When disfluency is—and is not—a desirable difficulty: The influence of typeface clarity on metacognitive judgments and memory

Carole L. Yue · Alan D. Castel · Robert A. Bjork

Abstract There are many instances in which perceptual disfluency leads to improved memory performance, a phenomenon often referred to as the *perceptual-interference effect* (e.g., Diemand-Yauman, Oppenheimer, & Vaughn (Cognition 118:111–115, 2010); Nairne (Journal of Experimental Psychology: Learning, Memory, and Cognition 14:248–255, 1988)). In some situations, however, perceptual disfluency does not affect memory (Rhodes & Castel (Journal of Experimental Psychology: General 137:615–625, 2008)), or even impairs memory (Glass, (Psychology and Aging 22:233–238, 2007)). Because of the uncertain effects of perceptual disfluency, it is important to establish when disfluency is a "desirable difficulty" (Bjork, 1994) and when it is not, and the degree to which people's judgments of learning (JOLs) reflect the consequences of processing disfluent information. In five experiments, our participants saw multiple lists of blurred and clear words and gave JOLs after each word. The JOLs were consistently higher for the perceptually fluent items in within-subjects designs, which accurately predicted the pattern of recall performance when the presentation time was short (Exps. 1a and 2a). When the final test was recognition or when the presentation time was long, however, we found no difference in recall for clear and blurred words, although JOLs continued to be higher for clear words (Exps. 2b and 3). When fluency was manipulated between subjects, neither JOLs nor recall varied between formats (Exp. 1b). This study suggests a boundary condition for the desirable difficulty of perceptual disfluency and indicates that a visual distortion, such as blurring a word, may not always induce the deeper processing necessary to create a perceptual-interference effect.

C. L. Yue (✉) · A. D. Castel (✉) · R. A. Bjork
Department of Psychology, University of California,
Los Angeles, CA, USA
e-mail: caroleleigh@ucla.edu
e-mail: castel@ucla.edu

The sense of fluency, or the subjective ease with which a person processes information, impacts a variety of judgments about that information. For example, items that are perceived as more perceptually fluent (e.g., larger in font size or greater in visual clarity) are more likely to be judged as typical members of a category (Oppenheimer & Frank, 2008) or as having been previously studied (Whittlesea, Jacoby, & Girard, 1990). The effect of perceptual fluency extends to memory predictions: More easily processed words are usually predicted to be more recallable or recognizable in the future (Begg, Duft, Lalonde, Melnick, & Sanvito, 1989; Hirshman & Mulligan, 1991; Nairne, 1988; Rhodes & Castel, 2008).

While people consistently judge less fluent items as being more difficult to remember or recognize on a future test, actual recall is often surprisingly unaffected by, or is even improved by, perceptual disfluency (Diemand-Yauman, Oppenheimer, & Vaughan, 2010; Hirshman & Mulligan, 1991; Rhodes & Castel, 2008; Slamecka & Graf, 1978). Hirshman and Mulligan (1991) and Mulligan (1996) examined the effects of perceptual interference on memory using pattern masks to occlude words almost immediately after their initial presentation. These researchers found that words in the perceptual-interference condition were recognized and recalled better than words that were not visually obscured. Such improved performance on explicit memory tests as a result of interfering with perceptual processing has been termed the *perceptual-interference effect* (Mulligan, 1996).

The discrepancy between memory predictions, known as *judgments of learning* (JOLs), and performance on measures of learning is especially relevant for students and educators, because faulty beliefs can lead to improper study strategies and inefficient learning. The mnemonic effects of disfluency are counterintuitive to many people in learning situations;

generally, a more fluent experience during study implies greater comprehension and leads to greater confidence in one's memory for the material (Benjamin, Bjork, & Schwartz, 1998; Maki, Foley, Kajer, Thompson, & Willert, 1990; Rawson & Dunlosky, 2002). Diemand-Yauman et al. (2010), however, demonstrated that text in a disfluent typeface (e.g., Monotype Corsiva) was remembered better than text in a clear typeface (e.g., Arial). Similarly, Sungkhasettee, Friedman, and Castel (2011) found that people recalled inverted words better than upright words, even though their JOLs did not reflect awareness of the benefit of disfluency.

Evidence that disfluency can be undesirable as well as desirable

In Diemand-Yauman et al. (2010) and Sungkhasettee et al. (2011), it appears that disfluency acted as a "desirable difficulty" (Bjork, 1994)—that is, a learning condition that makes encoding more difficult, but also engages processes that support learning and improve long-term retention. One explanation for why disfluency can be a desirable difficulty is that interfering with the perceptual processing of an item leads to additional, higher-level processing, which strengthens the associations among visual, semantic, and acoustic information in the perceptual system. The strengthening of these associations leads to better memory for the items in the interference condition than in the intact condition (Hirshman & Mulligan, 1991; Mulligan, 1996). Similarly, Alter, Oppenheimer, Epley, and Eyre (2007) speculated that a low degree of fluency during a learning event may act as a cue to engage in deeper, more elaborative processing. Thus, perceptual disfluency, such as a difficult-to-read typeface or the need to mentally invert a word in order to read it, induces the type of effortful processing that improves memory for those "difficult" items.

Results such as those in Diemand-Yauman et al. (2010) and Sungkhasettee et al. (2011) might tempt educators to apply similar manipulations in their classrooms in order to aid their students' learning—for example, creating lecture slides in an unusual font or even leaving the projector slightly out of focus. Perceptual disfluency does not, however, consistently lead to improved memory. Rhodes and Castel (2008), for example, found that while words presented in larger font were given higher JOLs than words presented in smaller font, recall was unaffected. In addition, Lindenberger, Scherer, and Baltes (2001) created perceptual disfluency by giving middle-aged adults sensory filters to simulate impaired visual and auditory acuity, but found that such disfluency neither enhanced nor impaired later recall of text passages or paired associates. Some studies even indicate negative performance effects of perceptual disfluency: Stimulus degradation causes an increase in reaction times for lexical decision tasks (Plourde & Besner, 1997; Yap & Balota, 2007), and performance on memory tasks that demand a

high amount of cognitive processing show a direct correlation with visual acuity: As sensory function decreases, so does performance on cognitively demanding working and long-term memory tasks (Glass, 2007).

One possible reason for such mixed effects of perceptual disfluency may be that some manipulations affect only lower-level visual processes, leaving higher-level cognition unaffected or impaired. Yap and Balota (2007), for example, hypothesized that the visual quality of a stimulus (i.e., how degraded or clear it is) affects an early stage of encoding and causes slower reaction times for lexical decision tasks (see also Sternberg, 1969). In combination with Glass's (2007) findings, this theory predicts that some types of perceptual disfluency may impair, rather than improve, cognitive processes, particularly if the disfluency involves distorting the stimuli. Similarly, Hirshman, Trembath, and Mulligan (1994) suggested that a strong focus on the visual aspects of stimuli could prevent other types of processing from occurring, and thus impair recall.

Such explanations are in line with classic models of memory that have suggested that encoding processes occur in a limited-capacity channel (Atkinson & Shiffrin, 1968; Baddeley, 1981). Anything that adds to the effort required at encoding but that is irrelevant to the learning experience is extraneous load and detracts from performance on memory tasks (Sweller, van Merrienboer, & Paas, 1998). Following this logic, if the task of visually interpreting and encoding distorted stimuli exceeds working memory limits, then we may see lower recall for those items; if, however, the task is within working memory limits but does not induce extra processing of the disfluent information, we may expect to see similar performance levels between fluent and disfluent conditions, as in the findings obtained by Rhodes and Castel (2008) and Lindenberger et al. (2001).

Goals of the present research

The experiments reported in the present article were designed to further explore the effects of perceptual fluency on metacognitive predictions of future recall performance and on actual recall performance. Words were visually distorted and presented in a blurred font to create disfluency. While many different measures of disfluency (e.g., irregular or distorted font, pattern masking, or text printed by a printer low on toner) have been used in previous research (Diemand-Yauman et al., 2010; Guenther, 2012; Mulligan, 1996; Oppenheimer & Frank, 2008), we chose to use the blurring manipulation because it was conceptually similar to manipulations used in a variety of previous studies (e.g., Glass, 2007; Guenther, 2012; Lindenberger et al., 2001; Oppenheimer & Frank, 2008). In addition, out-of-focus text is not an improbable occurrence in a classroom—although instructors may not purposefully leave their projectors out of focus, or students

may not intentionally sit farther than their eyesight will allow them to comfortably read the blackboard, incidental occurrences of these events are realistic instances in which learners may face visual distortion of the to-be-learned material.

On the basis of prior research (e.g., Hirshman & Mulligan, 1991; Rhodes & Castel, 2008), we assumed that the perceptual disfluency would lead distorted words to be given lower JOLs than would corresponding words presented in a clear typeface. The predictions for memory performance were not so clear; given the inconclusive pattern of findings on the mnemonic effects of perceptual fluency, three outcomes were possible: (1) If perceptual disfluency in the form of blurring is a desirable difficulty, we would expect better recall for blurred words as compared to clear words; (2) if this form of perceptual disfluency is more akin to the impaired visual acuity of Lindenberger et al. (2001), we would expect no difference in recall for blurred and clear words; (3) if remembering blurred words requires too high of a cognitive demand (cf. Glass, 2007), we would expect worse recall for blurred than for clear words.

Experiment 1a

All participants in Experiment 1a saw words presented in either a blurred or a clear font. Immediately following the presentation of each word, the participants verbally gave JOLs on a scale of 1–100. We used a multilist method so that any experience-based changes in JOLs, recall, or resolution would be evident (cf. Castel, 2008; deWinstanley & Bjork, 2004). If participants did perceive any mnemonic effects of disfluency, it is likely that their JOLs on subsequent lists would change accordingly and that resolution would improve across lists. Since blurred words should be perceived as less fluent than clear words, we anticipated that JOLs would be lower for the blurred words than for the clear words. The pattern of recall performance would help determine how visual distortion affects memory, providing further insight as to the extent and/or limitations of the perceptual-disfluency effect.

Method

Participants A group of 25 undergraduates enrolled in introductory-level psychology courses at the University of California, Los Angeles, participated for course credit. Each participant was tested individually.

Materials A selection of 110 words were taken from the English Lexicon Project database and normed for frequency and length (Balota et al., 2007). The words had an average log HAL frequency of 9.6 (i.e., relatively common) and an average length of 5.29 letters. The words were randomly divided into four lists of 26 words each, with the six remaining words being used as examples at the beginning of the experiment. Each list contained equal numbers of clear and blurred words randomly distributed throughout the list. The first and last words of each list were eliminated from the analysis to account for primacy and recency effects, leaving 24 key words, 12 blurred and 12 clear, in each list. The lists were counterbalanced so that each list appeared equally often in the first, second, third, and fourth positions. Blurred words were distorted using a computer program to disperse the pixels in each letter by 10%, which pilot data had indicated was enough to noticeably blur the words, but not enough to impede people's ability to read them. See the Appendix for an example of the stimuli.

Procedure The clarity of the words was manipulated within subjects. Participants saw four lists of 26 words, and each word was presented for 0.5 s in black, size 44 font on a white background. Equal numbers of words were presented in regular (i.e., clear) font and in blurred font. The clarity of the words was counterbalanced so that all words were presented equally often across participants in clear or in blurred font. After each word was presented, the participants were given 2 s to give a JOL by rating, on a scale of 1 (*not at all confident*) to 100 (*completely confident*), how confident they were that they would recall that item on a later test. Participants were instructed to say "0" if they had not been able to read the word. Prior to the first list, participants viewed six example words (three blurred and three clear) and were asked to practice making confidence judgments. Any remaining questions were answered before the participants began with the first list.

Immediately after each list, the participants engaged in a 10-s distractor task that involved counting backward by multiples of three. The starting number differed on each list. After the distractor task, participants were asked to recall out loud as many words as they could remember from the previous list. The participants did not receive feedback. This process was repeated three times, for a total of four lists.

Results

The alpha level was set to .05 for all inferential statistics, and all effect sizes are reported in terms of η_p^2 for ANOVAs or of Cohen's *d* for *t* tests. Across all 25 participants, in 28 cases a participant responded "0" in the judgment phase, to indicate that he or she had not seen the previous word. Twenty-two of those words were blurred and six were clear. These words were removed from all analyses, but we report the unconditionalized data as well. No single participant said "0" to more than five out of the 104 words that he or she saw, and no single word received a JOL of "0" more than twice.

Judgments of learning JOLs were analyzed in a 2 (format: blurred, clear) × 4 (list: 1, 2, 3, 4) ANOVA. As can be seen in

the top panel of Fig. 1, participants gave significantly higher JOLs to clear words ($M = 58.76$, $SE = 3.79$) than to blurred words ($M = 48.34$, $SE = 3.71$), $F(1, 24) = 18.84$, $\eta_p^2 = .44$. There was also a significant effect of list, $F(3, 72) = 21.74$, $\eta_p^2 = .48$, indicating that JOLs decreased across lists, regardless of format. We also found a borderline interaction between list and format, $F(3, 72) = 2.77$, $p = .05$, $\eta_p^2 = .10$. Paired-samples t tests indicated that within each list, clear words were given significantly higher JOLs than were blurred words, with an average d of 0.5. Within each format, participants gave the highest JOLs in List 1 (blurred, $M = 59.94$, $SE = 5.05$; clear, $M = 74.22$, $SE = 4.62$), but the JOLs decreased in subsequent lists, leveling off in Lists 3 and 4.

Recall The recall data were analyzed in a 2 (format) × 4 (list) ANOVA and are reported as percentages. As is shown in the bottom panel of Fig. 1, there was no interaction or effect of list, but there was a marginal effect of format, $F(1, 24) = 3.16$, $p = .09$, $\eta_p^2 = .12$. More clear words ($M = 29.05$, $SE = 2.46$) were recalled than blurred words ($M = 25.27$, $SE = 2.25$).

Resolution We used the Goodman–Kruskal gamma correlation as a nonparametric measure of the association between format, JOLs, and recall (Nelson, 1984). This analysis was conducted to examine participants' metacognitive accuracy—that is, whether they were actually more likely to recall a particular word if they gave it a higher JOL than if they gave it a lower JOL. For both blurred and clear words, resolution was significantly different from zero, $G = .27$,

$SE = .05$, $t(24) = 5.70$, and $G = .27$, $SE = .06$, $t(24) = 4.30$, respectively. These data indicate that participants were generally more likely to remember words to which they had given higher JOLs. There was no main effect of format or of list on resolution, but there was an interaction between list and format, $F(3, 45) = 3.20$, $\eta_p^2 = .176$: Resolution increased for blurred words and decreased for clear words across lists.

Unconditionalized data The same analyses with all words included revealed an identical pattern: Clear words received significantly higher JOLs than did blurred words, $F(1, 24) = 20.26$, $\eta_p^2 = .46$, and JOLs decreased across lists regardless of format, $F(3, 72) = 20.77$, $\eta_p^2 = .46$. We also saw a marginal interaction with a small effect size between list and format, $F(3, 72) = 2.31$, $p = .08$, $\eta_p^2 = .09$. There was no interaction or effect of list on recall, but there was a marginal effect of format on recall, $F(1, 24) = 3.76$, $p = .06$, $\eta_p^2 = .14$. The gamma correlations were significantly different from zero, $G = .30$, $SE = .05$, $t(24) = 6.28$, and $G = .29$, $SE = .04$, $t(24) = 4.72$, for blurred and clear words, respectively. There was no main effect of format or of list on resolution, but the interaction between those two variables was marginal, $F(3, 48) = 2.28$, $p = .09$, $\eta_p^2 = .13$. The unconditionalized data showed the same pattern as in the main analysis, of the resolution for blurred words increasing across lists as the resolution for clear words decreased across lists. The similarities between the conditionalized and unconditionalized data indicated that item effects did not sway our conditionalized results.

Discussion

In Experiment 1a, we examined JOLs and recall for blurred and clear words. As expected, we found that JOLs were higher for clear than for blurred words. Interestingly, resolution improved for blurred words and worsened for clear words across lists, indicating that participants were able to adjust their JOLs appropriately for the disfluent words. Participants may have been aware that they were not remembering the blurred words very often, and adjusted their JOLs on subsequent lists accordingly; why resolution did not also improve for clear words, however, is uncertain.

We did not find support for a desirable-difficulty explanation, but we also did not clearly support one of the other two theories. Since clear words were recalled only marginally more than blurred words, it was unclear whether this level of visual distortion was only a basic manipulation of visual acuity, as in Lindenberger et al. (2001), or whether deciphering the blurred words required a high enough cognitive demand to impair recall, as in Glass (2007). Another possibility is that the within-subjects design masked any unique benefit for disfluent items. For example, Alter et al. (2007) showed that merely having a title in a disfluent font
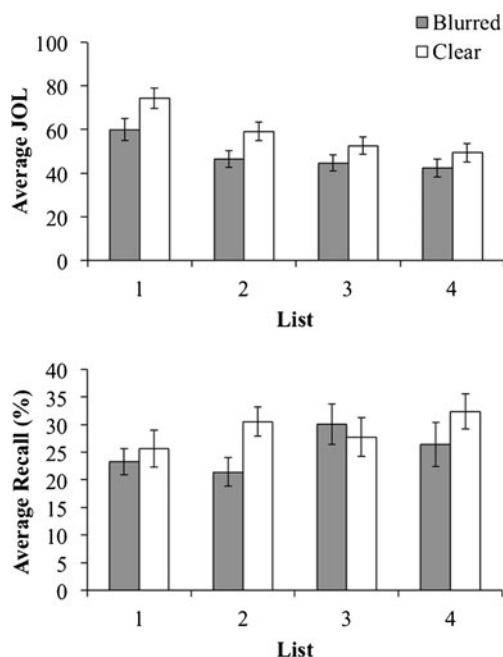

Fig. 1 Average JOLs and percentages of recall for blurred and clear words as a function of list in Experiment 1a. In all figures, error bars represent standard errors of the means

caused participants to process the content below it more deeply. If seeing even one blurred word induced participants to deeply process all of the words on that list, we would not see a perceptual-interference effect within subjects. Experiment 1b addressed this issue, in that the disfluency manipulation was presented to only half of the participants.

Experiment 1b

In Experiment 1a, we found that when participants saw both fluent and disfluent items, they gave higher JOLs to the fluent items. Instead of a perceptual-interference effect on memory, we found that clear words were recalled slightly better than blurred words. To test whether the appearance of disfluency induced deeper processing that extended beyond the disfluent items to other items in the list, for Experiment 1b we used a between-subjects design.

Method

Participants A group of 26 undergraduates enrolled in introductory-level psychology courses at the University of California, Los Angeles, participated for course credit. Each participant was tested individually.

Materials and procedure The materials were identical to those used in Experiment 1a, with the exception that two lists were randomly selected from the previous four. The procedure was also identical, again with the exception that participants saw and were tested on only two lists.

Results

The alpha level was set to .05 for all inferential statistics, and all effect sizes are reported in terms of η_p^2 for ANOVAs or of Cohen's *d* for *t* tests. Across all 26 participants, there were only 14 instances of a "0" response in the judgment phase, indicating that the participant had not seen the previous word. Eight of those instances were blurred, and six were clear. These words were removed from all analyses, but we report unconditionalized data as well. No single participant responded "0" to more than four out of the 52 words that he or she saw, and no single word received a JOL of "0" more than twice.

Judgments of learning The results were analyzed in a 2 (format) × 2 (list) mixed-subjects ANOVA. The top panel of Fig. 2 illustrates that there was no effect of format on JOLs, $F(1, 24) < 1$, but JOLs did drop significantly from the first list to the second list, $F(1, 24) = 8.49$, $\eta_p^2 = .26$. We found no interaction between list and format, $F(1, 24) < 1$.
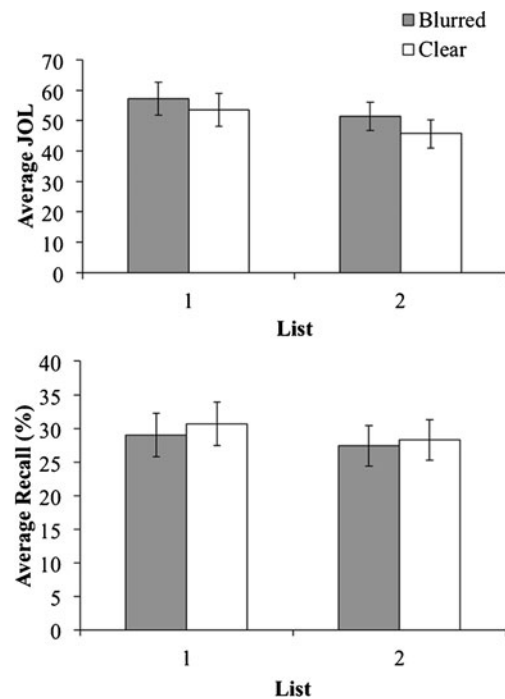

Fig. 2 Average JOLs and percentages of recall for blurred and clear words as a function of list in Experiment 1b

Recall The recall data were analyzed in a 2 (format) × 2 (list) mixed-subjects ANOVA. As can be seen in the bottom panel of Fig. 2, format had no effect on recall, $F(1, 24) < 1$. There was also no difference in recall between the lists, $F(1, 24) = 1.04$, $p = .32$, and no interaction, $F(1, 24) < 1$.

Resolution For both blurred and clear words, resolution was significantly different from zero, $G = .40$, $SE = .07$, $t(12) = 5.66$, and $G = .30$, $SE = .09$, $t(12) = 3.16$, respectively, again indicating that participants' JOLs were generally sensitive to recall. A 2 × 2 ANOVA revealed no significant differences in resolution between formats, $F(1, 24) < 1$, or lists, $F(1, 24) < 1$.

Unconditionalized data The same analyses on the unconditionalized data revealed similar results: JOLs decreased from List 1 to List 2, $F(1, 24) = 9.57$, $\eta_p^2 = .29$, and neither recall nor resolution varied by list or format.

Discussion

Even using a between-subjects manipulation, we did not see a mnemonic benefit for disfluent items. These results suggest that the lack of a perceptual-interference effect observed in Experiment 1a—with even a trend in the opposite direction—was not the result of participants engaging in generally deeper processing upon presentation of the blurred words (cf. Alter et al., 2007). Since we observed no effect of format on JOLs or on recall using a between-subjects manipulation in the present experiment, it is likely that the comparison of fluent and

disfluent words within a list, as in the prior experiment (Exp. 1a), does indeed affect JOLs and, possibly, resolution and recall. In the following experiments, we continued to examine possible reasons for the impaired recall evident in Experiment 1a.

Given that visual distortion did not appear to be a desirable difficulty in Experiments 1a and 1b, one possible explanation was that blurriness is only a minor manipulation of visual acuity (Lindenberger et al., 2001), in which case we would expect no difference in recall between blurred and clear words, regardless of other manipulations (e.g., study length). It should be noted, however, that there were more instances in which participants did not see a blurred word than did not see a clear word, indicating that participants did experience strong enough visual disfluency to impair perceptual identification of the word. While these instances occurred only a few times, it may be that identifying, processing, and encoding the words required a high enough cognitive demand to result in lower recall for the blurred words (Glass, 2007).

Although participants indicated that they had not seen the word in only very few instances, the presentation time may have been too brief to allow participants sufficient time to process a blurred word before moving on to the next item. In an effort to determine whether the poorer recall for blurred words was influenced more by the short presentation time or the distortion itself, in Experiments 2a and 2b we extended the presentation time of the stimuli to 2 s and conducted both recall and recognition tests. If distorted words continued to require too high a cognitive demand to interpret and encode, we should see the same pattern of results as in Experiment 1a. If the longer presentation allowed participants sufficient time to accurately recognize and process each word, we might see no difference in recall between formats—in fact, it is possible that participants would have the time that they needed to engage in additional processing for the disfluent words, leading to a benefit for blurred words (cf. Alter et al., 2007).

Experiment 2a

Experiment 1a had suggested that visually distorting words is not a desirable difficulty for learning, and in fact could harm recall. It is possible, however, that having more time to process the words could reduce the working memory resources expended to decipher the word, allowing time for equal, or even relatively deeper, processing of blurred words. In that case, we might see a perceptual-disfluency effect in which blurred words were recalled better than clear words. It is also possible that the longer presentation time could make the stimuli seem more like text that one might see if one had low visual acuity (e.g., without prescription glasses). In that case, we would expect no difference in recall, since sensory

impairment by itself does not lead to reduced memory (Lindenberger et al., 2001). Replicating the results from Experiment 1a, however, would indicate that even with extra time, processing visually distorted words within a series of blurred and clear words creates such a high demand on encoding processes that long-term memory is impaired.

Method

Participants A group of 25 undergraduates enrolled in introductory-level psychology courses at the University of California, Los Angeles, participated for course credit. Each participant was tested individually.

Materials and procedure The materials were identical to those used in Experiment 1a. The procedure was also identical, with the exception that all stimuli were shown for 2 s instead of 0.5 s.

Results

The alpha level was set to .05 for all inferential statistics, and all effect sizes are reported in terms of η_p^2 for ANOVAs or of Cohen's d for t tests. Across all 25 participants, there were only 14 instances of a "0" response in the judgment phase, indicating that the participant had not seen the previous word. Eight of those words were blurred and six were clear. These words were removed from all analyses, but we report unconditionalized data as well. No single participant said "0" to more than five out of the 104 words that he or she saw, and no single word received a JOL of "0" more than twice.

Judgments of learning Results were analyzed in a 2 (format) × 4 (list) ANOVA. As can be seen in the top panel of Fig. 3, there was a significant effect of format, $F(1, 24) = 8.56$, $\eta_p^2 = .26$, such that clear words were given significantly higher JOLs ($M = 50.84$, $SE = 4.21$) than blurred words ($M = 44.65$, $SE = 3.67$). JOLs significantly decreased across lists, $F(3, 72) = 23.70$, $\eta_p^2 = .50$, replicating the results from Experiment 1. There was a marginal interaction between list and format, $F(3, 72) = 2.50$, $p = .07$, $\eta_p^2 = .09$: The difference between JOLs for blurred and clear words was significant in List 1, $t(24) = 3.67$, $d = 0.5$, but diminished across lists.

Recall The recall data were analyzed in a 2 (format) × 4 (list) ANOVA. As can be seen in the bottom panel of Fig. 3, there was no effect of list on recall, $F(3, 72) < 1$. There was, however, a significant effect of format on recall, $F(1, 24) = 5.05$, $\eta_p^2 = .17$: Clear words were recalled significantly more often ($M = 26.45$, $SE = 1.89$) than blurred words ($M = 22.5$, $SE = 1.59$). There was no interaction between list and format, $F(3, 72) < 1$.
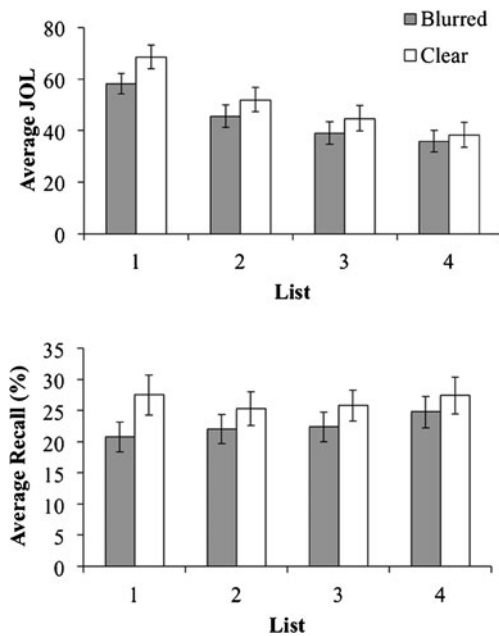
Fig. 3 Average JOLs and percentages of recall for blurred and clear words as a function of list in Experiment 2a

Resolution For one participant, we were unable to calculate a gamma correlation for clear words because he or she gave the exact same JOL to every clear word, but the elimination of those data should not affect the overall analysis. For blurred words, resolution was significantly different from zero, $G = .19$, $SE = .06$, $t(24) = 3.42$. Resolution was also significantly different from zero for clear words, $G = .24$, $SE = .05$, $t(23) = 5.02$. These gamma correlations were not significantly different between word formats, $t(23) < 1$, indicating that although participants gave numerically more accurate JOLs for clear words, there was no significant difference in resolution.

We found a marginal effect of list on resolution, $p = .07$. Participants had the best resolution on List 4 ($G = .43$, $SE = .08$), which differed significantly from the resolutions on List 3 ($G = .07$, $SE = .1$), $d = 0.7$, and List 1 ($G = .26$, $SE = .09$), $d = 0.6$, but not from that on List 2 ($G = .29$, $SE = .12$). There were no other significant differences between lists.

Unconditionalized data When all words were included in the analyses, the data demonstrated the same pattern: JOLs were significantly higher for clear than for blurred words, $F(1, 24) = 9.68$, $\eta_p^2 = .29$, and they decreased across lists, $F(3, 72) = 21.93$, $\eta_p^2 = .48$. An interaction between list and format was present, similar to the one in the conditionalized data, $F(3, 72) = 3.30$, $\eta_p^2 = .12$. Recall was significantly higher for clear than for blurred words, $F(1, 24) = 5.15$, $\eta_p^2 = .18$, but there was no effect of list and no interaction. The resolution also indicated that, in general, participants were more likely to recall words to which they had given higher JOLs (i.e., resolution was significantly different from zero), $G = .20$, $SE = .05$, $t(23) = 3.67$, for blurred

words, and $G = .25$, $SE = .05$, $t(23) = 5.26$, for clear words. Neither format nor list had an effect on resolution in the unconditionalized data. Just as in Experiments 1a and 1b, the similarity between the conditionalized and unconditionalized data indicated that item selection effects had not contributed to our results in the conditionalized data.

Discussion

Again, these results did not replicate previous findings that free recall was superior for words in a perceptual-interference condition (Hirshman & Mulligan, 1991; Hirshman et al., 1994); in fact, recall was worse for blurred than for clear words. One possible reason for our results is that blurring may only affect lower-level visual processes—if participants are forced to attend to the low-level visual information in the distorted stimuli, they may not be able to perform the type of processing necessary to effectively encode those words. This explanation is in line with Glass's (2007) theory of sensory and cognitive effort and with classic models of memory (Atkinson & Shiffrin, 1968; Baddeley, 1981).

Looking at the present data in light of these theories, attempting to encode blurred words could activate visual processes that take up a significant portion of working memory. While the effort of this act is sufficient to induce a feeling of disfluency, the learner does not have either the time or the capacity to engage in deeper processing. This interpretation may explain why both JOLs and average recall were lower for blurred words. Alternatively, recall may not be as sensitive to the perceptual-interference effect as are other explicit memory evaluations; similar studies have found stronger perceptual-interference effects on recognition tests than on free recall tests (Mulligan, 1996; Nairne, 1988). For this reason, Experiment 2b involved a recognition test instead of a recall test.

Experiment 2b

In previous studies, perceptual interference has been shown to enhance performance on yes/no recognition tests, even when recall does not show the same benefits (Hirshman & Mulligan, 1991; Nairne, 1988). The proposed reason for this discrepancy is that during the initial perceptual identification process, the learner is focusing on surface-level aspects of the word in order to identify it. Doing so would aid later recognition, but not recall, for fluent items, given that recall relies more on item elaboration than on perceptually distinctive features (Nairne, 1988). In Experiment 2b, we tested whether blurring words would lead to similar benefits in recognition memory. An aural recognition test was administered in order to reduce any misleading effects of transfer-appropriate processing (i.e., words that were clear at

encoding would be more likely to be correctly recognized if they were also clear at test).

Method

Participants A group of 26 undergraduates enrolled in introductory-level psychology courses at the University of California, Los Angeles, participated for course credit. Each participant was tested individually.

Materials Half of the words used in Experiments 1a and 2a were randomly selected for presentation in Experiment 2b. The remaining 52 words were used as foils in the final recognition test. The final test consisted of a list of 104 words—the 52 that participants had seen previously and the 52 foils—read out loud in a random order to the participants. The same six words used in the previous experiments were used as an example of the task at the beginning of this experiment.

Procedure The procedure for Experiment 2b was very similar to the that for Experiment 2a, except that participants were told that they would be taking a recognition test instead of a recall test. The participants were informed that after seeing a series of words, they would hear words out loud and be asked to say "Yes" if they had seen those words in the presentation, and "No" if they had not. Instead of separating the words into lists, we presented all 52 words in one list, with one aural recognition test at the end. Similar to the previous experiments, participants were asked to rate their confidence, on a scale of 1–100, that they would be able to remember each word for a later recognition test. They were told to say "0" if they had not seen the word. Just as in Experiment 2a, participants were given 2 s to view the word and 2 s to make a JOL.

In the test phase, participants were read a word out loud and asked to say "Yes" if they had seen the word on the list or "No" if they had not. The second phase of the test gauged incidental learning of the perceptual aspects of the stimuli; if participants acknowledged seeing a word on the list, they were asked to state whether the word had been clear or blurred. Even if they were not sure, participants were encouraged to take their best guess.

Results

Again the alpha level was set to .05 for all inferential statistics, and all effect sizes are reported in terms of Cohen's *d*. Across all 26 participants, there were only 17 instances of "0" responses in the judgment phase, indicating that the participant had not seen the previous word. Sixteen of those words were blurred and one was clear, and those

words were removed from all analyses. No single participant said "0" to more than four out of the 52 words that he or she saw.

Judgments of learning Since all words were presented in a single list, we analyzed the results with a paired-samples *t* test. As can be seen in the left two bars in Fig. 4, JOLs were higher for clear ($M = 63.56$, $SE = 3.70$) than for blurred ($M = 54.65$, $SE = 3.27$) words, $t(25) = 3.34$, $d = 0.5$.

Recognition The false alarm rate was low ($M = .10$, $SE = .02$) and $d' = 2.32$, indicating high discriminability between targets and lures. A paired-samples *t*test showed that, although the hit rate for clear words was numerically higher than the hit rate for blurred words, there was no significant difference in hit rates between the two formats (blurred, $M = .77$, $SE = .03$; clear, $M = 80$, $SE = .03$), $t(25) = 1.2$, $p = .23$. See the middle two bars in Fig. 4.

Format identification accuracy Accuracy in the incidental learning task was conditionalized on hits; a given participant's accuracy was the proportion of the number of words accurately identified as clear or blurred out of the number of overall hits in that category. Although participants were more likely to say that words had been presented in a clear format ($M = .59$, $SE = .02$) than in a blurred format ($M = .41$, $SE = .02$), a measure of discriminability ($d' = 1.24$) indicated that participants were not randomly guessing in this portion of the test. As can be seen in the right two bars in Fig. 4, participants were more accurate at recalling that clear words had been clear ($M = .80$, $SE = .03$) than that blurred words had been blurred ($M = .63$, $SE = .04$), $t(25) = 3.56$, $d = 1.0$.

Unconditionalized data When no words were removed from the analyses, the statistical patterns remained the same, demonstrating that item effects were not a concern in the
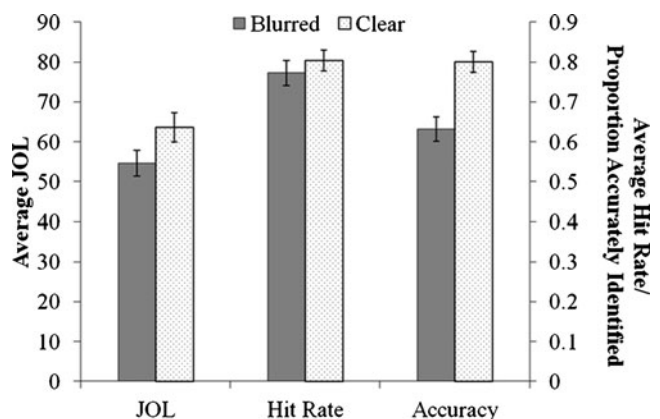

Fig. 4 Average JOLs, hit rates, and format identification accuracy for blurred and clear words in Experiment 2b. JOLs are measured on the left-hand axis, and hit rate and the proportion accurately identified are measured on the right-hand axis

original analyses. Participants gave significantly higher JOLs to clear words than to blurred words, $t(25) = 3.64$, $d = 0.6$. The hit rate was marginally higher for clear words than for blurred words, $t(25) = 1.96$, $p = .06$, $d = 0.4$, and format identification accuracy was significantly higher for clear than for blurred words, $t(25) = 3.57$, $d = 1.0$.

Discussion

Again, consistent with prior research, JOLs were higher for fluent words than for disfluent words (e.g., Rhodes & Castel, 2008). Consistent with Experiments 1a, 1b, and 2a, but inconsistent with prior research (e.g., Mulligan, 1996), recognition performance did not show a benefit for perceptual disfluency. In fact, blurred and clear words were equally likely to be recognized on the final test. Since recognition tests tend to be even more sensitive to the effects of disfluency (e.g., Hirshman & Mulligan, 1991), this result suggests that our inability to find a perceptual-interference effect in Experiments 1a, 1b, and 2a was likely not the result of an insensitive test.

The format identification test demonstrated that participants had better source memory for clear words than for blurred words, which could be explained in one of two ways: First, the cognitive effort required to identify and encode the word might have been high enough to prevent participants from also encoding its format (Glass, 2007), and second, the blurring might have been too minimal to encode at a conscious level, leading participants to guess more often that a word had been clear than that it had been blurred (Lindenberger et al., 2001). Participants' tendency to identify a word as clear rather than blurred could indicate that when participants were confident that they had an episodic memory of seeing the word, they remembered it as clear. This interpretation may be consistent with other biases that occur when people make judgments at retrieval that are based on prior processing fluency (e.g., Benjamin et al., 1998; Castel, Rhodes, McCabe, Soderstrom, & Loaiza, 2012).

Experiment 2a showed a detrimental effect for perceptual disfluency, while Experiment 2b showed very little effect of format on memory. It may be that disrupting the visual appearance of an item within a list does interfere with conceptual processing and item elaboration, effects that tend to be observed in a free recall test, but not a recognition test (Nairne, 1988). On the basis of this interpretation, in Experiment 3 we employed a free recall test in which participants were allowed even longer to process the stimuli.

Experiment 3

In Experiment 3, we extended the presentation time to 5 s for two reasons: (1) to more closely match the presentation times in related research (e.g., Rhodes & Castel, 2008; Sungkhasettee et al., 2011) and (2) to allow sufficient time for deeper processing. This extended time might equalize processing for the blurred and clear words, effectively eliminating any benefit for clear words. If that pattern were to occur, it would suggest that visual distortions do impair encoding processes with short presentation times, but that difficulty can be overcome given more processing time. In addition, we would be able to more closely compare our results with a broader range of previous research.

Method

Participants A group of 24 undergraduates enrolled in introductory-level psychology courses at the University of California, Los Angeles, participated for course credit. Each participant was tested individually.

Materials and procedure The materials were identical to those used in Experiments 1a and 2a. The procedure was also identical, with the exception that all stimuli were shown for 5 s each.

Results

Again the alpha level was set to .05 for all inferential statistics, and recall data are here reported as percentages. All effect sizes are reported in terms of η_p^2, for ANOVAs, or of Cohen's d, for t tests. Across all 24 participants, there were only 16 instances of "0" responses in the judgment phase, indicating that the participant had not seen the previous word. Fifteen of those words were blurred and one was clear. These words were removed from all analyses, and we also report the unconditionalized data. No single participant said "0" to more than three out of the 104 words that he or she saw, and no single word was given a "0" judgment more than twice across all participants.

Judgments of learning The results were analyzed in a 2 (format) × 4 (list) ANOVA. We found a main effect of format, $F(1, 23) = 7.70$, $\eta_p^2 = .25$; JOLs were significantly higher for clear ($M = 48.38$, $SE = 3.92$) than for blurred ($M = 43.99$, $SE = 3.73$) words, which can be seen in the top panel of Fig. 5. There was also a main effect of list, $F(3, 69) = 18.91$, $\eta_p^2 = .45$, indicating that JOLs decreased as the lists progressed. There was no interaction between list and format.

Recall The recall data were also analyzed in a 2 (format) × 4 (list) ANOVA. We found no interaction, but there was a marginal effect of list, $F(3, 69) = 2.61$, $p = .06$, with a small effect size, $\eta_p^2 = .10$. Paired-samples t tests indicated that recall on List 2 ($M = 34.85$, $SE = 2.60$) was significantly higher than recall on List 1 ($M = 28.60$, $SE = 2.41$), $t(1140) = 2.19$, $d = 2.5$, and on List 4 ($M = 28.74$, $SE = 2.29$), $t(1145) =$
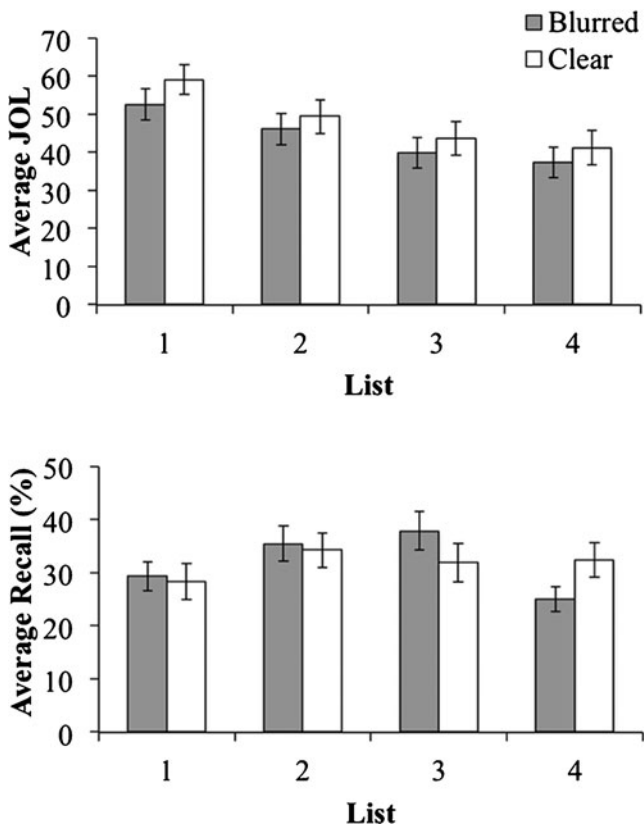
Fig. 5 Average JOLs and percentages of recall for blurred and clear words as a function of list in Experiment 3

2.16, $d = 2.5$. List 2 recall was marginally higher than List 3 recall ($M = 29.88$, $SE = 3.07$), $t(1141) = 1.82$, $p = .07$, $d = 1.7$. See the bottom panel of Fig. 5.

Most importantly, there was no significant effect of format on recall; although recall for clear words ($M = 31.73$, $SE = 2.45$) was numerically higher than recall for blurred words ($M = 29.43$, $SE = 2.08$), the effect was not significant, $F(1, 23) = 1.76$, $p = .20$, $\eta_p^2 = .07$. Altogether, these results suggest that although recall improved for List 2 and then returned to List 1 levels, the format of the word did not influence participants' ability to recall the word.

Resolution Again, for one participant we were unable to calculate a gamma correlation for clear words because he or she gave the exact same JOL to every clear word. For both blurred and clear words, resolution was significantly different from zero, $G = .38$, $SE = .05$, $t(23) = 7.21$, and $G = .31$, $SE = .05$, $t(22) = 6.58$, respectively. Resolution did not differ between formats, $t(22) = 1.26$, $p = .22$, $d = 0.3$. These data show that although participants were generally more likely to remember words to which they had given higher JOLs, resolution did not differ between blurred and clear words. We also found no effect of list on resolution, $F < 1$, indicating that participants' resolutions did not change across lists.

Unconditionalized data To ensure that item effects did not sway our conditionalized results, we conducted the same analyses on the data with all words included. As in previous experiments, the statistical patterns remained the same. We found similar effects of format and list on JOLs: JOLs were higher for clear than for blurred words, $F(1, 24) = 8.14$, $\eta_p^2 = .25$, and $F(3, 72) = 18.56$, $\eta_p^2 = .44$. Even with all words included, there was no effect of format on recall, $F(1, 24) < 1$, but there was a similar effect of list on recall, $F(3, 72) = 2.92$, $\eta_p^2 = .11$. The resolution data also exhibited the same pattern seen in the conditionalized data.

Discussion

When the presentation time was extended to 5 s, participants' JOLs remained sensitive to the blurring manipulation, but their recall was unaffected. It appears that rather than act as a desirable difficulty, visually distorting a word requires a longer presentation time in order for participants to achieve the same level of recall as for clear words. This result suggests that encoding blurred words may require a high amount of effort at an early stage of processing, but longer study time can eliminate the discrepancy in recall between formats without informing metacognitive judgments.

General discussion

In four experiments, we measured the effect of perceptual disfluency, defined as blurring of words, on recall and recognition. We did not show a mnemonic benefit for perceptual disfluency on either form of test; in fact, we found that perceptually fluent and disfluent items were at least equally likely to be remembered (Exps. 1a, 1b, 2b, and 3), and sometimes fluent items were more likely to be remembered (Exp. 2a) than perceptually disfluent items. With the exception of Experiment 1b, JOLs were consistently greater for the clear words than for the blurred words, supporting the notion that JOLs are based on ease of processing and perceptual fluency (Begg et al., 1989). The reason that we did not observe this pattern in Experiment 1b was likely the between-subjects design; when participants were faced with words at the same level of fluency or disfluency, they had no reason to vary their JOLs, and perhaps were not even considering that the words could be presented in a better or a worse format. However, it is interesting to note that, even when there was no difference in memory performance between the two types of presentation in our within-subjects experiments (cf. Rhodes & Castel, 2008), JOLs were always greater for the clear words, suggesting that some additional, and perhaps unappreciated, processing may occur for disfluent information, consistent with other work that has shown the unappreciated benefits of desirable difficulties (e.g., Sungkhasettee et al., 2011).

Overall, the results from the present set of experiments suggest that perceptual disfluency is not universally a desirable difficulty. There could be several possible reasons for the impairment (Exp. 2a) or lack of effect (Exps. 1a, 1b, 2b, and 3) on memory. First, the benefits of perceptual disfluency are thought to arise as a result of deeper processing (Alter et al., 2007). It may be that blurring words, unlike inverting them or presenting them in a different font, engages processes relevant to visual perception but does not induce higher-order semantic processing (Hirshman et al., 1994). Such dynamics would also be in line with Glass's (2007) findings that for highly demanding cognitive tasks, decreased visual acuity can lead to impaired memory. In Experiment 2a, the brief length of the study time, as well as the length of the list, may have created a sufficiently demanding sensory task that it prevented participants from engaging in the deeper processing necessary to create a perceptual-interference effect. In Experiment 3, participants had sufficient time to process both types of words, and the visual degradation no longer impaired their ability to encode the stimuli, providing results more similar to those of Rhodes and Castel (2008) and Lindenberger et al. (2001).

Although previous studies have shown an advantage for disfluency at relatively short presentation intervals (e.g., Mulligan, 1996), and even that a brief presentation can result in better memory than a longer presentation (Nairne, 1988), we did not find similar results in the present context. In the present studies, the shortest presentation was 0.5 s (500 ms), which was significantly longer than the intervals used in the prior work by Nairne (1988) and Mulligan (1996), in which stimuli were presented for 100 ms or less. It could be that extremely brief presentation is necessary to induce the generative processing evident in Nairne (1988), whereas even half a second is long enough for participants to read a word instead of generating it from a perceptual trace, thus eliminating any possible benefit of generation. The lack of difference in word identification across different presentation times supports this explanation; if a 0.5-s interval is long enough for participants to read a word—even a distorted one—then they should have been equally able to identify words at presentation times of 0.5, 2, and 5 s, which was the case. Although this lack of difference may seem to indicate that the blurring manipulation had no effect on fluency, the consistent JOL pattern indicates that blurred words were likely perceived as being less fluent to participants, relative to the clear words.

Another difference between the present experiments and previous work is that the participants in the present study made JOLs for each word. The processing required to make a JOL has been shown to cause participants to process words differently than they otherwise would and to modify memory (Dougherty, Scheck, Nelson, & Narens, 2005; Naveh-Benjamin & Kilb, 2012). Thus, the present procedure differed from prior work (e.g., Mulligan, 1996; Nairne, 1988)

not simply in terms of presentation time, but also in the processing that occurred during study (particularly in the JOL stage); it may be that the act of making the metacognitive judgment affected processing in a way that led to similar or worse memory for blurred as compared to clear words.

Along these lines, the second possible reason for memory impairment for disfluent words is that blurring may have induced participants to engage in reduced processing, and not in generative processing as we expected. Initially, we anticipated that blurring the words would elicit a form of the generation effect. In this phenomenon, people are more likely to remember a word if they see only part of it and must generate the whole word on their own (Jacoby, 1978; Slamecka & Graf, 1978). For example, people tend to remember the word "pumpkin" better if they see "p_mpk_n" and must fill in the blanks than if they simply read the word in its entirety. Given multiple study–test trials, however, participants are able to improve processing of the nongenerated items enough to eliminate the advantage of generation (deWinstanley & Bjork, 2004). The proposed reason for this effect is that participants are able to monitor their learning enough to realize that more effective processing could be used to remember the read-only items.

If awareness and modification of differential processing had occurred in the present studies, however, we would expect to see the same improvement in recall for blurred words after multiple study–test trials. Since that pattern did not emerge, we must seek another explanation. One possibility is that at test, even if participants recalled the word itself, they were unable to correctly recall the format in which the item had initially been presented; if that were the case, they would have no reason to attempt to change their subsequent processing of blurred items. We can see some evidence for this explanation in the incidental-format identification test in Experiment 2b. When participants correctly recognized a word, they were more likely to say that it had been clear than that it had been blurred. Perhaps the mere fact that they remembered a word modified their memory of how it had been presented, demonstrating an interesting bias in memory and explaining why we did not see improvements in recall of blurred words across lists; if participants had realized that they were recalling clear words better, they would likely have changed their JOLs to reflect that realization and/or altered how they were processing the blurred words. As it is, however, JOLs dropped uniformly for both clear and blurred words, indicating that participants did not realize that there was any difference between fluent and disfluent words in their memory. This explanation also applies to our finding that resolution was no different between blurred and clear words; participants were generally unaware of recalling the clear words more often, so they were not able to give more accurate JOLs. It is also possible that participants used only a limited range of the JOL scale, making it more difficult to see any marked improvements in JOLs or resolution.

A third reason that we did not see a benefit for disfluent words could be that the delay between study and test was too short. Several studies on desirable difficulties, such as testing and spacing, have shown benefits at longer testing intervals even when an immediate test did not reveal those benefits (e.g., Glenberg, 1976; Roediger & Karpicke, 2006). In the domain of perceptual disfluency, Diemand-Yauman et al. (2010) demonstrated a mnemonic advantage for perceptually disfluent items at a delay of 15 min or longer, although there was no immediate-test comparison group. While it remains unclear whether differing retention intervals would lead to a different set of findings, future research could incorporate varying delays in order to examine how the effects of disfluency may change over time

Although blurring did not effectively induce the type of processing that would lead to better recall or recognition, it did affect JOLs as expected in four of our five experiments. To explain this result, we look to Koriat's (1997) cue-utilization framework. According to this framework, people use various types of cues to make predictions of future learning, but those cues do not necessarily represent factors that will affect later test performance. The visual degradation of words, therefore, may have a direct effect on JOLs in terms of people's apriori theories of how clarity affects memory, even if memory is unaffected. If participants strongly consider the clarity of a word when they make a JOL, they may activate certain theories that suggest that less perceptually clear items are less likely to be remembered, leading to the pattern evident in the present experiments: higher JOLs for clear than for blurred words, even when there was in fact no difference in memory performance for the clear and blurred words.

It is also possible that the cue of visual distortion has a direct effect on JOLs by affecting the ease of processing of that item. Especially given that we only found differences in JOLs with a within-subjects design, it is likely that participants engaged in comparative processes when studying a mixed-format list. If they had a subjective sense that reading the blurry words was more difficult than reading the clear words, that experience would also have led to the observed JOL pattern.

Conclusions

The present study suggests that is necessary to adopt a more cautious attitude toward disfluency than has been present in recent literature: Not all types of perceptual difficulties are desirable. Disfluency being undesirable in some instances may reflect learners having apriori theories that influence encoding (e.g., that blurred items are not important or do not usually need to be remembered) and/or the inability of certain disfluency manipulations to induce deeper processing (cf. Hirshman et al., 1994; Rhodes & Castel, 2008).

Given that the recall of blurred words equaled that of clear words when participants were given longer to process each word, a likely explanation is that visually distorting words created too high a demand on the cognitive processes necessary to encode words presented rapidly in list form (Glass, 2007). Instead of creating a desirable difficulty, blurring impaired learning when processing time was strictly limited and had no effect when the processing time was long enough. Interestingly, people may engage in additional processing for the blurred words that then brings the recall level to that of the cleared words, suggesting some form of additional, and beneficial, encoding of the blurred words. Finally, the present work provides insight regarding the practical implications for educational settings. While some types of disfluency may be desirable, such as presenting textual information in an unusual or distinctive font (Diemand-Yauman et al., 2010), visual distortions, such as blurred or out-of-focus text, can potentially impair learning. In other words, perceptual disfluency does not always aid recall, and thus it will be necessary to clarify when disfluency is a desirable difficulty and when it impedes learning before recommending classroom implementations.

Appendix

All of the words were originally in Calibri, size 44 black font. Blurred words were created using a computer program that dispersed the pixels by 10 percent

Example of a clear word:

stamp

Example of a blurred word:

References

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.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 2, pp. 89–195). New York, NY: Academic Press.

Baddeley, A. (1981). The concept of working memory: A view of its current state and probable future development. *Cognition, 10,* 17–23. doi:10.1016/0010-0277(81)90020-2

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B.… Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods, 39,* 445–459. doi:10.3758/BF03193014

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.

Benjamin, A. S., Bjork, R. A., & Schwartz, B. L. (1998). The mismeasure of memory: When retrieval fluency is misleading as a metamnemonic index. *Journal of Experimental Psychology. General, 127,* 55–68. doi:10.1037/0096-3445.127.1.55

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.

Castel, A. D. (2008). Metacognition and learning about primacy and recency effects in free recall: The utilization of intrinsic and extrinsic cues when making judgments of learning. *Memory & Cognition, 36,* 429–437. doi:10.3758/MC.36.2.429

Castel, A. D., Rhodes, M. G., McCabe, D. P., Soderstrom, N. C., & Loaiza, V. M. (2012). *The fate of being forgotten: Forgotten information is judged as less important*. Manuscript under revision.

deWinstanley, P. A., & Bjork, E. L. (2004). Processing strategies and the generation effect: Implications for making a better reader. *Memory & Cognition, 32,* 945–955. doi:10.3758/BF03196872

Diemand-Yauman, C., Oppenheimer, D. M., & Vaughan, E. B. (2010). Fortune favors the **bold** (*and the italicized*): Effects of disfluency on educational outcomes. *Cognition, 118,* 111–115.

Dougherty, M. R., Scheck, P., Nelson, T. O., & Narens, L. (2005). Using the past to predict the future. *Memory & Cognition, 33,* 1096–1115.

Glass, J. M. (2007). Visual function and cognitive aging: Differential role of contrast sensitivity in verbal versus spatial tasks. *Psychology and Aging, 22,* 233–238.

Glenberg, A. M. (1976). Monotonic and nonmonotonic lag effects in paired-associate and recognition memory paradigms. *Journal of Verbal Learning and Verbal Behavior, 15,* 1–16.

Guenther, R. K. (2012). Does the processing fluency of a syllabus affect the forecasted grade and course difficulty? *Psychological Reports, 110,* 1–9.

Hirshman, E., & Mulligan, N. (1991). Perceptual interference improves explicit memory but does not enhance data-driven processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 507–513.

Hirshman, E., Trembath, D., & Mulligan, N. (1994). Theoretical implications of the mnemonic benefits of perceptual interference. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 608–620.

Jacoby, L. L. (1978). On interpreting the effects of repetition: Solving a problem versus remembering a solution. *Journal of Verbal Learning and Verbal Behavior, 17,* 649–667. doi:10.1016/S0022-5371(78)90393-6

Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of*

Experimental Psychology. General, 126, 349–370. doi:10.1037/0096-3445.126.4.349

Lindenberger, U., Scherer, H., & Baltes, P. B. (2001). The strong connection between sensory and cognitive performance in old age: Not due to sensory acuity reductions operating during cognitive assessment. *Psychology and Aging, 16,* 196–205.

Maki, R. H., Foley, J. M., Kajer, W. K., Thompson, R. C., & Willert, M. G. (1990). Increased processing enhances calibration of comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 609–616.

Mulligan, N. (1996). The effects of perceptual interference at encoding on implicit memory, explicit memory, and memory for source. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 1067–1087.

Nairne, J. S. (1988). The mnemonic value of perceptual identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 248–255. doi:10.1037/0278-7393.14.2.248

Naveh-Benjamin, M., & Kilb, A. (2012). How the measurement of memory processes can affect memory performance: The case of remember/know judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38,* 194–203.

Nelson, T. O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychological Bulletin, 95,* 109–133.

Oppenheimer, D. M., & Frank, M. C. (2008). A rose in any other font wouldn't smell as sweet: Effects of perceptual fluency on categorization. *Cognition, 106,* 1178–1194.

Plourde, C. E., & Besner, D. (1997). On the locus of the word frequency effect in visual word recognition. *Canadian Journal of Experimental Psychology, 51,* 181–194.

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.

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

Roediger, H. L., III, & Karpicke, J. D. (2006). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science, 1,* 181–210. doi:10.1111/j.1745-6916.2006.00012.x

Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning and Memory, 4,* 592–604. doi:10.1037/0278-7393.4.6.592

Sternberg, S. (1969). Memory-scanning: Mental processes revealed by reaction-time experiments. *American Scientist, 57,* 421–457.

Sungkhasettee, 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

Sweller, J., van Merrienboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10,* 251–296.

Whittlesea, B. W. A., Jacoby, L. L., & Girard, K. (1990). Illusions of immediate memory: Evidence of an attributional basis for feelings of familiarity and perceptual quality. *Journal of Memory and Language, 29,* 716–732. doi:10.1016/0749-596X(90)90045-2

Yap, M. J., & Balota, D. A. (2007). Additive and interactive effects on response time distributions in visual word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 274–296. doi:10.1037/0278-7393.33.2.274