unlimited study time condition were told that they would have an unlimited amount of time to memorize the terminology and definition and were instructed to push the spacebar when they desired to terminate their study of a particular item. Prior to the term and definition being presented together, the term was presented in isolation (in regular, 26-pt unbolded font) and participants were instructed to indicate how familiar the item was. This was done to ensure that levels of familiarity did not influence judgments across item types.

¹ Pilot testing revealed that on average participants terminated study after approximately 9 s when they were allowed to determine their own study time.

After participants finished reading the instructions, they began the task. First, the term was presented in isolation and participants were to rate the item for familiarity from 1 to 5 (with 1 being *very unfamiliar* and 5 being *very familiar*). Subsequently, participants were presented with each term and corresponding definition. In the unlimited time condition participants were allowed to study the item for as long as the desired and pressed the spacebar when they wanted to terminate studying of an item. In the timed condition, studying terminated after 7 s passed. Once study termination occurred, participants rated their confidence (on a scale from 0% to 100%) that they would later be able to recall the term when provided with its corresponding definition. After participants typed their confidence rating, the computer presented the next term.

Following the presentation of all terms and definitions, participants in both conditions completed a filler task for 5 min that required them to type states of the United States. After the filler task, participants began the cued-recall test, during which they were given definitions and asked to type the corresponding term. After completing the memory test, participants were debriefed and thanked for their participation in the study.

Statistical procedure. For Experiment 1A (and all subsequent experiments), we report analyses for mean recall performance, JOLs, and study duration across both conditions and separately for each item type.

In addition, we performed a mediation analysis to examine whether processing fluency mediates the relationship between perceptual features and JOLs (for a similar analytic approach examining predictors of metacognitive monitoring or control processes, see Hines, Touron, & Hertzog, 2009; Tauber & Rhodes, 2012). Because mediation occurs at the item level (i.e., study duration and JOL occur for each item), we assessed mediation for lower level variables using a 1–1–1 design via multilevel structural equation modeling (MSEM) as described by Preacher, Zyphur, and Zhang (2010). The MSEM procedure estimates the direct effect of item type on both JOLs (*c*) and study duration (*a*), the direct effect of study duration on JOLs (*b*), and the indirect effect of item type on JOLs via study duration (*ab*). To conclude mediation is present, it must be demonstrated that bolding is predictive of study duration (*a*), study duration predicts JOLs after controlling for item type (*b*), and including study duration significantly reduces the relationship between item type and JOLs (*ab*). Full mediation occurs when the relationship between item type and JOLs (*c*) is eliminated after controlling for study duration, whereas partial mediation occurs when the relationship between item type and JOLs (*c*) remains after controlling study duration. The MSEM procedure also provides a measure of the dependency among observations (intraclass correlation coefficient, or ICC) that occur because of between-individual and within-individual variances in responses (Hox, 2002). Although there is no specified rule, an ICC greater than .10 suggests that multilevel modeling is justified compared with the conventional regression analysis that ignores the hierarchical structure of the data.

To align our results with much of the extant metamemory literature, we additionally report gamma correlations between recall and JOLs, study duration and JOLs, and study duration and recall. A gamma correlation is a nonparametric measure of association on an item-by-item basis that is calculated for each participant and aggregated across all participants in each condition (Nelson, 1984).

Results

Recall. To assess the effects of font bolding and study condition on recall (Figure 1), we submitted mean recall accuracy to a 2 (Bolding: bolded vs. unbolded) \times 2 (Condition: timed vs. unlimited) mixed-factorial analysis of variance (ANOVA). Recall performance was marginally better following unlimited study time than timed study time, $F(1, 72) = 3.82, p = .06, \eta^2 = .05$, but there was no effect of bolding on recall performance, $F(1, 72) < 1, p = .78$. There was a marginal interaction of bolding and condition, $F(1, 72) = 3.65, p = .06, \eta^2 = .04$. However, there were no differences in bolded and unbolded word recall for the timed or untimed conditions, $t(38) = 1.24, p = .22, d = .20, 95\% \text{ CI} [-.04, .43]$, and $t(34) = 1.44, p = .16, d = .25, 95\% \text{ CI} [.06, .43]$, respectively. Thus, although recall was better with unlimited study, there was no effect of bolding on memory performance.

Predictions. To assess the effects of font bolding and study condition on monitoring, we submitted JOLs to the same mixed-factorial ANOVA (Figure 1). Participants in the unlimited study time condition gave higher confidence ratings than participants in the timed condition, $F(1, 72) = 23.43, p < .001, \eta^2 = .25$. There was no main effect of bolding on confidence, $F(1, 72) = 2, p = .16, \eta^2 = .03$. However, there was a significant interaction between bolding and condition on confidence, $F(1, 72) = 3.87, p = .05, \eta^2 = .05$. In the unlimited condition, the mean confidence rating for bolded items was greater than for unbolded items, $t(34) = 2.86, p = .01, d = .48, 95\% \text{ CI} [.30, .67]$. However, in the timed condition, there were no differences in confidence ratings for bolded or unbolded items, $t < 1, p = .73, d = .06, 95\% \text{ CI} [-.09, .20]$. Thus, bolded items were given higher JOLs than unbolded items in the unlimited study time condition despite no differences in recall performance.

Study duration. To assess the influence of bolding information on processing fluency, we examined the amount of time allocated to studying bolded and unbolded items for participants in the untimed condition (Figure 2). This analysis revealed that less time was spent studying bolded than unbolded items, $F(1, 34) = 10.71, p = .002, d = .55, 95\% \text{ CI} [.36, .75]$.

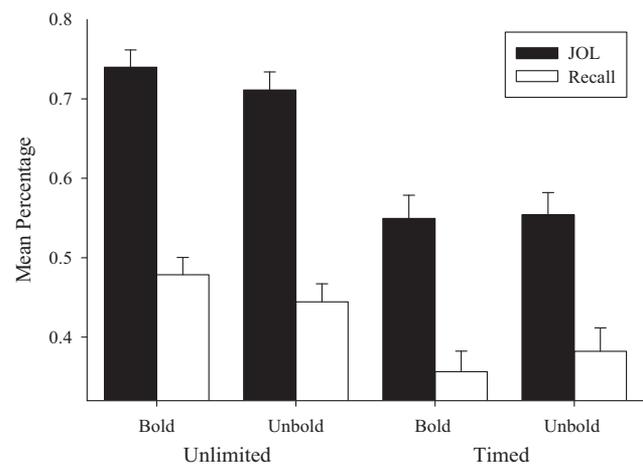


Figure 1. Predicted and actual recall performance across item type for both conditions in Experiment 1A. Error bars reflect SEs. JOL = judgment of learning.

Mediation analysis. Given that bolded items were studied for less time and given higher JOLs than unbolded items, we performed MSEM to assess whether study duration mediates the relationship between bolding and JOLs for participants in the untimed condition (see Method section for details). There was a moderate ICC for JOLs ($\rho = .219$), suggesting that multilevel modeling was justified. The analysis revealed that bolding was a significant predictor of study duration, $a = -1,282.35$, $SE(a) = 374.84$, $t(1393) = 3.42$, $p = .001$. Furthermore, study duration was a significant predictor of JOLs after controlling for bolding, $b = -.001$, $SE(b) = .00024$, $t(1393) = 4.23$, $p < .001$. It is important that the indirect effect of bolding on JOLs through study duration was significant, $ab = .876$, $SE(ab) = .314$, $t(1393) = 2.79$, $p = .01$, indicating that mediation was present. However, there was still a significant relationship between bolding and JOLs after controlling for study duration, $c = 2.29$, $SE(c) = .89$, $t(1393) = 2.58$, $p = .01$. Thus, study duration only partially mediated the relationship between bolding and JOLs.

Gamma correlations. To assess the predictive accuracy of judgments, we calculated gamma correlations between JOLs and recall accuracy separately for bolded and unbolded items across conditions (Table 1). Relative accuracy was reliably greater than zero in all cases (all $ps < .001$), suggesting that participants were generally accurate in their predictions. To examine whether accuracy differed across conditions, we submitted mean gammas to a 2 (Bolding: bolded vs. unbolded) \times 2 (Condition: timed vs. unlimited) mixed-factorial ANOVA.² This analysis revealed no effect of condition or bolding, $F(1, 70) < 1$, $p = .50$, $\eta^2 = .01$, and $F(1, 70) = 2$, $p = .16$, $\eta^2 = .03$. Furthermore, there was only a marginal interaction of bolding and condition, $F(1, 70) = 4.32$, $p = .09$, $\eta^2 = .042$. Thus, participants were generally accurate in their predictions but this did not differ as a function of item type or study condition.

We also examined the relationship of study duration with both recall and JOLs (Table 1). There was a strong negative correlation between study duration with both recall and JOLs that reliably differed from zero ($ps < .001$). Thus, consistent with previous research (e.g., Koriat et al., 2006), items processed more easily were given higher JOLs and better remembered.

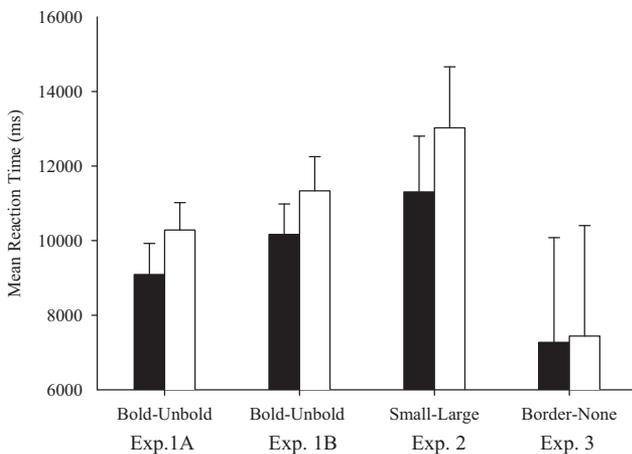


Figure 2. Average time spent studying different item types for each experiment. Error bars reflect SEs.

Table 1
Gamma Correlations (SEs) Between Metacognitive Judgments, Memory Performance, and Study Duration in Experiments 1–3

Experiment	Condition	Item type	Gamma correlation			
			JOL, Recall	Recall, Fluency	JOL, Fluency	
1A	Untimed	Bold	.48 (.06)			
		Unbold	.66 (.04)			
		Aggregate		-.40 (.05)	-.53 (.04)	
	Timed	Bold	.54 (.06)			
		Unbold	.52 (.06)			
		Aggregate		—	—	
1B	Untimed	Bold	.58 (.05)			
		Unbold	.59 (.06)			
		Aggregate		-.37 (.05)	-.56 (.04)	
	2	Untimed	Small	.52 (.07)		
			Large	.59 (.07)		
			Aggregate		-.34 (.08)	-.37 (.05)
3	Untimed	Border	.58 (.05)			
		No Border	.57 (.05)			
		Aggregate		-.22 (.06)	-.29 (.05)	

Note. All correlations reliably differ from zero. “Aggregate” reflects gamma correlations computed across all items, regardless of item type.

Familiarity. Finally, although not the primary goal of the present study, we also examined whether any baseline differences in familiarity ratings were apparent for items that were (subsequently) presented in bolded or unbolded font to a 2 (Bolding: bolded vs. unbolded) \times 2 (Condition: timed vs. unlimited) mixed-factorial ANOVA. There was no effect of bolding, $F(1, 72) < 1$, $p = .50$, $\eta^2 = .05$, nor an interaction of bolding and condition, $F(1, 72) < 1$, $p = .48$, $\eta^2 = .01$. Thus, differences in JOLs cannot be because of baseline differences in familiarity for the items between items or across conditions. There were also no significant baseline differences in any of the subsequent experiments (all $ps > .05$) and therefore these analyses will not be mentioned further. However, it is possible that making familiarity judgments prior to studying the items may have influenced JOLs. That is, if a participant determined that an item was unfamiliar they subsequently may have given a lower JOL for the item regardless of actual memorability. Alternatively, the extra presentation of the term may have influenced subsequent recall performance. Therefore, before continuing to the discussion section we first report the results from Experiment 1B where we evaluated JOLs using the identical design as in the untimed condition of Experiment 1A, with the exception that familiarity judgments were not collected.

² Two participants in the untimed condition had perfect recall performance. Consequently, the differences in degrees of freedom between the mean analyses and gamma correlation analyses occurs because participants with perfect performance were not included in the latter because the ordinal association between two variables cannot be computed when there are no discordances. Note that this issue occurred in subsequent experiments as well.

Experiment 1B

Method

Participants. Thirty Arizona State University undergraduate psychology students participated for partial course credit. One participant was removed from the analyses because of 0% performance on the cued-recall test, resulting in 29 participants.

Material, design, and procedure. The materials and procedure for Experiment 1B were identical to that of Experiment 1A except that no familiarity ratings were made on the cue prior to studying the term and its corresponding definition. Furthermore, study time was always unlimited and terminated by the participant (i.e., there was no 7-s timed condition).

Results

Recall and predictions. To assess the effects of font bolding on memory and monitoring, recall performance and JOLs were submitted to separate repeated measures ANOVAs (Figure 3). There were no difference in recall performance across item types, $F(1, 28) = 2.77, p = .12, d = .31, 95\% \text{ CI } [.12, .50]$. However, confidence ratings were greater for bolded than unbolded items, $F(1, 28) = 9.03, p = .01, d = .56, 95\% \text{ CI } [.36, .75]$. Thus, as with Experiment 1A, bolded items received higher JOLs despite the fact that both item types were remembered equally well.

Study duration. To assess the effect of bolding on processing fluency, study durations for bolded and unbolded items were submitted to a repeated measures ANOVA (Figure 2). This analysis revealed that participants spent less time studying bolded items than unbolded items, $F(1, 28) = 4.38, p = .05, d = .39, 95\% \text{ CI } [.17, .61]$.

Mediation analysis. As with Experiment 1A, we performed MSEM to assess whether study duration mediates the relationship between bolding and JOLs. There was a moderate ICC for JOLs ($\rho = .307$), suggesting that multilevel modeling was justified. The analysis revealed that bolding was a significant predictor of study duration, $a = -993.17, SE(a) = 466.67, t(1158) = 2.13, p = .03$.

Furthermore, study duration was predictive of JOLs after accounting for bolding, $b = -.001, SE(b) = .00026, t(1158) = 3.89, p < .001$. It is important that the indirect effect of font size on JOLs through study duration was significant, $ab = .673, SE(ab) = .355, t(1158) = 1.90, p = .05$, suggesting that mediation was present. However, there was still a significant relation between bolding and JOLs after controlling for study duration, $c = 2.52, SE(c) = .95, t(1158) = 2.66, p = .01$. Thus, study duration only partially mediated the relationship between bolding and JOLs.

Gamma correlations. Gamma correlations between JOLs, recall, and study duration are displayed in Table 1. The correlation between JOLs and recall for both bolded and unbolded items reliably differed from zero ($ps < .001$), but did not differ as a function of item type, $F(1, 28) < 1, p = .82, d = .04, 95\% \text{ CI } [-.45, .54]$. In addition, there was a strong negative correlation between study duration with both recall and JOLs that reliably differed from zero ($ps < .001$), indicating that items processed more easily were given higher JOLs and better remembered.

Discussion of Experiments 1A and 1B

The results from Experiment 1A and 1B demonstrate the presence of metacognitive illusions using educationally relevant material and a much less salient perceptual manipulation than has been used previously. When participants were allowed to control the amount of time allocated to each item during study, items presented in bold were mistakenly regarded as more memorable than unbolded items. Furthermore, in the untimed condition bolded items were studied for less amount of time than unbolded items. Although it could be argued that JOLs and/or recall performance may have been influenced by familiarity judgments, participants demonstrated a similar pattern of results for bolded items in Experiment 1B that did not include familiarity judgments. The finding that bolded items were studied less and given higher JOLs suggests that participants provides strong support for the fluency hypothesis, and suggests that participants may have mistaken the ease of processing as diagnostic of later ease of retrieval that resulted in inflated JOLs for bold relative to unbolded items (Rhodes & Castel, 2008). However, mediation analyses revealed that processing fluency only partially mediated the relationship between bolding and JOLs, such that the relationship between bolding and JOLs remained even after controlling for study duration. Thus, in addition to processing fluency, participants may have relied on beliefs of how perceptual information may influence memory when making JOLs (Mueller et al., 2014). Previous research has similarly demonstrated that participants may rely on both observable item characteristics (e.g., item concreteness) and fluency (e.g., image generation latencies) when assessing learning (Hertzog et al., 2003). It is interesting, however, there were no differences in JOLs across item type in the timed condition of Experiment 1A when participants could have presumably still relied on beliefs. Therefore, when only given a limited time to study information participants may have only focused on more relevant information associated with future retrievability resulting in more effective monitoring for bolded items. That is, had we extended study time further (beyond the amount of time it takes to read the term and definition), more beliefs about perceptual features involved with the terms (i.e., bolding) may have played a larger role in JOL decision. Nevertheless, the results from these

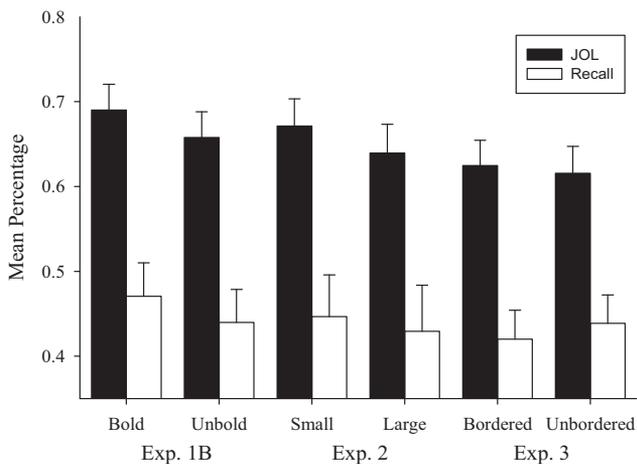


Figure 3. Predicted and actual recall performance across item types in Experiments 1B, 2, and 3. Error bars reflect SEs. JOL = judgment of learning.

experiments suggest that subtle differences in perceptual information can serve to inappropriately bias metacognitive monitoring and control processes. Experiment 2 was designed to test whether or not metacognitive illusions occur using educational material with a much more salient perceptual manipulation utilized in previous research.

Experiment 2

Previous research suggests that items presented in larger fonts consistently receive higher JOLs relative to smaller fonts, despite showing no differences in recall between the two item types. Although the font size effect has often been interpreted as reflecting increased processing fluency for larger items (e.g., Kornell et al., 2011; Rhodes & Castel, 2008), Mueller et al. (2014) found no differences in study duration across item types even though larger items were given higher JOLs. Thus, in Experiment 2 we presented both terms and definitions in large and small font sizes with unlimited study time to further explore how perceptual information influences JOLs and study duration with educationally relevant material.

Method

Participants. Thirty Arizona State University undergraduate psychology students participated for partial course credit. One participant was removed from the analyses because of 0% performance on the cued-recall test, resulting in 29 participants.

Materials, design, and procedure. Materials were identical to those used in Experiment 1A. The only difference is that for each participant, 20 terms and corresponding definitions were randomly presented in regular (unbolded) 24-pt Arial font, whereas the other 20 terms and corresponding definitions were presented in regular 48-pt Arial font (Appendix). Terms presented in isolation during familiarity ratings were presented in an intermediate regular 36-pt Arial font. All study and test stimuli were in black font and centered on a white background. The procedure for Experiment 2 was identical to that of Experiment 1A except that study time was always unlimited and terminated by the participant.

Results

Recall and predictions. To assess the effects of font size on memory and monitoring, recall performance and JOLs were submitted to separate repeated measures ANOVAs (Figure 3). There was no difference in overall recall performance for small and large items, $F(1, 28) < 1$, $p = .52$, $d = .12$, 95% CI [-0.06, .31]. However, confidence ratings were greater for small than for large items, $F(1, 28) = 3.98$, $p = .05$, $d = .37$, 95% CI [.17, .57].

Study duration. To assess of font size on processing fluency, study durations for small and large items were submitted to a repeated measures ANOVA (Figure 2). This analysis revealed that participants spent longer studying large items than small items, $F(1, 28) = 10.89$, $p = .003$, $d = .61$, 95% CI [.41, .81].

Mediation analysis. We performed MSEM to assess whether study duration mediates the relationship between font size and JOLs. There was a moderate ICC for JOLs ($\rho = .296$), suggesting that multilevel modeling was justified. The analysis revealed that font size was a significant predictor of study duration, $a =$

1718.95 , $SE(a) = 512.11$, $t(1158) = 3.36$, $p = .001$, and that study duration was a significant predictor of JOLs after controlling for font size, $b = .001$, $SE(b) = .00025$, $t(1158) = 4.08$, $p < .001$. It is important that the indirect effect of font size on JOLs through study duration was significant, $ab = -.670$, $SE(ab) = .227$, $t(1158) = 2.96$, $p = .003$, and controlling for study duration eliminated the relationship between font size and JOLs, $c = -2.52$, $SE(c) = 1.57$, $t(1158) = 1.60$, $p = .11$. These results suggest that study duration fully mediated the relationship between font size and JOLs.

Gamma correlations. Gamma correlations between JOLs, recall, and study duration are displayed in Table 1. Predictive accuracy of JOLs was reliably greater than zero for both small and large items ($ps < .001$), but did not differ as a function of item type, $F(1, 27) < 1$, $p = .56$, $d = .11$, 95% CI [-0.33, .56]. Furthermore, there was a strong negative correlation between study duration with both recall and JOLs that reliably differed from zero ($ps < .001$), indicating that items processed more easily were given higher JOLs and better remembered.

Discussion

The results from Experiment 2 extend the findings of Experiment 1 demonstrating that font size does not influence memory performance but does bias metacognitive monitoring and control processes. Interestingly, contrary to previous research (e.g., Kornell et al., 2011; Rhodes & Castel, 2008) the results from the current experiment found that small items were actually perceived as more memorable than large items. Furthermore, small items were studied less than large items. In conjunction with the SEM analyses showing that study duration fully mediated the relationship between font size and JOLs, these results suggest that participants relied primarily on processing fluency to inform JOLs rather than beliefs about how perceptual information should influence memorability. Thus, participants may have mistaken the ease with which small items were processed as diagnostic of future memorability. We suspect that the opposite direction of the font size effect compared with previous research is likely because of the amount of information to be studied on each trial. That is, the longer study times for large items presumably reflects the amount of time necessary to read the information that subtended more of the visual field while learning both a term and its definition, which is likely not an issue when studying words or word pairs that are typically used. Regardless, the results from Experiments 1 and 2 demonstrate that both subtle and salient perceptual information, respectively, bias monitoring processes in educationally relevant material. Thus, Experiment 3 was designed to determine whether any form of perceptual saliency influences perceived memorability, or if this effect is limited only to perceptual features that affect study duration.

Experiment 3

In previous research investigating metacognitive illusions (as well as in Experiments 1 and 2 of the current study) the perceptual features that have been manipulated were inherent to the item itself. That is, within-item features (e.g., font size or type) of the to-be-remembered information varied in its perceptual saliency. However, it is reasonable to assume that external perceptual fea-

tures (e.g., background color) may also influence metamemory processes. Given that ubiquity of tables and figures in textbooks and other education material, in Experiment 3 we manipulated external perceptual features by presenting a border around half of the items while holding internal features (i.e., font) constant across all items. It is possible that beliefs about how borders influence memorability may result produce differences in JOLs across item types. However, based on the results from the previous studies, we expect that borders surrounding the text should have little influence on the ease with which the actual study *item* is processed and therefore not influence JOLs.

Method

Participants. Thirty-five Arizona State University undergraduate psychology students participated for partial course credit.

Materials, design, and procedure. Materials and procedure were identical to those used in Experiment 2, except that all terms and definitions were presented in regular (unbolded) 24-pt Arial font. All study and test stimuli were in black font and centered on a white background. However, 20 terms and definitions were surrounded by a light blue border with black outlines (similar to a picture frame), whereas the other 20 items were presented without a border (Appendix). Terms presented in isolation during familiarity ratings were presented in regular 24-pt Arial font without a border.

Results

Recall, predictions, and study duration. To assess the effects of bordering on performance, recall performance, JOLs, and study duration for bordered and unbordered items were submitted to separate repeated measures ANOVAs (Figure 3). These analyses revealed no differences for bordered and unbordered items in recall, $F(1, 34) < 1, p = .49, d = .12, 95\% \text{ CI } [-.15, .38]$, confidence ratings, $F(1, 34) < 1, p = .64, d = .08, 95\% \text{ CI } [-.12, .28]$, or study duration, $F(1, 34) < 1, p = .60, d = .09, 95\% \text{ CI } [-.13, .31]$.

Mediation analysis. Although there were no differences in JOLs or study duration across item types, we performed MSEM to assess the relationship between bordering, study duration, and JOLs. There was a moderate ICC for JOLs ($\rho = .279$), suggesting that multilevel modeling was justified. The analysis revealed that bordering was neither predictive of study duration nor JOLs, and that the indirect effect of bordering on JOLs through study duration was not significant, $ts < 1, ps > .60$. Thus, as expected from the mean analyses, there was no evidence of mediation. However, consistent with the previous experiments study duration was a significant predictor of JOLs after controlling for bordering, $b = -.001, SE(b) = .00038, t(1393) = 2.66, p = .01$.

Gamma correlations. Gamma correlations between JOLs, recall, and study duration are displayed in Table 1. Relative accuracy was reliably greater than zero for both bordered and unbordered items ($ps < .001$), but did not differ as a function of item type, $F(1, 32) < 1, p = .85, d = .07, 95\% \text{ CI } [-.35, .49]$. Furthermore, there was a strong negative correlation between study duration with both recall and JOLs that reliably differed from zero ($ps < .001$), indicating that items processed more easily processed were given higher JOLs and better remembered.

Discussion

The results from Experiment 3 converge with those of Experiments 1 and 2 in the finding that perceptual information has little impact on subsequent recall performance. However, these results also demonstrate that external perceptual features have little influence on perceived memorability of learned information, with no differences in JOLs for bordered or unbordered items. It is interesting that there were also no differences in the amount of time allocated to studying either item type. Equivalent study duration for bordered and unbordered items suggests that it is not necessarily the beliefs about how perceptually salient features may influence memory that results in metacognitive illusions, per se, but rather biases in monitoring may be a result of how perceptual information actually influences processing fluency. As a consequence, because there were no differences in study duration across item types, participants did not perceive one item type to be particularly more memorable than the other.

General Discussion

The current study sought to examine how perceptual information influences metacognitive monitoring and control processes using educationally relevant material. The results consistently demonstrated that perceptual information had little impact on subsequent memory performance. Nevertheless, when participants were allowed to control their study time, items that were processed more easily (as evidenced by study duration) consistently resulted in monitoring biases. Together, these results suggest that participants may mistake the ease in which educationally relevant information is processed as diagnostic of actual memorability despite the fact that perceptual features have little impact on memory performance (Rhodes & Castel, 2008). Although metacognitive illusions from perceptual information have been found in previous research, such research used study scenarios that are dissimilar to real-life studying. The current findings suggest that when given control of study duration people may often misuse perceptual information when evaluating how well they have learned material when studying educational materials (e.g., bolding in text or font size in note slides), which may ultimately lead to inaccurate beliefs about one's own learning and future memory performance.

Belief-Based Versus Fluency-Based Monitoring

In the current study, participants may have used perceptual saliency to make analytic inferences of future recallability (i.e., belief-based JOLs), whereas processing fluency experienced during the task could be used to make implicit inferences about future memory (i.e., fluency-based JOLs). The results suggest that reliance beliefs of how perceptual information may influence performance may not be the primary reason for monitoring differences across item types (Mueller et al., 2014). Bolded, large, and bordered items are arguably more salient than unbolded, small, and unbordered items, respectively, and should consequently lead to the belief that the salient items should be more memorable. However, despite identical study materials between the timed and untimed conditions in Experiment 1A, no monitoring differences occurred when participants were unable to control study duration. Furthermore, large items actually engendered *lower* JOLs than

small items in Experiment 2, and there were no monitoring differences between bordered and unbordered items in Experiment 3. However, consistent across experiments was that higher JOLs were given to item types that received less study, and study duration mediated the relationship between item type and JOLs. These findings suggest that the assessment of the mnemonic cue of how easily information was encoded was the primary influence on monitoring errors across experiments (Begg et al., 1989; Hertzog, Dunlosky, Robinson, & Kidder, 2003; Koriat, 2008; Undorf & Erdfelder, 2011).

The results from the current study, however, are not entirely inconsistent with an analytic processing account of how perceptual information influences JOLs. Although processing fluency was predictive of JOLs in Experiments 1A and 1B, the relationship between item type and JOLs remained even after controlling for study duration. Thus, beyond an implicit attribution of processing fluency to perceived memorability, participants may have also relied on analytic processes when making JOLs (for similar findings see Hertzog et al., 2003). For example, given that bolded information in textbooks is generally deemed important because it is often the subject of later testing, participants may misattribute perceived importance from prior learning episodes as diagnostic of future memorability. Alternatively, perceived variation in processing speed across item types may nevertheless trigger analytic processes that lead people to believe that more quickly processed information should be easier to remember irrespective of implicit processes (Mueller et al., 2014). That is, an a priori belief that large items (Experiment 2) should be more memorable than small items may quickly be overshadowed by the relative processing difficulty across item types, leading participants to discount the validity of the prior belief online and instead rely on a heuristic that easier to process items should presumably be more easy to remember. However, previous research has demonstrated that in contrast to when assessing one's own learning, study duration is not predictive of JOLs when the participant is instructed to make JOLs for another participant while watching their study session (Undorf & Erdfelder, 2011, 2013). That is, even though participants could see how long it took the other participant to learn the item, study duration was not predictive of performance presumably because subjective feelings of processing fluency could not be used to inform JOLs. Although we favor the processing fluency over the analytical processing account of the current findings, it is not possible to fully dissociate the two because we did not assess participants' beliefs. Thus, future research should assess both processing fluency and beliefs when examining the influence of perceptual information on JOLs.

The reasoning behind the conflicting results between the Mueller et al. (2014) study and the current study are not entirely clear. As mentioned previously, Mueller et al. found that larger items were given higher JOLs than smaller items despite no differences in study duration between the two. Furthermore, they did not find the typical negative correlation between study duration and JOLs (e.g., Hertzog et al., 2003; Koriat, 2008; Koriat et al., 2006; Undorf & Erdfelder, 2011, 2013). It may be the case that perceptual information has a larger influence on processing fluency with the longer study materials as used in the current study (see Miele & Molden, 2010). Thus, similar to Mueller et al., had we measured study duration for *only* the terms in Experiments 1A and 1B it is possible that only a small (nonsignificant) difference would

emerge across item types. However, if this small numerical difference carries over to processing of the (unbolded) definition, it may actually produce a large difference in *overall* encoding time given the length of the definition to be studied. In any manner, given the applied ramifications of how perceptual information can serve to negatively influence perceived learning and study termination decisions, an important endeavor for future research is to better specify under what conditions perceptual information does or does not influence processing fluency.

Self-Regulated Learning

Although we were primarily interested in the influence of perceptual information on monitoring processes, the results from the current study are also informative of how learners preferentially allocate study time. Many of the extant theories of self-regulated learning suggest that learners have a predetermined level of mastery of material that is desired, and that control processes to terminate study are directly influenced by processes that monitor the ease with which information is learned (Nelson & Leonesio, 1988; Nelson & Narens, 1990; but see Koriat et al., 2006). For example, the *region of proximal learning* (RPL) theory posits that study time is preferentially allocated to items that are just beyond the learner's current understanding and would therefore benefit from additional study (Metcalfe & Kornell, 2003, 2005). Thus, the RPL theory predicts that more study time should be devoted to items of intermediate difficulty, as easy items should benefit from only minimal study and difficult items may never be learned regardless of additional study. Alternatively, the *discrepancy reduction* theory posits that learners try to reduce the discrepancy between their current state of learning and a desired level of mastery, thereby devoting more to difficult items because the discrepancy between current and desired levels of mastery is the greatest (Thiede & Dunlosky, 1999). In accord with the discrepancy reduction theory, results from Experiments 1A and 1B found that items regarded as less memorable received more study. It is important, however, the additional study was "labor-in-vain" (Nelson & Leonesio, 1988) because it did not improve subsequent memory. These findings suggest that errors in calibration can negatively affect control decisions even during situations in which relative accuracy is reasonably high (e.g., Metcalfe & Finn, 2008; Miele, Finn, & Molden, 2011; Rhodes & Castel, 2009).

In contrast to the aforementioned theories, Koriat and colleagues have suggested that learners do not necessarily have a predetermined level of mastery that is desired, but rather item difficulty is monitored based on the control decision to terminate study (Koriat, 2008; Koriat et al., 2006). More specifically, learners devote as much time as necessary to encode an item and the realization that an item requires more effort to commit to memory (i.e., memorizing effort) allows the learner to appreciate that the item will be difficult to recall. The results from Experiments 1 and 2 found that regardless of perceptual characteristics, items that received more study resulted in lower JOLs. Furthermore, items that were studied for longer were actually remembered worse. Thus, presumably participants relied on an "easily learned, easily remembered" heuristic in which they assessed the amount of time and effort invested in learning the item and used this information to (correctly) predict that these items would be more difficult to remember. Such findings suggest that biases in monitoring are not necessarily because

of faulty beliefs by the participant, but rather the actual ease with which information is processed leads to a subjective feeling that this information will be easier to remember (Koriat & Bjork, 2006).

Metacognition in Ecologically Valid Settings

From a practical standpoint, these findings suggest that learners may often be inaccurate when assessing how well information is learned and will be subsequently retrieved. Previous research in the educational domain suggests students may engage in a variety of study techniques that appear to be more effective than they really are. For example, a prominent study technique—rereading the material (Karpicke, Butler, & Roediger, 2009)—can create the illusion of enhanced learning by increasing processing fluency whereby successive rereading attempts become faster and easier. However, fluency does not directly translate to enhanced subsequent memory performance. Similarly, highlighting material upon initial learning creates the illusion of elaborative processing but has been shown to be no more effective for subsequent memory performance than reading without highlighting (Peterson, 1992). The results from the current study suggest that even subtle perceptual manipulations in textbooks and other educational material may serve to inappropriately lead people to believe information is better learned.

Although we did not find evidence that processing fluency influenced subsequent memory performance, this is not always the case (e.g., Sungkhasettee, Friedman, & Castel, 2011). As mentioned previously, Diemand-Yauman et al. (2011) found that more fluently processed information was actually remembered worse in both controlled experimental and classroom settings. Although the authors did not measure JOLs, previous research has found that participants are more confident in their ability to solve problems in easy-to-read fonts than hard-to-read fonts, despite the fact that performance for hard-to-read material is actually better because of deeper processing strategies (Alter, Oppenheimer, Epley, & Eyre, 2007). Thus, it has been suggested that perceptual *disfluency* may act as a “desirable difficulty” (Bjork, 1994), whereby information that is more difficult to process triggers analytic reasoning processes that facilitate long-term retention of learned information (Alter et al., 2007). Thus, one possible negative ramification of perceptual manipulations in educational material that increases processing fluency is that learners may believe they have mastered the material resulting in less effortful processing or premature termination of study.

Although the results from Diemand-Yauman et al. (2011) might suggest that a beneficial practice for educators would be to present materials in fonts or typefaces that are more difficult to read, more recent studies have either found no influence of fluency on subsequent memory using both educationally relevant (e.g., Carpenter, Wilford, Kornell, & Mullaney, 2013; Eitel, Köhl, Scheiter, & Gerjets, in press; Miele & Molden, 2010; Serra & Dunlosky, 2010) and common memory research (e.g., Kornell et al., 2011; McDonough & Gallo, 2012; Mueller et al., 2014; Rhodes & Castel, 2008, 2009) materials, or that disfluency may even *hurt* performance (e.g., Rawson & Dunlosky, 2002; Yue et al., 2013). Given the complexities of monitoring and control dynamics involved in learning, we suspect that minor differences in manipulated perceptual features (see Yue et al., 2013; Sungkhasettee et al., 2011) and

experimental designs across studies likely produce important differences in encoding and retrieval processes. For example, Diemand-Yauman et al. manipulated perceptual features between-subjects and did not allow for self-controlled encoding, thereby eliminating the possibility to use perceived fluency between easy- and hard-to-read fonts as a basis for learning and study termination decisions. In addition, the authors did not require participants to make assessments of future memory performance. Recent unpublished research from our laboratory has indicated that making metacognitive judgments fundamentally alters encoding and retrieval processes depending on learning conditions and the nature of the subsequent memory test and that predictions can be used as a method to enhance subsequent memory performance (see also Dougherty, Scheck, Nelson, & Narens, 2005; Spellman & Bjork, 1992; Yue et al., 2013). Thus, when making online (overt) assessments of how well information is learned across item types that vary in perceptual characteristics, participants may engage similar processing strategies across item types but nevertheless spend longer studying more difficult to process items to try to equate memory performance for both items. In any manner, we suggest that an important avenue for future research is to determine under what conditions fluency helps/hinders subsequent memory for learned information in order to better inform educators of best practices to implement in classroom environments.

Conclusions

The primary purpose of the present was to create a study scenario that mimics real-life study situations. Study items were commonly tested educational materials that featured a variety of perceptual manipulations (bolding, font size, borders) that are often present in textbooks and other educationally relevant materials. In addition, across experiments participants were allowed to preferentially allocate study time to certain items and make the metacognitive control decision to terminate study, as is usually the case when studying occurs in everyday situations. These findings suggest that when participants are allowed to control study time certain perceptual information may influence the ease in which information is processed, resulting in shorter study duration and enhanced perceived memorability. Given the ubiquity of perceptually salient features in educationally relevant material that may influence encoding fluency, it is important for educators to inform students of the possibilities of illusions of learning and suggest alternative methods to avoid such biases.

References

- Alter, A. L., & Oppenheimer, D. M. (2009). Uniting the tribes of fluency to form a metacognitive nation. *Personality and Social Psychology Review: An Official Journal of the Society for Personality and Social Psychology*, *13*, 219–235.
- Alter, A. L., Oppenheimer, D. M., Epley, N., & Eyre, R. N. (2007). Overcoming intuition: Metacognitive difficulty activates analytic reasoning. *Journal of Experimental Psychology: General*, *136*, 569–576. doi:10.1037/0096-3445.136.4.569
- Arbuckle, T. Y., & Cuddy, L. L. (1969). Discrimination of item strength at time of presentation. *Journal of Experimental Psychology*, *81*, 126–131. doi:10.1037/h0027455
- Begg, I., Duft, S., Lalonde, P., Melnick, R., & Sanvito, J. (1989). Memory predictions are based on ease of processing. *Journal of Memory and Language*, *28*, 610–632. doi:10.1016/0749-596X(89)90016-8

- Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe & A. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 185–205). Cambridge, MA: MIT Press.
- Carpenter, S. K., Wilford, M. M., Kornell, N., & Mullaney, K. M. (2013). Appearances can be deceiving: Instructor fluency increases perceptions of learning without increasing actual learning. *Psychonomic Bulletin & Review*, *20*, 1350–1356. doi:10.3758/s13423-013-0442-z
- Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning & Verbal Behavior*, *11*, 671–684. doi:10.1016/S0022-5371(72)80001-X
- Diemand-Yauman, C., Oppenheimer, D. M., & Vaughan, E. B. (2011). Fortune favors the bold (and the italicized): Effects of disfluency on educational outcomes. *Cognition*, *118*, 111–115. doi:10.1016/j.cognition.2010.09.012
- Dougherty, M. R., Scheck, P., Nelson, T. O., & Narens, L. (2005). Using the past to predict the future. *Memory & Cognition*, *33*, 1096–1115. doi:10.3758/BF03193216
- Dunlosky, J., & Metcalfe, J. (2009). *Metacognition*. Thousand Oaks, CA: Sage
- Eitel, A., Kühl, T. I. M., Scheiter, K., & Gerjets, P. (in press). Disfluency meets cognitive load in multimedia learning: Does harder-to-read mean better-to-understand? *Applied Cognitive Psychology*.
- Hart, J. T. (1965). Memory and the feeling of knowing experience. *Journal of Educational Psychology*, *56*, 208–216. doi:10.1037/h0022263
- Hart, J. T. (1967). Memory and the memory-monitoring process. *Journal of Verbal Learning and Verbal Behavior*, *6*, 685–691.
- Hertzog, C., Dunlosky, J., Robinson, A. E., & Kidder, D. P. (2003). Encoding fluency is a cue used for judgments about learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 22–34. doi:10.1037/0278-7393.29.1.22
- Hines, J. C., Touron, D. R., & Hertzog, C. (2009). Metacognitive influences on study time allocation in an associative recognition task: An analysis of adult age differences. *Psychology and Aging*, *24*, 462–475. doi:10.1037/a0014417
- Hox, J. J. (2002). *Multilevel analysis: Techniques and applications*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Karpicke, J. D., Butler, A. C., & Roediger, H. L. (2009). Metacognitive strategies in student learning: Do students practice retrieval when they study on their own? *Memory*, *17*, 471–479. doi:10.1080/09658210802647009
- Kelley, C. M., & Rhodes, M. G. (2002). Making sense and nonsense of experience: Attributions in memory and judgment. In B. Ross (Ed.), *The psychology of learning & motivation* (pp. 293–320). New York, NY: Academic Press. doi:10.1016/S0079-7421(02)80010-X
- King, J. E., Zechmeister, E. B., & Shaughnessy, J. J. (1980). Judgments of knowing: The influence of retrieval practice. *American Journal of Psychology*, *93*, 329–343. doi:10.2307/1422236
- King, L. (2010). *The science of psychology: An appreciative view*. New York, NY: McGraw-Hill.
- Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to JOLs. *Journal of Experimental Psychology: General*, *126*, 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*, 416–428. doi:10.3758/MC.36.2.416
- 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*, 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*, 1133–1145. doi:10.1037/0278-7393.32.5.1133
- 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*, 36–69. doi:10.1037/0096-3445.135.1.36
- 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*, 787–794. doi:10.1177/0956797611407929
- Leonesio, R. J., & Nelson, T. O. (1990). Do different metamemory judgments tap the same underlying aspects-of memory? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 464–467. doi:10.1037/0278-7393.16.3.464
- Mazzoni, G., & Comolli, C. (1993). Strategies in study time allocation: Why is study time sometimes not effective? *Journal of Experimental Psychology: General*, *122*, 47–60. doi:10.1037/0096-3445.122.1.47
- Mazzoni, G., Cornoldi, C., & Marchitelli, G. (1990). Do memorability ratings affect study-time allocation? *Memory & Cognition*, *18*, 196–204. doi:10.3758/BF03197095
- McDonough, I. M., & Gallo, D. A. (2012). Illusory expectations can affect retrieval-monitoring accuracy. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*, 391–404. doi:10.1037/a0025548
- Metcalfe, J., & Dunlosky, J. (2008). Metamemory. In H. L. Roediger, III (Ed.), *Learning and memory: A comprehensive reference* (pp. 349–362). Oxford, England: Elsevier.
- Metcalfe, J., & Finn, B. (2008). Evidence that judgments of learning are causally related to study choice. *Psychonomic Bulletin & Review*, *15*, 174–179. doi:10.3758/PBR.15.1.174
- Metcalfe, J., & Kornell, N. (2003). The dynamics of learning and allocation of study time to a region of proximal learning. *Journal of Experimental Psychology: General*, *132*, 530–542. doi:10.1037/0096-3445.132.4.530
- Metcalfe, J., & Kornell, N. (2005). A region of proximal learning model of study time allocation. *Journal of Memory and Language*, *52*, 463–477. doi:10.1016/j.jml.2004.12.001
- Metcalfe, J. & Finn, B. (in press). Metacognition and control of study choice in children. *Metacognition and Learning*.
- 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*, 320–324. doi:10.1177/0956797610397954
- Miele, D. B., & Molden, D. C. (2010). Naive theories of intelligence and the role of processing fluency in perceived comprehension. *Journal of Experimental Psychology: General*, *139*, 535–557. doi:10.1037/a0019745
- Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of behavior*. New York, NY: Henry Holt. doi:10.1037/10039-000
- 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
- Nelson, T. O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychological Bulletin*, *95*, 109–133. doi:10.1037/0033-2909.95.1.109
- Nelson, T. O., & Leonesio, R. J. (1988). Allocation of self-paced study time and the “labor-in-vain effect.” *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 676–686. doi:10.1037/0278-7393.14.4.676
- Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. Bower (Ed.), *The psychology of learning and motivation* (pp. 125–173). New York, NY: Academic Press. doi:10.1016/S0079-7421(08)60053-5
- Nelson, T. O., & Narens, L. (1994). Why investigate metacognition? In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 1–25). Cambridge, MA: MIT Press.

- Peterson, S. (1992). The cognitive function of underlining as a study technique. *Reading Research and Instruction, 31*, 49–56. doi:10.1080/19388079209558078
- Preacher, K. J., Zyphur, M. J., & Zhang, Z. (2010). A general multilevel SEM framework for assessing multilevel mediation. *Psychological Methods, 15*, 209–233. doi:10.1037/a0020141
- Pressley, M., & Ghatala, E. S. (1990). Self-regulated learning: Monitoring learning from text. *Educational Psychologist, 25*, 19–33. doi:10.1207/s15326985ep2501_3
- Rawson, K. A., & Dunlosky, J. (2002). Are performance predictions for text based on ease of processing? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*, 69–80. doi:10.1037/0278-7393.28.1.69
- Rhodes, M. G., & Castel, A. D. (2008). Memory predictions are influenced by perceptual information: Evidence for metacognitive illusions. *Journal of Experimental Psychology: General, 137*, 615–625. doi:10.1037/a0013684
- Rhodes, M. G., & Castel, A. D. (2009). Metacognitive illusions for auditory information: Effects on monitoring and control. *Psychonomic Bulletin & Review, 16*, 550–554. doi:10.3758/PBR.16.3.550
- Serra, M. J., & Dunlosky, J. (2010). Metacomprehension judgments reflect the belief that diagrams improve learning from text. *Memory, 18*, 698–711. doi:10.1080/09658211.2010.506441
- Serra, M. J., & Metcalfe, J. (2009). Effective implementation of metacognition. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition and education* (pp. 278–298). New York, NY: Routledge.
- Spellman, B. A., & Bjork, R. A. (1992). When predictions create reality: Judgments of learning may alter what they are intended to assess. *Psychological Science, 3*, 315–316. doi:10.1111/j.1467-9280.1992.tb00680.x
- Sungkhassetee, V. W., Friedman, M. C., & Castel, A. D. (2011). Memory and metamemory for inverted words: Illusions of competency and desirable difficulties. *Psychonomic Bulletin & Review, 18*, 973–978. doi:10.3758/s13423-011-0114-9
- Tauber, S. K., & Rhodes, M. G. (2012). Multiple bases for young and older adults' judgments of learning in multitrial learning. *Psychology and Aging, 27*, 474–483. doi:10.1037/a0025246
- Thiede, K. W., Anderson, M. C. M., & Theriault, D. (2003). Accuracy of metacognitive monitoring affects learning of texts. *Journal of Educational Psychology, 95*, 66–73. doi:10.1037/0022-0663.95.1.66
- Thiede, K. T., & Dunlosky, J. (1999). Toward a general model of self-regulated study: An analysis of selection of items for study and self-paced study time. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*, 1024–1037. doi:10.1037/0278-7393.25.4.1024
- 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*, 1264–1269. doi:10.1037/a0023719
- Undorf, M., & Erdfelder, E. (2013). Separation of encoding fluency and item difficulty effects on judgments of learning. *The Quarterly Journal of Experimental Psychology, 66*, 2060–2072. doi:10.1080/17470218.2013.777751
- Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. In D. J. Hacker, J. Dunlosky, & A. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 277–304). Hillsdale, NJ: Erlbaum.
- Yue, C. L., Castel, A. D., & Bjork, R. A. (2013). When disfluency is—and is not—a desirable difficulty: The influence of typeface clarity on metacognitive judgments and memory. *Memory & Cognition, 41*, 229–241. doi:10.3758/s13421-012-0255-8

(Appendix follows)

Appendix

Examples of Perceptual Feature Manipulation During Learning in Experiments 1–3

Experiment 1A & 1B: Font Bolding

Unbolded

Bolded

• superego: The Freudian structure of personality that serves as the harsh internal judge of our behavior; what we often call conscience.

• **superego**: The Freudian structure of personality that serves as the harsh internal judge of our behavior; what we often call conscience.

Experiment 2: Font Size

Small

Large

• superego: The Freudian structure of personality that serves as the harsh internal judge of our behavior; what we often call conscience.

• **superego**: The Freudian structure of personality that serves as the harsh internal judge of our behavior; what we often call conscience.

Experiment 3: Font Bordering

Unbordered

Bordered

• superego: The Freudian structure of personality that serves as the harsh internal judge of our behavior; what we often call conscience.

• **superego**: The Freudian structure of personality that serves as the harsh internal judge of our behavior; what we often call conscience.

Received February 3, 2013
 Revision received May 16, 2014
 Accepted May 21, 2014 ■