

Picture-Perfect Is Not Perfect for Metamemory: Testing the Perceptual Fluency Hypothesis With Degraded Images

Miri Besken

Bilkent University and Heinrich-Heine-Universität Düsseldorf

The perceptual fluency hypothesis claims that items that are easy to perceive at encoding induce an illusion that they will be easier to remember, despite the finding that perception does not generally affect recall. The current set of studies tested the predictions of the perceptual fluency hypothesis with a picture generation manipulation. Participants identified mixed lists of intact images and images whose certain parts were deleted (generate condition) and made predictions about their subsequent memory performance, followed by a recall test. The intact condition always produced higher memory predictions and shorter identification latencies than the generate condition, consistent with the perceptual fluency hypothesis (Experiments 1 to 3). The actual memory performance for generate images was higher than intact images when aggregate judgments of learning (JOLs) were used (Experiment 1) and equivalent to intact images when item-by-item JOLs were used (Experiment 2 to 3). In Experiment 3, introducing a manipulation that facilitates naming latency for generate images did not increase JOL ratings, providing evidence that not all manipulations that facilitate the ease of perception produce higher JOLs. In Experiment 4, the role of a priori beliefs for the picture generation manipulation was assessed through an online questionnaire. Reading a scenario about the manipulation produced no JOL differences for intact and generate images. The results of the 4 experiments reported here are generally consistent with the perceptual fluency hypothesis of metamemory, and are discussed in terms of experience-based and theory-based processes in metamemory judgments and *Koriat's (1997) cue utilization framework*.

Keywords: perceptual fluency, metamemory, memory, generation, pictures

Metamemory refers to a wide variety of experiential processes, beliefs, and heuristics that people use in order to monitor, control, and regulate their memory (*Koriat & Helstrup, 2007*). These experiential processes, beliefs, and heuristics are important because they guide how cognitive resources will be allocated during the course of learning. For example, even though aging is generally associated with impaired memory performance, older adults can learn to perform almost as well as younger adults on certain types of memory tasks if they monitor their learning performance and pay attention to the information deemed important by the experimenter (*Castel, 2007; Castel, Benjamin, Craik, & Watkins, 2002*).

One common way to assess metamemory is through judgments of learning (JOLs)—participants' predictions of their performance in a subsequent memory test. Even though these predictions sometimes align with actual memory performance, certain cues lead to inaccurate predictions (e.g., *Benjamin, Bjork, & Schwartz, 1998; Koriat, 1997; Susser & Mulligan, 2015*). A recent hypothesis claims that perceptual fluency is one of these cues that produces

disparity between memory predictions and actual memory performance (*Rhodes & Castel, 2008*). According to this hypothesis, if an item is easier to perceive at the time of encoding, participants make higher memory predictions for these items, despite the fact that ease of perception does not generally influence subsequent recall. In order to test this hypothesis, *Rhodes and Castel (2008)* presented participants with large-font (48 font) and small-font (18 font) words, followed by JOLs. Participants produced higher JOLs for large-font words than small-font words, but their recall performance for large- and small-font words did not significantly differ. This finding has been replicated many times in successive studies (e.g., *Kornell, Rhodes, Castel, & Tauber, 2011; McDonough & Gallo, 2012; Susser, Mulligan, & Besken, 2013*). The hypothesis has also been supported through the use of different visual perceptual manipulations. For example, clear and intact words typically produce higher JOLs than their blurred (*Yue, Castel, & Bjork, 2013*) or backward-masked (*Besken & Mulligan, 2013*) counterparts. Similar demonstrations have also been shown in auditory modality. Words presented in high volume or words that are heard intact generally produce higher JOLs than words presented in low volume (*Rhodes & Castel, 2009*) or words that are interspersed with silences (*Besken & Mulligan, 2014; Susser et al., 2013*).

In all these manipulations, easily perceived items produced higher JOLs compared with items whose perception was not as clear, and provided evidence for the perceptual fluency hypothesis. However, all these manipulations have used visual or auditory words as study items. The perceptual fluency hypothesis in metamemory can also be assessed with pictures, a critical omission

This article was published Online First February 4, 2016.

Miri Besken, Department of Psychology, Bilkent University and Mathematical and Cognitive Psychology, Institute of Experimental Psychology, Heinrich-Heine-Universität Düsseldorf.

Correspondence concerning this article should be addressed to Miri Besken, Faculty of Economics, Administrative and Social Sciences, Department of Psychology, Bilkent University, Çankaya 06800, Ankara, Turkey. E-mail: mbesken@bilkent.edu.tr

for two important reasons. First, pictures and words are qualitatively different and encoding differences between them may not impact metacognitive processes to the same degree. Second, degradation for pictures may be compensated differently than degradation for words, which in turn might impact subjective online difficulties in dissimilar ways.

First, pictures and words are hypothesized to be qualitatively different and may engage different types of cognitive processes during encoding. For example, Paivio (1975, 1986) argued that pictures are encoded through both verbal and imaginal codes, whereas words are primarily encoded through verbal codes. Other researchers have argued that pictures are represented more distinctively and elaborately than words (e.g., Mintzer & Snodgrass, 1999; Nelson, Reed, & Walling, 1976). According to this view, both pictures and words access a common semantic code; however, pictures are represented in richer sensory-perceptual codes and have more direct and easier access to this common semantic code than words. Words, on the other hand, activate more phonemic and orthographic processes before accessing semantic codes (e.g., Dewhurst & Conway, 1994; Nelson, Reed, & McEvoy, 1977). Either way, processing verbal materials may be limited to language-specific processes that do not require extensive conceptual processing, which results in representations with impoverished sensory-perceptual details and weaker records of semantic processing. In a similar line of reasoning, Amit, Algom, and Trope (2009) argued that pictures are similar to their physical referents and are subject to the same perceptual processes that are applied to their referent objects, whereas the representation of words are more distant from their referent objects. Accordingly, perceptual analysis of pictures might induce more specific and concrete representations than words. Koriat (1997) argued that certain characteristics of to-be-learned materials may affect JOLs because they serve as cues that reveal the items' a priori ease or difficulty, and he classified imagery as a relatively diagnostic cue of subsequent memory performance. Pictures are generally higher in imagery value than words and induce more specific and concrete representations; thus, they might present participants with cues that are valuable for judging memorability, more so than words. For example, in a recent study, Swahili words were rated as easier to study, to understand, and to link to novel words if they were paired with a picture rather than a word, despite the lack of superior cued-recall performance for picture-word pairs at test (Carpenter & Olson, 2012). This is an initial hint that pictures might induce diagnostic cues that affect metacognitive processes differently than words, even when there are no actual memory differences. Likewise, Ackerman (1981) suggested that picture encoding relies on more perceptual-sensory skills that are mastered earlier in life than visual word encoding. Hence, picture encoding might reflect automatic sensory processes more than word encoding. To the extent that picture encoding is more automatic than word encoding, it should be less amenable to subjective experience-based processes that affect metacognitive judgments.

Second, the subjective ease with which we process degraded stimuli might be qualitatively different for verbal and pictorial materials. Previous research has shown that objective perceptual difficulties may reduce JOLs for verbal materials (e.g., Besken & Mulligan, 2013, 2014). However, the mechanism may not function similarly for our visual world. While processing the visual world around us, we do not always see the objects in their entirety. Some

parts of the objects are blocked by other objects, yet we have little difficulty identifying them (Brown & Koch, 1993). Perhaps people do not experience as much subjective difficulty when they identify degraded images of objects, unlike degraded verbal materials. In this respect, some evidence shows that degraded and intact pictures might produce similar feelings of difficulty, leading to metacognitive errors in source recognition tests. For example, Foley, Foley, Durso, and Smith (1997) presented participants with drawings of objects that were either complete or incomplete at encoding. In a surprise source memory test, participants were asked whether the test item was complete or incomplete during encoding. Participants mistakenly thought that incomplete items were presented in their complete form before, yet this misattribution error did not exist for items presented in their complete form. This finding was replicated in subsequent research (Foley, Foley, Scheye, & Bonacci, 2007). This metacognitive misattribution error might partially be caused by the lack of distinctive recollections about filling in the gaps for the incomplete pictures during encoding. Consequently, this might imply that processing of incomplete pictures feels subjectively less demanding at encoding. Considering that the perceptual fluency hypothesis emphasizes *subjective experiential effort* during encoding, it is not warranted that the perceptual fluency hypothesis will be validated with degraded images. In a different line of work, Westerman, Miller, and Lloyd (2003) showed that when there is a switch from pictures at encoding to verbal material at recognition, participants reduce their use of the fluency heuristics while making recognition judgments. They contended that the fluency heuristic is subject to metacognitive control, and may be attenuated when the participant does not find it useful. The employment of degraded and intact pictorial material at encoding might also affect metacognitive judgments, and the fluency of pictures may not be used as a heuristic to guide metacognitive judgments at encoding. For these reasons, it is important to assess whether the perceptual fluency hypothesis extends to degraded pictorial materials at encoding as well. Thus, the first goal of the current study was to evaluate the perceptual fluency hypothesis for metamemory through the use of perceptually fluent and disfluent images.

There are various methods for making the perception of images disfluent. One can degrade an image by fragmenting through deletion or occlusion (e.g., Foley et al., 2007), by superimposing noise or filter on the image (e.g., Vokey, Baker, Hayman, & Jacoby, 1986), or by backward-masking it (e.g., Breitmeyer, 1984; Intraub, 1984; for a review of other available methods of visual perceptual degradation, see Kim & Blake, 2005). Previous research shows that certain manipulations of image degradation can also enhance memory. For example, in a classic study, Kinjo and Snodgrass (2000) presented participants with a mixed list of intact and generate pictures. In the intact condition, the pictures were presented in their normal form. In the generation condition, half of each picture was deleted. Participants were asked to identify or generate the names of the pictures and remember them for a later memory test. The subsequent recall and recognition tests yielded superior memory performance for the names of generate pictures than for the intact pictures. Typically, the picture generation paradigm produces higher memory performance for the generate items than the intact items, yet its effects on metamemory are not known. If the perceptual fluency hypothesis extends to pictorial material, using an encoding manipulation that reduces memory predictions while increasing memory performance can also assure

that actual memory performance and memory predictions are supported by different underlying mechanisms (Dunn & Kirsner, 1988). The current study uses the picture generation paradigm in order to test the effects of perceptual fluency on memory and metamemory.

The second goal of the study was to test the claims of the perceptual fluency hypothesis by investigating how experience-based and theory-based processes contribute to the effect. Theoretically, two different sources of information are hypothesized to contribute to JOLs: theory-based processes and experience-based processes (Koriat, Bjork, Sheffer, & Bar, 2004). Experience-based processes refer to the specific subjective online difficulties of the participant while encoding specific items. Theory-based processes, on the other hand, denote heuristics and beliefs that participants might have about how their memory operates without reference to a specific learning episode (for a related framework, see analytic and nonanalytic bases of memory predictions in Matvey, Dunlosky, & Guttentag, 2001). The font-size effect, the initial paradigm used for the testing the claims of the perceptual fluency hypothesis, was originally conceptualized as an experience-based process (Rhodes & Castel, 2008). According to this conceptualization, participants experienced categorically different levels of subjective difficulty in the perception of large- and small-font words, and these difficulties mediated participants' JOLs (Bjork, Dunlosky, & Kornell, 2013; Rhodes & Castel, 2008). Yet recent findings provide evidence that the font-size effect might be driven by a *belief* that large-font words are easier to remember than small-font words, rather than by experience-based perceptual difficulties (Mueller, Dunlosky, Tauber, & Rhodes, 2014). Three different lines of proof are offered for this finding. First, there are no processing differences for large- and small-font words, as measured by a lexical decision task. Second, reading scenarios about the font-size experiment induces higher subsequent memory predictions for large-font words compared with small-font words for participants who are not exposed to the experiment. Third, pre-JOLs reveal that participants rate the memorability of large-font words higher than small-font words, even before they are exposed to the stimuli.

Even though the perceptual fluency hypothesis might not account for the font-size effect, this does not rule out that online perceptual difficulties during encoding may affect metamemory judgments for at least some manipulations. For example, Besken and Mulligan (2014) showed that perception (as measured by naming latencies and identification rates) for words interspersed with silences (disfluent condition) was slower and less accurate than intact words (fluent condition) in an auditory generation paradigm. Moreover, this objective measure of perceptual disfluency (as measured by naming latency differences between generate and intact words) was correlated with the decrease in the JOLs produced by this manipulation. Similarly, JOLs for the backward-masked words (disfluent condition) were lower than the intact words (fluent condition) for the perceptual interference paradigm (Besken & Mulligan, 2013), in which backward-masked words are known to produce lower accuracy and slower naming latencies than intact words (Hirshman, Trembath, & Mulligan, 1994), implying a role for the experience-based processes. Accordingly, it is not possible to rule out the effect of experience-based perceptual processes on JOLs, at least for some manipulations.

There are various methods for examining the contribution of theory-based and experience-based processes to JOLs. As mentioned with the previous examples, one method for examining the contribution of theory-based and experience-based processes to the perceptual fluency effect is to measure objective ease of perception through processing latency, and to examine whether these latencies account for the JOL differences between fluent and disfluent items (e.g., Besken & Mulligan, 2013, 2014; Mueller et al., 2014). This method has also been used for other fluency manipulations, such as relatedness (Mueller, Tauber, & Dunlosky, 2013; Undorf & Erdfelder, 2015), identicalness (Mueller, Dunlosky, & Tauber, 2015), and motoric fluency (Susser & Mulligan, 2015).

A second method involves administering questionnaires about possible experimental scenarios with fluency manipulations and asking participants to indicate their predictions about the scenario to disentangle the effects of experience and belief on JOLs (e.g., Mueller et al., 2014). The underlying assumption for this method is that a participant can only use their a priori beliefs about the specific fluency manipulation to guide their memory predictions, as they are not exposed to the procedure. Yet having preexisting beliefs about a manipulation does not always rule out the effect of experience-based processes on JOLs. To explain, one might have a priori beliefs about a manipulation, but these beliefs may still be altered through experience.

Perhaps a more direct way to clarify the contributions of perceptual fluency and beliefs could be realized by dissociating these effects through a manipulation in which participants' awareness of the experience-based processes is reduced or eliminated. Reber, Winkielman, and Schwarz (1998) presented participants with pictures and asked them to rate the prettiness of the pictures. Unbeknownst to the participants, before each picture, a preceding contour that either matched or mismatched the target item was shown subliminally for a brief period of time. Results revealed that both perceptual fluency (as measured by recognition speed) and affective judgments (as measured by the prettiness of the picture) were higher for the pictures preceded by matching contours. Such a manipulation, in which perceptual fluency of the material increases unbeknownst to the participants, could also shed light on metamemory judgments. If identification of materials is faster for the condition with a matching prime than a mismatching prime, and this difference is reflected in the metamemory judgments, this would constitute evidence for the role of experience-based processes. Yet if participants' identification latency is not reflected in their metamemory judgments, this would constitute evidence that experience-based processes may not always be taken into account. The current project aims to examine the contributions of experience-based processes and theory-based processes to metacognitive judgments in encoding pictures comprehensively by employing all the aforementioned methods.

The four experiments reported here used a picture generation paradigm that is modeled after Foley et al. (2007). Participants were presented with mixed lists of intact and generate images. The images consisted of colored photographs of common objects. In the intact condition, the images were presented on the computer screen in their normal form. In the generation condition, approximately 50% of the image was deleted by overlaying a checkerboard pattern, giving the image a fragmented appearance. Participants were asked to type the names of images during encoding and remember them for a later memory test. In the first two experiments, this picture generation paradigm was used along with

a recall test to assess the validity of perceptual fluency hypothesis for pictorial material. Experiment 1 used aggregate JOLs at the end of the list. Experiment 2 used item-by-item JOLs. If the perceptual fluency hypothesis extends to pictorial material, intact images should produce higher memory predictions compared with generate images. Moreover, if this effect is produced by objective experiential difficulties during perception, this should be reflected in identification latencies. To foreshadow, both experiments produced higher JOLs and faster identification latencies for intact items than for generate items, providing evidence in favor of the perceptual fluency hypothesis. Experiments 3 and 4 investigated the relative contribution of experience-based and theory-based processes to this paradigm. Experiment 3 introduced a new variable to the current study design that facilitates experience-based processes, and which might leave theory-based processes unchanged, to investigate the effects of experience-based processes on JOLs, as in Reber et al. (1998). Last of all, Experiment 4 employed an online questionnaire to examine the effects of theory-based processes on subsequent memory predictions, specifically on a priori beliefs about intact and generate images.

Experiment 1

In Experiment 1, participants were presented with mixed lists of generate and intact images. This phase was followed by an aggregate JOL rating at the end of the list and a recall test. The experiment had three goals. The first goal was to see whether introducing a perceptual difficulty by fragmenting the images with deletion throughout the image (rather than deleting one half of the picture, as in Kinjo & Snodgrass, 2000) enhances memory performance. The second goal of the experiment was to see how this perceptual difficulty affects aggregate JOLs at the end of the list. The perceptual fluency hypothesis predicts higher JOLs for intact images (fluent) than generate (disfluent) images. In this case, this picture generation manipulation should produce a double-crossed dissociation between memory and metamemory such that the generation condition reduces JOLs while increasing actual recall. The third goal of the experiment was to see whether the perceptually difficult generation condition was in fact *perceived* later than the intact condition. During the experiment, participants were asked to type in the name of the objects as soon as they identified them. The program recorded participants' latency for first key press and total typing time. If intact images are in fact perceived more fluently, they should produce faster identification latencies, compared with generate images, as measured by first key press latency.

Method

Participants. Twenty-four native speakers of German between the ages of 18 and 35 years from the Heinrich-Heine-Universität Düsseldorf community participated in the experiment. They were compensated with either course credit or a payment of \$3.26. Studies of the generation effect usually use 16 to 24 participants in a within-subjects design (e.g., Kinjo & Snodgrass, 2000). The sample size was estimated through previous studies of the generation effect.

Materials and design. The encoding condition (generate vs. intact) was manipulated within subjects. For the pilot study, 92

colored photographs of everyday items on a white background were selected from two different photograph databases: the Bank of Standardized Stimuli (Brodeur, Dionne-Dostie, Montreuil, & Lepage, 2010) and 360 high-color images (Moreno-Martínez & Montoro, 2012). The images were edited with the GIMP GNU Image Manipulation Program (Version 2.8.10; Kimball & Mattis, 2013). First, each original photograph was resized to 500×500 pixels to guarantee uniformity. These images constituted the intact condition. A generate version of each image was created by superimposing a checkerboard pattern mask on the intact images. The mask consisted of a total of 224 (18×18) squares, with half of the squares transparent to show parts of the original image, and half of them white to fit the background color of the screen and erase parts of the image. The generate versions of the images were piloted with 20 native speakers of German and selected to ensure that the images were identified, named with the same word, and typed within 6 s on the vast majority of trials.

The final material consisted of 36 critical items, along with two practice items at the beginning of the encoding phase to clarify the procedure, four filler items following the practice phase for the primacy effects, and four filler items at the end of the encoding phase for recency effects, adding up to a total of 46 items for the encoding phase.

For a given study list, the critical images were randomly assigned to the intact or generate condition, with the requirement that the average number of letters for the name of the image in each condition did not significantly differ from each other. The order of the critical images was randomized anew for each participant, with the restriction that no more than two items from the same encoding condition were presented consecutively. Two versions of the study list were created, counterbalancing the images across encoding conditions, and they were presented to an equal number of participants. Practice, primacy, and recency items were excluded from all analyses.

Procedure. Participants were tested on individual computers, either alone or in groups of two to four people. The experiment consisted of four phases: the encoding phase, the aggregate JOL phase, the distractor phase, and the testing phase. All of the instructions were presented on the screen. The experimenter answered questions if there were any.

The experiment started with the encoding phase. Participants were given instructions on the screen that they would be presented with two types of images: intact images or images whose certain portions were deleted. They were asked to type in the name of the image as soon as they identified it, and to press "Enter" and remember the images for a later memory test. Participants were told that each trial would last for a total of 6 s and that they should type in the name as quickly as they could, and that spelling mistakes were not important.

Each trial started with a blank screen for 40 ms, followed by an intact or a generate image. In both conditions, once the image was presented on the screen, the participants typed in the name of the image underneath as soon as they identified it and pressed "Enter." Once the participant pressed "Enter," the outer frame of the image turned gray to acknowledge the participant's completion of typing. The program proceeded on to the next trial 6 s after the onset of the image, regardless of whether the participant finished typing or not. Typing rather than vocal responses was used in order to make group testing possible. Two measures of identification were col-

lected from participants: For each trial, the program recorded the first key press latency (the time that elapsed from the onset of the image until the participant's first key press) and the total typing time (the time that elapsed from the participant's first key press until they pressed enter). Response latency measured through typing onset and typing duration has been used as an index of retrieval fluency in previous research (e.g., Serra & Dunlosky, 2005). No feedback was given to the participants about identification accuracy. An example of the two encoding conditions and the procedure for Experiment 1 are displayed in Figure 1.

Following the encoding phase, participants were asked to make predictions about the upcoming memory test separately for intact and generate images. They were told that the list that they had just studied was comprised of images, half presented intact and half generate (the nature of the two types of trials was reminded in detail to elicit these judgments). Participants were prompted to enter a prediction for what percent of each type of word they would recall out of 100. The nature of the memory test was made clear by saying that they would not be presented with the images again during the test phase. The order of the two memory predictions was counterbalanced across participants. After the predictions, participants were given a 3-min distractor task in which they solved arithmetic problems presented on the computer screen one at a time.

The distractor phase was followed by the recall test, in which participants were asked to recall and type in as many words from the encoding phase as possible. Once they typed in the word and pressed "Enter," they could see their response on the screen and proceed on to typing the next word. The time limit for the recall

phase was 5 min, but it could also be self-terminated earlier by pressing the "ESC" key.

Results

All descriptive statistics for Experiment 1 are presented in Table 1. For all analyses, the alpha level was set at .05. Participants must have identified and completed typing at least 80% of the generate study items in order to be included in this and subsequent experiments. All participants satisfied the necessary prerequisite. During the encoding phase, the identification rates for both intact pictures (99%) and generate pictures (99%) were quite high. A sign test revealed that the pictures were identified equally well in both conditions ($p = 1.00$).

For encoding, the median reaction times (RTs) for the first key press and the total typing time were calculated for each participant separately for intact and generate items, excluding trials on which participants misidentified the critical images or did not type at all (.7%). A paired samples t test showed that the encoding condition had a significant effect on first key press latency, $t(23) = 7.04$, $p < .001$, Cohen's $d = 2.04$. Participants started typing the names of intact items faster than generate items. Thus, it took participants longer to identify the generate images. The encoding condition had no significant effect on the total typing time, $t(23) = .14$, $p = .89$, $d = .04$, showing that once participants identified the image, their typing speed was equivalent for intact and generate items.

High identification rates at study make it possible to analyze recall data without conditionalizing on correct study identification. Because unconditionalized and conditionalized produce the same

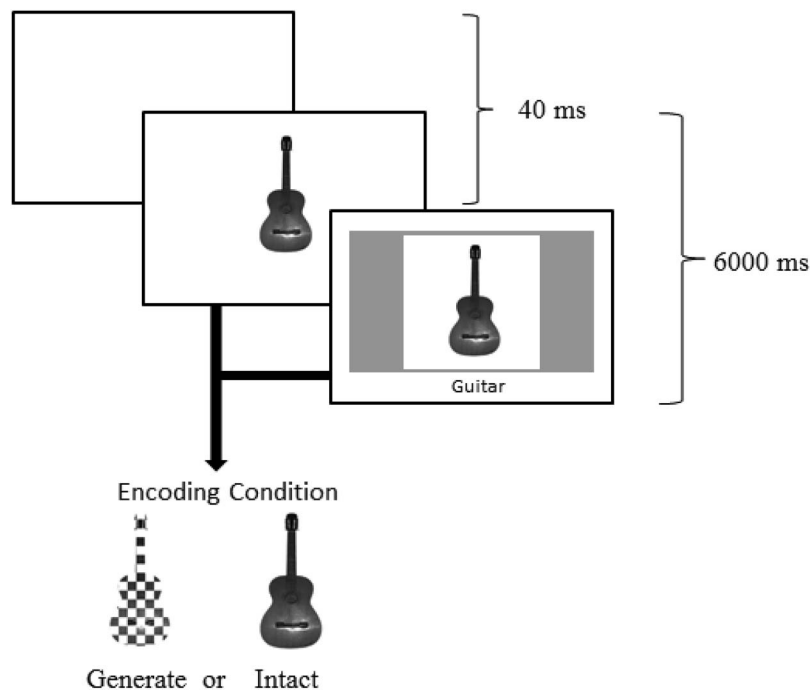


Figure 1. The general design, encoding conditions, and procedure for Experiment 1 and Experiment 2. Experiment 1 used aggregate JOLs at the end of the encoding phase, and Experiment 2 used self-paced item-by-item JOLs after each image. The images used for encoding condition were colored in Experiments 1 and 2. JOLs = judgments of learning.

Table 1
Means and Standard Deviations (in Parentheses) for Intact and Generate Images for Experiments 1, 2, and 4

Exp. #	Identification latency (ms)		Total typing time (ms)		Memory predictions (out of 100)		Proportion correct recall	
	Intact	Generate	Intact	Generate	Intact	Generate	Intact	Generate
1	1,152.67 (274.67)	1,372.27 (282.33)	1,694.21 (1,366.39)	1,688.10 (1,286.18)	44.04 (19.07)	37.96 (19.00)	.45 (.11)	.54 (.13)
2	1,049.06 (131.43)	1,273.00 (156.04)	1,721.73 (1,043.10)	1,757.92 (980.65)	63.50 (20.24)	59.14 (20.19)	.52 (.12)	.52 (.16)
4	—	—	—	—	65.11 (18.89)	62.33 (18.20)	—	—

Note. Exp. = experiment.

results, only unconditionalized recall data are reported for this and all subsequent experiments. A paired samples *t* test revealed a main effect of encoding condition on the aggregate JOLs, $t(23) = 2.77$, $p = .011$, $d = .80$, with higher recall predictions for intact items than generate items. However, participants recalled a higher portion of items from the generate condition than the intact condition, $t(23) = 2.86$, $p = .01$, $d = .85$. The current experiment demonstrates the typical generation effect for the memory performance: Participants have higher recall performance for images with missing fragments (generate condition) than images presented in their normal, intact form. This finding mirrors the findings of verbal materials in visual and auditory modality, in which introducing a perceptual difficulty, such as deleting letters (e.g., Glisky & Rabinowitz, 1985; Nairne, Riegler, & Serra, 1991) or replacing certain parts of the words with silences, enhances memory performance (Besken & Mulligan, 2014). It also replicates Kinjo and Snodgrass's (2000) picture generation effect, using a different method for fragmenting the images, providing further proof that the memory enhancement for generate material extends beyond verbal materials.

Nevertheless, participants are clearly not aware of the benefits of introducing perceptual difficulties during the learning of images. On the contrary, they believe that perceptual disfluency will affect actual memory performance negatively. This finding is consistent with the perceptual fluency hypothesis, which argues that higher perceptual fluency is indicative of higher memory predictions. Along with the memory performance, this set of results forms a double dissociation in which a picture generation manipulation decreases memory predictions while increasing actual memory performance, adding to the body of literature in which perceptual difficulties produce double-crossed dissociations between metamemory and memory (Besken & Mulligan, 2013, 2014).

When aggregate JOLs at the end of the list are used to assess metamemory, they may not always reflect the online difficulties that participants experience during a specific episode (Koriat et al., 2004); thus, higher aggregate JOLs for intact than generate items might be a consequence of a priori beliefs about the memorability of intact and generate images. Yet the current experiment provides objective measures of perceptual difficulties. Participants' identification latencies, as measured by the first key press latency, are slower for generate images than for intact images, implying that participants take a longer time in processing images with missing fragments. Moreover, once they identify the name of item denoted in the image, their total typing time does not differ significantly for intact and generate items, showing that the effect is specific to initial identification. Thus, even if the aggregate JOLs might

represent theory-based processes, participants experience objective difficulties in identifying generate pictures. This is consistent with Besken and Mulligan's (2014) findings for auditory generation effect, in which naming latencies for intact words presented over the headphones are faster than words that are interspersed with silences. Most previous studies assessing the merits of the perceptual fluency hypothesis have not used objective measures of perceptual fluency, and no studies have used it for pictorial materials.

Experiment 2

Experiment 1 showed that the effect of the picture generation manipulation on aggregate JOLs at the end of the list is consistent with the perceptual fluency hypothesis, and that generate items are in fact processed more slowly than intact items. However, JOLs can also be assessed on an item-by-item basis immediately after the presentation of each item. It is important to find out whether the same pattern of results would also be found with this measure, especially considering that some researchers have proposed that aggregate and item-by-item JOLs might rely on different mechanisms (e.g., theory-based vs. experience-based information; Koriat et al., 2004). If the perceptual fluency hypothesis applies to immediate item-by-item JOLs, then intact images should receive higher memory predictions than generate images and should provide further support for online perceptual difficulties.

Identification latencies for intact and generate images, as measured by first key press latency, were routinely recorded in Experiment 2 to measure objective perceptual fluency. As in Experiment 1, identification latencies should be faster for intact images than for generate images. In addition to providing evidence for an objective measure of perceptual fluency, identification latencies can also be used to assess the stipulations of the perceptual fluency hypothesis. If experience-based processes are influential in JOLs, to the extent that the generation manipulation slows down perception (experience-based processes), a parallel reduction should be observed in the JOLs for those items. The correlational and mediational analyses in this experiment will assess the contribution of the experience-based and theory-based processes to metamemory judgments.

When aggregate JOLs at the end of the list are used, memory performance for generate items is usually higher than for intact items. However, previous research showed that item-by-item JOLs may sometimes eliminate the memory superiority for generate items (Begg, Vinski, Frankovich, & Holgate, 1991; Besken & Mulligan, 2013, 2014; Matvey et al., 2001). If the present manipulation relies on a similar mechanism, immediate item-by-item

JOLs may reduce or eliminate the memory superiority for the generation condition.

Method

Participants. Twenty-four native speakers of German from the Heinrich-Heine-Universität Düsseldorf community participated in the study for course credit or a compensation of \$3.26.

Materials, design, and procedure. The materials were identical and the design was similar to Experiment 1. The only difference was that immediate item-by-item JOLs were used rather than aggregate JOLs at the end of the list. Participants were instructed to type the name of the object depicted in the image within 6 s, followed by a JOL for each item immediately after each trial, indicating their confidence for their subsequent memory performance from 0 (*not confident at all*) to 100 (*very confident*) for the item that they have just seen. The nature of the memory test was made clear by indicating that they would not be presented with the images again. The item-by-item JOLs were self-paced. Pressing “Enter” initiated the next trial. The distractor and the testing phases were identical to Experiment 1.

Results

All descriptive statistics for Experiment 2 are presented in Table 1. All participants satisfied the necessary prerequisite of 80% identification rate and were included in the experiment. During the encoding phase, the identification rates for both intact images (99%) and generate images (99%) were quite high. A sign test for paired samples revealed that the images were identified equally well in both conditions ($p = .453$).

For median identification latency, the trials in which participants misidentified the critical images or in which they did not type at all were excluded (1.1%). Paired-samples t tests showed that first key press latency was significantly faster for intact images than generate images, $t(23) = 12.79$, $p < .001$, $d = 3.84$. The total typing time did not differ by encoding condition, $t(23) = .93$, $p = .362$, $d = 0.28$.

For item-by-item JOLs, trials in which participants misidentified the critical images or failed to type in the name of the object or failed to type in the JOL were excluded (1.3%). As in Experiment 1, participants produced significantly higher JOLs for intact items than for generate items, $t(23) = 3.69$, $p = .001$, $d = 0.97$. However, their recall performance did not differ by encoding condition, $t(23) = 0.08$, $p = .937$, $d = 0.00$. The results of Experiment 1 show that the picture generation manipulation produces an effect size of $d = .85$ for recall. The power of the present analysis to detect an effect of that size ($n = 24$, $\alpha = .05$, one-tailed) is .93.

The perceptual fluency hypothesis predicts that, to the extent that the generation manipulation produces disfluency, a parallel reduction in JOLs should also be observed. In order to test this, two scores were calculated: the difference between the median identification latencies between generate and intact trials (Identification Latency_{intact} – Identification Latency_{generate}; $M = 223.94$, $SD = 85.81$), and the difference between the mean item-by-item JOLs for generate and intact trials (JOL_{intact} – JOL_{generate}; $M = 4.36$, $SD = 5.79$). The correlation between the identification latency difference and JOL difference was $r(22) = -.44$, $p = .03$.

This indicates that the bigger the difference was between identification latencies for intact and generate items, an associated reduction was also observed for JOLs. In other words, the reduction in fluency was correlated with the reduction in JOLs, providing support for the perceptual fluency hypothesis.

Another issue relevant to the contribution of experience-based and theory-based processes to JOLs for this manipulation can also be assessed through mediation analyses, using the identification latencies for intact and generate items. To explain the relationship between perceptual fluency and JOLs, one can compute point-biserial correlations between JOLs and encoding condition (intact vs. generate) at the item level within subjects. For each participant, a point-biserial correlation was computed, with the encoding condition dummy coded (0 = generate, 1 = intact). A positive correlation between encoding condition and JOLs for each participant indicates that as the level of fluency increases, JOLs increase as well. This is only another way of demonstrating that intact items receive higher JOLs than generate items on average, as previously shown with paired samples t test. These within-subjects correlations reflect all the factors that contribute to making JOLs, including both experience-based processes and theory-based processes. If perceptual fluency (an instance of experience-based processes) makes a direct impact in this relationship, then partialing out the effect of fluency (as measured by identification latency) should significantly decrease this correlation for each participant. However, if the relationship is driven by theory-based processes about the memorability of a class of items, then the correlation should remain the same. Alternatively, it is possible that both fluency and beliefs contribute to the JOLs independently. In this case, partialing out the identification latency should reduce, but not eliminate, this correlation. As expected, the mean within-subjects interitem correlation between encoding condition and JOLs was positive ($M = .12$, $SD = .15$) and significantly different from zero, $t(23) = 3.76$, $p = .001$, $d = .77$, confirming the presence of average JOL differences between encoding conditions. The same analysis was conducted again, this time using the interitem correlations between encoding condition and JOLs for each participant, partialing out the effect of identification latencies on each item to control for the effect of perceptual fluency. The mean within-subjects interitem correlation was .08, with a standard deviation of .16. The pairwise t -test comparison of the first mean correlation between encoding condition and JOLs and the second mean correlation, which partialled out the effect of perceptual fluency, was significant, $t(23) = 2.30$, $p = .031$, $d = .66$. This finding indicates that perceptual fluency has a unique contribution to JOLs above and beyond other factors. However, the mean partial within-subjects interitem correlation also differed significantly from zero, $t(23) = 2.47$, $p = .021$, $d = .51$, indicating that even when the effect of perceptual fluency is controlled, participants still gave higher ratings to intact images than generate images. This implies that there are other factors than experience-based processes, which contribute to JOLs. A priori beliefs about the memorability of intact and generate items, or beliefs that were formed during the course of the experiment about intact and generate items, might explain the remainder of the positive correlation between encoding condition and JOLs.

As in Experiment 1, participants gave higher confidence ratings for intact images compared with generate images, providing further support to the perceptual fluency hypothesis for item-by-item JOLs. This result demonstrates that the perceptual fluency hypoth-

esis extends beyond simple verbal materials and covers more complex materials, such as pictures, which are hypothesized to have richer sensory-perceptual details and stronger records of semantic processing (e.g., Dewhurst & Conway, 1994). More importantly, the identification latencies for generate items were slower than intact items, providing proof that perceptual fluency, as measured by identification latencies, is indeed related to JOLs, consistent with Besken and Mulligan (2014).

Most importantly, the data set of this experiment made it possible to investigate the relationship between perceptual fluency, identification latencies, and JOLs. The perceptual fluency hypothesis emphasizes the importance of experience-based processes in making JOLs. By this token, the more difficult it is for a participant to identify the disfluent items compared with a baseline, the lower their JOLs should be compared with the baseline. The negative correlation between average JOL difference and average identification latency difference for generate and intact item suggests that participants face experience-based difficulties while perceiving the items. Still, this analysis only reveals that JOLs and identification latencies are highly and significantly correlated. Further mediational analysis showed that both experience-based and theory-based processes contribute to the JOLs independently. Thus, even when experience-based perceptual fluency manipulation affects JOLs, beliefs also make an independent contribution to metamemory above and beyond objective perceptual fluency differences.

With regard to actual recall performance, requiring participants to provide item-by-item JOLs eliminated the superior memory performance for generate items. This is consistent with previous research using visual (Begg et al., 1991; Besken & Mulligan, 2013; Matvey et al., 2001) and auditory (Besken & Mulligan, 2014) perceptual manipulations. Sometimes, item-by-item JOLs eliminate the typical generation effect for memory. One possible explanation is that JOLs might induce deeper encoding for the intact items as well, eliminating the difference between encoding conditions (Matvey et al., 2001).

Experiment 3

The mediational analyses in Experiment 2 showed that both experience-based and theory-based processes contribute to JOLs. Certainly, employing these analyses to show the relationship between perceptual fluency, identification latencies, and JOLs is useful for examining the independent contribution of each of these processes. However, the mediational analyses do not precisely explain *how* perceptual difficulties affect JOLs in relation to experience-based processes and theory-based processes. There are two possible explanations for these findings. The first explanation is that perceptual fluency directly affects JOLs. This explanation is coherent with a straightforward interpretation of the perceptual fluency hypothesis (Rhodes & Castel, 2008), but the hypothesis itself makes no explicit claims about this point. Alternatively, the experience of perceptual difficulty might influence a participant's theories about memory and induce a belief that a category of items that are difficult to perceive will be remembered less in subsequent memory tests; thus, beliefs make an independent contribution to this metamemorial illusion. Besken and Mulligan (2014) and Experiment 2 of this article provide some indirect proof for this alternative explanation. A similar type of argument has been made

for fluency manipulations that affect other types of judgments (Alter & Oppenheimer, 2009).

Perhaps a more direct way to clarify how the perceptual fluency hypothesis works is to dissociate the effects of experience and belief through a manipulation in which participants' awareness of the experience-based processes is reduced. As in Experiment 2, participants were presented with intact and generate images to be identified, followed by item-by-item JOLs. In addition, a new variable was introduced to the design in Experiment 3. Before each critical image was displayed, they were shown a preceding blackened contour that either matched or mismatched the critical image for a very brief period (see Figure 2 for an instance). In a similar design, Reber et al. (1998) found that when an image is primed with a contour that matched the critical image, the recognition was faster for matching items than mismatching items. In the current design specifically, a preceding contour that matches the critical generate image should facilitate the perception of the image, compared with a mismatching preceding contour, in line with findings of Reber et al. In other words, the preceding contours should facilitate the identification latencies for generate items if they match the critical image, producing a change in experience-based processes. If the straightforward explanation of the perceptual fluency hypothesis is valid, this facilitation in experience-based processes should also be directly reflected in participants' JOLs such that participants should give higher JOLs to items preceded by matching contours than those preceded by mismatching contours. This, in turn, should produce a linear relationship between identification latencies and JOLs.

However, if the alternative explanation of the perceptual fluency hypothesis is valid and perceptual difficulties induce a belief in which a class of perceptually disfluent items are given lower ratings than fluent items, a dissociation between identification latencies and JOLs should emerge such that participants give higher ratings to intact items than generate items; yet there should be no effect of the preceding contour type on the JOLs even though their identification latencies are faster for matching than mismatching contours, because participants may not be aware of the facilitatory experiential effect of matching preceding contours on generate items. Such a dissociation between identification latencies and JOLs shows that participants are not always able to take objective experience into account, especially when they cannot thoroughly perceive the preceding image.¹ This possibility provides support for the belief-mediated perceptual fluency hypothesis.

¹ There are two reasons to claim that participants may not thoroughly process the preceding contour. First, they do not see the preceding images in their normal form; they only see blackened contours, which they might or might not be able to identify, even if they are presented with them for a longer duration. Second, the behavioral and neuroimaging studies of visual perception, which use colored images of simple objects and scenes, show that processing of these stimuli generally takes around 100 to 200 ms (e.g., Martinovic, Gruber, & Müller, 2008; Thorpe, Fize, & Marlot, 1996). The preceding contours in Experiment 3 are presented for a duration of 40 ms, which is considerably shorter than the hypothesized time to process an image.

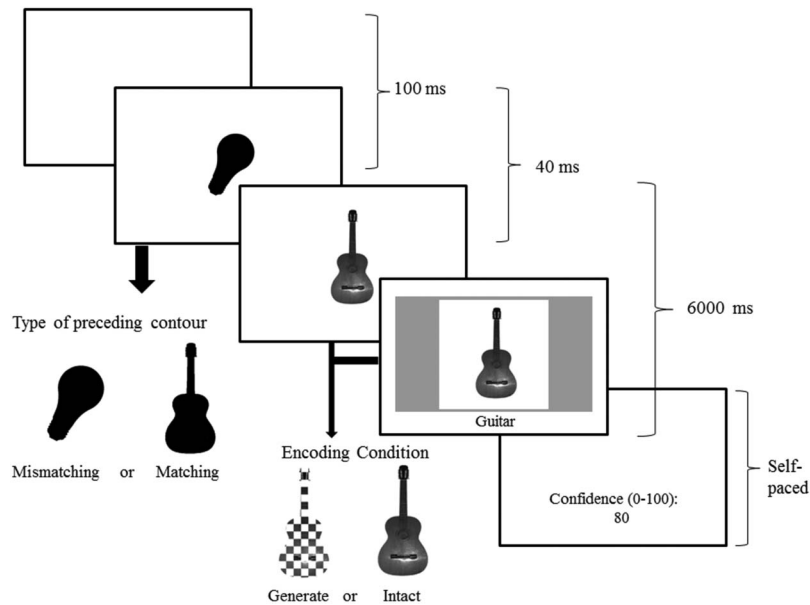


Figure 2. The design, material, and procedure for Experiment 3. The critical intact or generate image was preceded by a blackened contour, which either matched or mismatched the critical image, as depicted in the figure. The contour was presented for the duration of 40 ms. The images used for encoding condition were colored.

Method

Participants. Twenty-four native speakers of German from the Heinrich-Heine-Universität Düsseldorf community participated in the study for course credit or a compensation of \$3.26.

Materials and design. The design was a 2 (encoding condition: generate vs. intact) \times 2 (preceding contour: matching vs. mismatching) within-subject design. The critical and filler items were identical to Experiment 1 for the generation manipulation, with two extra items for the practice phase. Each critical item was preceded by an image, which was either the blackened contour of the critical item (matching preceding contour) or the blackened contour of an item from another category that has not been used as a critical image (mismatching preceding contour). The preceding contours comprised of the blackened images of 66 photographs from the pilot study (explained in Experiment 1), including 36 matching, 18 mismatching, and 12 filler contours. All preceding contours were painted black using the image manipulation program GIMP.

The two study lists from Experiment 1, which assigned critical items to intact or generate conditions, were divided further such that the images in both encoding conditions were preceded either by a matching or a mismatching contour. The lists were counterbalanced such that an item preceded by a matching contour in one version of the list was preceded by a mismatching contour in the other version. Thus, both encoding condition and preceding contour type were counterbalanced across participants, and each critical image was presented in all possible combinations to an equal number of participants. The order of items was randomized anew for each participant, with the restriction that no more than two items from the same encoding condition was presented consecutively.

Procedure. The instructions for Experiment 3 were similar to Experiment 2. The only difference was that participants were also told that they would see some contours for a very brief period before each critical image was presented, and they just needed to keep their eyes on the screen and start typing once they identified or recognized the critical image.

Each trial started with a blank screen for 100 ms, followed by the presentation of a matching or a mismatching blackened contour for 40 ms. This was followed by the critical image for 6,000 ms, requiring participants to type in the name of the item, followed by a self-paced item-by-item JOL rating. The program recorded participants' first key press latency and the total typing time, as in previous experiments. An example of the materials and procedure for Experiment 3 is displayed in Figure 2. The distractor and the testing phases were identical to previous experiments.

Results

All descriptive statistics for Experiment 3 are presented in Table 2. All participants satisfied the necessary prerequisite of 80% identification rate and were included in the experiment. During the encoding phase, the identification rates for all conditions were quite high and not significantly different from each other by a nonparametric Friedman test of differences, $\chi^2(3) = 1.07, p = .79$.

The median RTs for the first key press and the total typing time were calculated, excluding trials on which participants misidentified the critical images or they did not type in at all (1.3%). A repeated measures analysis of variance (ANOVA) with Encoding Condition and Type of Preceding Contour as the repeated factors showed a main effect for first key press latency, $F(1, 23) = 25.23$, *Mean Squared Error (MSE)* = 20910.43, $p < .001$, $\eta_p^2 = .52$, with first key press latency faster for intact images than generate im-

Table 2
Means and Standard Deviations (in Parentheses) for Intact and Generate Images That Are Preceded by Matching and Mismatching Contours for Experiment 3

Preceding contour type	Item type			
	Intact		Generate	
	Matching	Mismatching	Matching	Mismatching
Identification rate	.99 (.03)	.99 (.04)	.99 (.03)	.98 (.04)
Response time (ms)	1,084.04 (350.37)	1,063.38 (256.65)	1,162.85 (317.40)	1,281.1 (355.64)
JOLs (out of 100)	61.18 (23.56)	62.23 (21.25)	58.16 (22.59)	58.03 (19.53)
Proportion correct	.50 (.19)	.53 (.18)	.49 (.19)	.54 (.23)

Note. JOLs = judgments of learning.

ages. There was no main effect for the type of preceding contour, $F(1, 23) = 2.38$, $MSE = 23991.33$, $p = .136$, $\eta_p^2 = .094$. However, the interaction between encoding condition and type of preceding contour was significant, $F(1, 23) = 4.47$, $MSE = 25898.03$, $p = .046$, $\eta_p^2 = .163$. Follow-up paired samples t tests revealed that for intact images, the presentation of preceding contours did not affect first key press latency, $t(23) = .65$, $p = .655$, $d = 0.14$; yet for generate images, the presentation of a matching contour facilitated the first key press latency compared with a mismatching contour, $t(23) = 2.60$, $p = .016$, $d = 0.76$.

As in previous experiments, total typing time did not significantly differ by encoding condition, $F(1, 23) = .06$, $MSE = 153568.73$, $p = .812$, $\eta_p^2 = .003$, or type of preceding contour, $F(1, 23) = .16$, $MSE = 53707.32$, $p = .692$, $\eta_p^2 = .007$, or by their interaction, $F(1, 23) = .24$, $MSE = 129143.24$, $p = .631$, $\eta_p^2 = .010$.

For item-by-item JOLs, trials on which participants misidentified the critical images or failed to type in the name of the object or the JOL were excluded (1.7%). A repeated measures ANOVA for JOLs revealed a main effect for encoding condition, $F(1, 23) = 5.76$, $MSE = 54.24$, $p = .025$, $\eta_p^2 = .20$, with higher JOLs for intact items than generate items. There was no main effect for type of preceding contour, $F(1, 23) = .122$, $MSE = 42.19$, $p = .730$, $\eta_p^2 = .005$, and no interaction, $F(1, 23) = .199$, $MSE = 41.94$, $p = .660$, $\eta_p^2 = .009$.

A repeated measures ANOVA for the unconditionalized recall revealed no significant main effect for encoding condition, $F(1, 23) = 0.005$, $MSE = 0.028$, $p = .946$, $\eta_p^2 = .00$, no significant main effect for type of preceding contour, $F(1, 23) = 1.39$, $MSE = 0.027$, $p = .250$, $\eta_p^2 = .057$, and no significant interaction between these factors, $F(1, 23) = 0.029$, $MSE = 0.040$, $p = .867$, $\eta_p^2 = .001$. The encoding condition on recall in Experiment 1 produces $\eta_p^2 = .26$, revealing an effect size of $f = .59$. The power of the current analysis to detect an effect of that size is .98 ($n = 24$, $\alpha = .05$).

An important issue that pertains to the goals of this project is to disentangle the effects of experience-based and theory-based processes in making JOLs. As explained in Experiment 2, participants' individual JOL ratings may depend on both of these processes, and partialing out the effect of experience-based processes for each participant and comparing this with the total correlation between encoding type and JOLs should inform us about the contribution of experience-based processes. In order for Experiment 3 to be comparable with Experiment 2, a similar type of

mediational analysis was conducted. Because there was the additional variable of preceding contour type, the analysis was conducted in three steps rather than two steps. In the first step, a within-subjects interitem correlation was calculated between encoding condition (generate = 0, intact = 1) and JOLs for each participant, as in Experiment 2. The mean within-subjects interitem correlation ($M = .10$, $SD = .19$) differed significantly from zero, confirming the presence of average JOL differences between intact and generate items, $t(23) = 2.51$, $p = .019$, $d = .51$, as also shown previously with repeated-measures ANOVA. In the second step, a within-subjects interitem correlation was calculated between encoding type (generate = 0, intact = 1) and JOLs for each participant, controlling for the preceding contour type (dummy coded as mismatching = 0, matching = 1). If the preceding contour type makes a contribution to the total mean correlation between encoding type and JOLs, partialing out the effect of preceding contour should lower the interitem correlation for each participant, leading to a significant difference between the initial mean correlation and the correlation controlling for the effect of preceding contour. The partial mean correlation between encoding type and JOL, controlling for preceding contour ($M = .10$, $SD = .20$), did not differ significantly from the total mean correlation, showing that type of preceding contour does not contribute to JOLs, $t(23) = .90$, $p = .375$, $d = .00$. In the last step, a third within-subjects interitem correlation for each participant was calculated, this time controlling for both preceding contour type and identification latency.² If perceptual fluency is making a direct impact on JOLs, partialing out the effect of identification latency and preceding contour at the same time should significantly lower the correlation, compared with the mean correlation that controls only for the preceding contour. A paired samples t test showed that controlling for both preceding contour type and identification latency lowered the mean correlation ($M = .05$, $SD = .18$) significantly, $t(23) = 3.85$, $p = .001$, $d = 1.17$, providing evidence that perceptual fluency mediates the effect of encoding type on JOLs for the picture generation manipulation. The remaining mean correlation did not significantly differ from zero, $t(23) = 1.32$, $p = .199$, $d = .27$, showing that beliefs formed before the experiment or beliefs formed during the course of the experiment did not

² A similar analysis was conducted controlling only for identification latencies, rather than controlling for both variables. The results mirror the results explained here.

significantly contribute to making JOLs for the encoding condition.

A similar mediation analysis was conducted in three steps to examine the relationship between preceding contour type (mismatching = 0, matching = 1) and JOLs, looking at (a) the total mean correlation ($M = .02$, $SD = .20$), (b) the total mean correlation controlling for encoding condition ($M = .02$, $SD = .21$), and (c) the total mean correlation controlling for encoding condition and identification latency ($M = .01$, $SD = .18$). None of the correlations were significantly different than zero, $t_s < 1$. They also did not significantly differ from each other, $t_s < 1.2$.

The current experiment investigated whether experience-based processes affect JOLs directly or by mediating beliefs about fluency. To accomplish this, a new variable was introduced to the design. This variable affected objective measures of perceptual fluency, even though its influence on perceptual fluency may not have been very obvious to the participants. A direct effect of experience-based processes on JOLs requires an associated change in JOLs in the same direction. As in previous experiments, the perceptually disfluent generate condition produced slower identification latencies than the fluent intact condition, reflected in the JOLs. The mediational analyses yielded that perceptual fluency, as measured by identification latencies, accounted for most of the JOL differences between generate and intact images. In other words, just like in Experiment 2, presentation of intact and generate images made a unique contribution to JOLs above and beyond beliefs. However, identification latencies were also facilitated for generate images that were preceded by a matching contour compared with a mismatching contour. Yet the effect of the preceding contour manipulation that affected identification latencies was not reflected in the JOLs, ruling out the possibility that every manipulation that changes objective perceptual fluency also affects JOLs. The mediational analyses also showed that when participants are making JOLs, they do not consider the effects of preceding contour manipulation at all. This finding has two implications. First, in order for the participants to make JOLs based on perceptual fluency, the manipulation needs to be conspicuous to the participants. If the participants cannot perceive the manipulation on a conscious level, their JOLs are not affected by the manipulation. Second, this finding also demonstrates that perceptual fluency does not always impact JOLs directly. Objective measures of perceptual fluency might partially operate by mediating subjective beliefs about the subsequent memory performance. Experiment 3 still shows that experience-based processes are influential, because the identification latencies for generate items are in fact slower than intact items, and this difference mediates JOLs for the encoding condition. However, the dissociation between identification latencies and JOLs for the preceding contour manipulation constitutes evidence for the indirect belief-mediated effect of perception on metacognitive judgments (Mueller et al., 2014).

Some previous research has also shown that the differences in perceptual fluency may not always affect memory decisions (e.g., Miller, Lloyd, & Westerman, 2008; Westerman, Miller, & Lloyd, 2003). For example, Westerman et al. (2003) showed that a change from pictures at encoding to words at recognitions attenuates the use of perceptual fluency heuristic such that participants do not rely on perceptual fluency to recognize items. Westerman et al. (2003) categorized this as an instance of metacognitive control and

contended that participants might have beliefs about fluency that are difficult to change. In the current study, the finding is actually more indicative that if the participants are not aware of a perceptual fluency manipulation, they may not be able to develop beliefs about it; accordingly, their memory predictions will not be affected. However, once they can consciously perceive the manipulation, the experience-based processes can still mediate the relationship between type of item and JOLs.

Experiment 4

Experiment 3 showed a dissociation between identification latencies and memory predictions. The preceding contour manipulation, which might not have been very noticeable to participants, affected objective measures of perceptual fluency, but this difference was not reflected in JOLs at all. The more obvious encoding condition (intact vs. generate) manipulation also affected objective measures of perceptual fluency, and this difference was reflected in the JOLs. In other words, the impact of the two variables, both of which produced consistent objective perceptual fluency differences, was inconsistent on JOLs. This implies that at least some of the perceptual fluency effects can be explained by beliefs rather than objective experiences.

Koriat's (1997) cue utilization framework claims that participants have a priori beliefs about the effects of certain manipulations that affect JOLs. For example, one might believe that intact pictures will produce better memory than generate pictures, because intact pictures are easier to learn. Yet another possibility is that beliefs, an instance of theory-based processes, might be activated during the course of learning as the participants are exposed to the materials. Experiments 1 to 3 showed that beliefs explain some of the processes used to make JOLs. In all these experiments, beliefs made an impact on JOLs, yet it is not clear when these beliefs are formed. One possibility is that participants already have a priori beliefs about the effects of the picture generation manipulation on memory before the experiment begins. Another possibility is that these beliefs begin to form during the course of encoding intact and generate images, when participants are prompted for JOLs. The goal of this experiment was to determine whether the participants have to go through the experiment in order to give higher JOLs to intact items than generate items, or they hold an a priori belief that intact items are more memorable than generate items even before carrying out the experiment.

As indicated in the introduction, a straightforward methodology to assess a priori beliefs about memory is using a questionnaire in which participants read a scenario about a memory experiment and are asked to make predictions about different conditions (Koriat et al., 2004; Kornell et al., 2011; Mueller et al., 2014; Susser & Mulligan, 2015). For example, Mueller et al. (2014) presented participants with a scenario about a memory experiment for large- and small-font words and asked them to predict memory performance for these two conditions. The results revealed higher memory predictions for large-font words than small-font words, indicative of a priori beliefs about the memorability of words, because the participants had this belief even though they did not go through the experimental procedure.

In Experiment 4, an online questionnaire was used to assess whether participants had a priori beliefs about intact and generate images without being exposed to the experimental procedure. All

participants were given the same scenario and asked to make to predictions about their subsequent performance. If participants have a priori beliefs about the memorability of intact and generate items, intact items should receive higher memory predictions than generate items in the questionnaire. However, if higher predictions for intact images compared with generate images are mediated by online perceptual difficulties during encoding, there should be no difference between the encoding conditions.

Method

Participants. A total of 117 participants from the Heinrich-Heine-Universität Düsseldorf community participated in the questionnaire for a chance of winning a \$10.87 gift card.

Materials, design, and procedure. The data were collected through an online questionnaire. The online questionnaire was advertised by students taking an experimental practicum class at the university, as well as on various university Facebook pages. Participants could choose to participate by clicking on the online questionnaire link. Once they consented to participating, they read the following description in German:

Suppose that you were participating in a memory experiment. In the encoding phase of the experiment, you were presented with photographs of common items and asked to remember the names of the items in the photographs for a later memory test. Each item was presented for 6 seconds one at a time on the screen. Half of these photographs were presented in their normal form, with the photographs presented on the screen as a whole. The other half of the photographs were presented, with some parts of the photographs missing. You were able to identify the names of all of the photographs. Then, you were given a memory test five minutes later, and you were asked to recall the names of the items from the first part of the experiment as accurately as possible. You were not presented with the items again.

Please make two predictions about your memory performance. What percent of intact items do you think you would remember? What percent of fragmented (generate) items do you think you would remember? Please make both predictions out of 100. Indicate the percentages for each one in the slider below.

Participants were randomly assigned to one of the two conditions: In one condition, examples of one intact and one generate image accompanied the preceding description, whereas in the other condition, no examples were shown. The slider scale position for intact and generate items was counterbalanced across participants. Once participants made their judgments, they proceeded on to the next page. The next page of the questionnaire assured that participants chose the correct explanation for their choice (e.g., “I thought that I would remember the generate photos better than the intact photos”). The confirmation was followed by a screen where they were asked to indicate their reason for their prediction. This question was followed by demographics about native language, age, and whether they had participated in one of the previous experiments about this topic at Heinrich-Heine-Universität Düsseldorf.

Results

Eighteen participants were excluded from the study for not completing the survey ($n = 14$) or for not being a native speaker

of German ($n = 3$), or for having participated in one of the previous experiments ($n = 1$). Another eight participants were excluded for failing to give consistent responses on the questionnaire (e.g., they marked intact more than generate, even though their percentages were the other way around in the previous page). The remaining 91 participants had a mean age of 21.42 years, with a standard deviation of 2.43, and were all native speakers of German.

For the group who was shown examples of the material, the mean memory prediction for intact pictures was 65.67, with a standard deviation of 18.75, and the mean was 62.83, with a standard deviation of 17.35, for generate pictures ($n = 42$). For the group that was not shown any examples of the material, the mean memory prediction for intact pictures was 64.63, with a standard deviation of 19.18, and the mean was 61.90, with a standard deviation of 19.07, for generate pictures ($n = 49$). All other descriptive statistics for Experiment 4 are presented in Table 1. The means for predictions for each encoding condition (intact and generate) were calculated separately and submitted to a mixed subjects ANOVA, with the group condition (accompanying image vs. no accompanying image) as a between-subjects variable. The results showed no main effect for the encoding condition, $F(1, 89) = 1.57$, $MS_e = 223.93$, $p = .214$, $\eta_p^2 = .017$, no main effect for group, $F(1, 89) = 0.09$, $MS_e = 471.52$, $p = .761$, $\eta_p^2 = .001$, and no interaction between these two variables, $F(1, 89) < 0.01$, $MS_e = 223.93$, $p = .982$, $\eta_p^2 < .001$. The design in the current experiment is most similar to Experiment 1, because two JOLs were obtained from each participant. The encoding condition on JOLs had an $\eta_p^2 = .25$, producing an effect size of $f = .57$. The power of the current analysis to detect an effect size of that size is .99 ($n = 91$, $\alpha = .05$).

The current experiment investigated whether participants have a priori beliefs about the memorability of intact and generate images if they do not go through the experimental phase. Participants who did participate in previous experiments were presented with a scenario explaining the procedure and asked to make predictions about subsequent memory performance. The rationale was to see whether participants who had not carried out the experiment before had a priori beliefs that favored better memory performance for the intact, thus the more fluent, items. The results revealed no difference between encoding conditions, even when participants were shown examples of the materials during the presentation of the scenario. This finding is indicative that participants do not have uniform a priori beliefs about the memorability of intact and generate images. More possibly, the beliefs about intact and generate images are formed or are reinforced during exposure to the stimuli during the experiment. When participants have objective difficulties while processing the items, this might induce a belief that the generate items will be remembered less. The mediational analyses in both Experiments 2 and 3 showed that experience-based processes mediate the effect. The remaining effect that is observed for Experiment 2 is possibly induced *during* the experiment rather than before the experiment.

The online questionnaire in Experiment 4 is most comparable with Experiment 1, because in both experiments, only two JOLs (for intact and generate images) are obtained from each participant. Yet the two experiments produce two sets of different results. This difference between the results can most logically be explained through exposure to the experiment. In Experiment 4, participants

did not experience the objective perceptual difficulties because they were taking an online survey; on the other hand, the participants in Experiment 1 went through the experiment. This might indicate that going through the experience of participating in the experiment modifies the memory predictions for the participants, and provides support to the idea that the beliefs about the memorability of intact and generate pictures are formed during the experiment.

General Discussion

Predictions about our own memory performance rely on various cues. Some of these cues are valid and some of them are not (Koriat & Helstrup, 2007). A recent hypothesis about perceptual fluency argues that the perceptual characteristics of the items to be learned constitute a cue that impacts our memory predictions (Rhodes & Castel, 2008). According to this hypothesis, perceptually fluent items are predicted to be remembered better than perceptually disfluent items, even though perceptual fluency does not always affect actual memory performance, resulting in a metacognitive illusion. The experiments reported here aimed to assess the basic claims of the perceptual fluency hypothesis with an encoding manipulation that degraded the perception of images while increasing recall for the same material. The study had two main goals. The first goal was to investigate whether the perceptual fluency hypothesis can be generalized to perceptually richer pictorial material, because previous studies have only investigated the effect through the use of perceptually impoverished verbal material. The second goal of the study was to test the claims of the perceptual fluency hypothesis by investigating how experience-based and theory-based processes contribute to the effect.

These issues were examined with a picture generation manipulation. Half of the images were presented in their normal intact form (the fluent condition), whereas the other half were presented in a perceptually disfluent generate form. The results of the experiments are consistent with the perceptual fluency hypothesis. In all experiments in which participants went through the experimental procedure (Experiments 1 to 3), the intact images received higher JOLs than the generate images. Clearly, the intact images were judged to be more memorable than the less fluent generate items. These findings provide support for the generalization of the perceptual fluency hypothesis to pictorial material. This finding is important, because pictures and words are considered qualitatively different and may engage different types of encoding processes (Dewhurst & Conway, 1994; Nelson et al., 1977; Paivio, 1975, 1986), which in turn could impact metacognitive processes differently. Amit et al. (2009) argued that pictures are more similar to the objects that they represent and might be subject to the same perceptual processes that are applied to the objects. On the other hand, words do not perceptually look like their physical referents and may induce representations that are further away from the objects that they represent; thus, perceptual analysis of pictures might induce more specific and concrete representations than words. Even though words and pictures may be subject to different types of perceptual analyses during encoding, these differences are not reflected in the metamemory predictions. Just like verbal material, pictorial material also produces higher memory predictions for the more fluent, intact condition than the disfluent generate condition. The current findings validate that pictorial

materials that are closer to their referent objects than verbal materials are also analyzed in the same way. This confirms the *perceptual* characteristics of the perceptual fluency hypothesis in metamemory.

In addition to validation of the perceptual fluency hypothesis with perceptually enriched material, the current study provides objective measures of perceptual fluency. In Experiments 1 to 3, participants' perceptual fluency was assessed through two measures: first key press latency and total typing time. In all three experiments, identification latencies, as measured by first key press, were longer for generate images than intact items, providing support that it is in fact more difficult to perceive generate images. On the contrary, total typing time for intact and generate items was equivalent, showing that once participants identified the object shown in the image, the typing time did not make a difference—further proof that the effect is driven by initial perception of the image. To date, most research testing claims of perceptual fluency has not used objective measures of perception, except for a couple of studies (e.g., Besken & Mulligan, 2014; Mueller et al., 2014; Susser, Jin, & Mulligan, 2015), and no studies have investigated the identification latencies for the picture generation effect. Besken and Mulligan (2014) found that an auditory generation paradigm produced slower RTs for the more disfluent generate items than intact items, similar to the current study. Mueller et al. (2014), on the other hand, found that no difference between the identification of large- and small-font words in a lexical decision task. Both identification latencies and reportedly lower JOLs for generate than intact images show that perception of generate images is effortful at encoding and take longer than intact images.

The findings for recall performance are also in line with previous research using verbal materials. First of all, requiring aggregate JOLs at the end of the list produced higher performance for the perceptually difficult generate items compared with the easy-to-perceive, intact items, leading a double dissociation between memory predictions and actual recall performance, similar to verbal materials (Besken & Mulligan, 2013, 2014; Susser et al., 2013). Moreover, item-by-item JOLs eliminated the superior recall performance, as it usually does with perceptual generation manipulations (Begg et al., 1991; Besken & Mulligan, 2013, 2014; Matvey et al., 2001).

The second goal of the study was to assess the specific contributions of experience-based processes and theory-based processes to JOLs through the picture generation paradigm. As explained earlier, previous research indicates that JOLs are experience-based processes and theory-based processes (Koriat et al., 2004). Experience-based processes refer to the online subjective difficulties that participants go through while learning a specific item. On the other hand, theory-based processes refer to beliefs and heuristics about how memory operates beyond individuated experiential processes. The current design of the experiments reported here provides us with the opportunity to investigate the contribution of each process. The first question is whether experience-based processes contribute to JOLs. The experiments here provide evidence that they do. First, in Experiments 1 to 3, the identification latencies for generate items were always slower than for intact items, indicating some online objective difficulty. Moreover, correlational analyses in Experiment 2 showed that this objective difficulty in identification latencies was correlated with the reduction in JOLs. The partial correlation analyses in Experiments 2 and 3

also showed that when the effect of perceptual fluency, measured through identification latencies, was taken out, the correlation between perceptual fluency and JOLs decreased significantly, showing the independent contribution of experience-based processes. Last of all, the online survey in Experiment 4 revealed that without having the experience of carrying out the experiment, participants did not predict intact items to be more memorable than generate items, showing that participants do not have uniform a priori beliefs about perceptual fluency for this manipulation. Carrying out the experiment seems to be the only possible explanation for significantly different predictions across different encoding conditions, implying a role for experience-based processes.

The second question is whether theory-based processes influence JOLs. The answer to this question is also positive for two reasons. First, for the partial correlations in Experiment 2, in which the effect of experience-based processes (i.e., identification latencies) are taken out, the relationship between item type and JOLs is still significantly above zero, implying the presence of an effect beyond experience. This remaining effect is most probably theory-based processes (for similar explanations about theory-based processes, see Mueller et al., 2015; Undorf & Erdfelder, 2015). The second finding in favor of theory-based processes is that in Experiment 3, an additional variable (i.e., preceding contours that matched or mismatched the critical images) was included in the design. The matching preceding contours facilitated identification latencies for generate items more than the mismatching contour, yet this facilitation was not reflected in the JOLs, showing a dissociation between experience-based processes and JOLs. This type of dissociation between the two measures is indicative that identification latency, a variable that denotes experience-based processes, does not affect JOLs directly. This implies a role for the theory-based processes in the perceptual fluency hypothesis.

To sum up, both theory-based and experience-based processes are at work for producing metacognitive prediction differences between intact and generate images. Koriat's (1997) cue utilization provides an appropriate framework for interpreting the findings of the current set of experiments. In the cue-utilization framework, memory predictions are hypothesized to rely on three different types of cues: intrinsic, extrinsic,³ and mnemonic cues. Intrinsic cues refer to characteristics of the study items that disclose an item's a priori difficulty. In the current case, a priori beliefs about difficulty of intact and generate images may be classified as an intrinsic cue. Mnemonic cues, on the other hand, refer to cues that may signal to the participants the extent to which an item has been learned. In the current case, the perceptual fluency with which the participants process the material may be classified as a mnemonic cue. According to cue utilization framework, intrinsic cues can affect JOLs directly (e.g., beliefs about the difficulty of the intact and generate images may affect JOLs directly), or they can influence JOLs through their effects on mnemonic cues indirectly (e.g., the difficulty of intact and generate images is mediated by the ease of processing for these items). In the beginning of Experiment 3, two different options were mentioned as to how perceptual fluency might operate. The first explanation was that perceptual fluency directly affects JOLs. This explanation is coherent with a simple and direct interpretation of the perceptual fluency hypothesis (Rhodes & Castel, 2008), and corresponds to mnemonic cues in Koriat's cue utilization framework. An alternative interpretation could be that the experience of perceptual difficulty may affect a

participant's memory predictions by inducing a belief that the memory performance for a category of perceptually disfluent items will be lower in a subsequent test, making an independent contribution to this metamemorial illusion. This corresponds to intrinsic cues in Koriat's cue utilization framework. The evidence in Experiments 1 to 4 suggests that both of these factors seem to be at work. Participants rate objectively more disfluent generate items as less memorable, but this difference only comes into play when participants actually go through the process (Experiments 1 to 3), not when they have to make judgments about a scenario (Experiment 4), emphasizing the importance of experience-based processes. However, experience does not operate directly on JOLs. If participants cannot grasp the potential impact of a manipulation on perceptual fluency, the facilitated identification latencies do not affect JOLs (Experiment 3), implying that theory-based processes are also necessary to induce the perceptual fluency effect.

Even though the findings of the current experiment are in line with other generation manipulations (Besken & Mulligan, 2014), in which both experience-based and theory-based processes contribute independently to making JOLs, they are at odds with the findings from the font-size effect (Mueller et al., 2014). Both the font-size effect and the generation effect produce higher memory predictions for the "fluent" items than "disfluent" items. Yet these two manipulations have fundamental differences in terms of certain characteristics. First, the font-size manipulation does not produce objective perceptual difficulties, whereas the generation manipulation does. Second, participants have a priori beliefs about the memorability of large- and small-font words, whereas they do not have a priori beliefs about intact and generate images. These two fundamental differences presents us with a hint that the processes through which the generation effect impacts metacognitive judgments may be essentially different than the processes through which the font-size effect impacts metacognitive judgments. In the font-size effect, a priori beliefs play a predominant role in estimating the memory predictions for large- and small-font words and imply a role for the theory-based processes (Mueller et al., 2014). Yet in the generation effect, participants may first have to experience the objective difficulty and then develop beliefs about the memorability of the intact and generate images during the course of the experiment; only then will these beliefs be reflected in JOLs. These processes imply roles for both experience-based and theory-based processes in making JOLs. Thus, the underlying processes through which fluency manipulations affect JOLs may need to be considered separately for different types of manipulations.

The current project has two major contributions to literature. First, the current set of experiments show that the perceptual fluency effect for images works in a similar way as it does with verbal material. Second, the current set of experiments demonstrated that the contribution of experience-based and theory-based processes on JOLs for perceptual fluency is not linear, and per-

³ Koriat's cue utilization framework also explains extrinsic cues, cues that pertain to the conditions of encoding (e.g., number of times an item is presented, presentation time) or encoding operations applied to the material by the learner (e.g., levels of processing, interactive imagery). This is not mentioned in the discussion, as the conditions of learning were similar across encoding conditions, and the same instructions were given to all participants.

ceptual fluency manipulations that affect speed of perception do not always affect JOLs. By using a manipulation that is not very obvious to the participants, perceptual fluency, as measured by identification speed, was facilitated for the perceptually disfluent generate condition. If there were a direct relationship between objective measures of fluency and JOLs, this facilitation in identification should have also been reflected in the JOLs. Alternatively, if the effect is mediated by beliefs, the relationship between identification and JOLs should not be linear. The current findings show that the relationship is not linear, providing evidence for beliefs.

There might be two explanations as to how beliefs might moderate metacognitive judgments. One explanation is that identification latencies may impact JOLs solely by mediating beliefs about the manipulation. Alternatively, there might be a threshold at which participants can become aware of phenomenal experiences of objective measure of fluency in order to consider it in their JOLs. The first explanation is that identification latencies may impact JOLs solely by mediating beliefs about the manipulation. In the current case, because participants are not fully aware of the preceding contour manipulation and its potential effects on perceptual fluency, these potential effects do not factor into memory predictions. Thus, objective measures of fluency, such as identification latency, are necessary, but they are not sufficient to produce differences in JOLs for categorically different levels of fluency. Participants need to develop beliefs about the manipulation before or during the encoding process. In a very recent study, [Susser et al. \(2015\)](#) primed words with either matching or mismatching word primes backward masked by Xs. The results showed that when participants were not aware of the prime, perceptual fluency (as measured by naming latencies) for the words preceded by the matching prime increased, yet this difference was not reflected in JOLs; however, if the matching primes were presented long enough for participants to consciously process them, matching primes produced higher JOLs than mismatching primes. [Susser et al.](#) contended that most of the experience-based perceptual fluency effects can be accommodated by participants' beliefs about the effects of objective experience. Their finding, along with the finding from this study, provides evidence that participants can take their objective experience into account only when they become aware of it. Other research has also shown that perceptual fluency may not be taken into account during recognition ([Miller et al., 2008](#); [Westerman et al., 2003](#)). Thus, the theory-based processes to JOLs are essential for producing differences in perceptually fluent and perceptually disfluent material, whereas this might not always be the case with experience-based processes.

An alternative explanation for participants not to factor the identification latencies into JOLs for the preceding contour manipulation in Experiment 3 could be that there might be a threshold at which participants can subjectively become aware of an objective measure of fluency in order to consider it in their JOLs. For example, average identification differences between intact and generate images for Experiments 1, 2, and 3 (only for items preceded by nonmatching contours) have an average of 220 ms. The identification difference between generate items preceded by matching contours versus mismatching contours, on the other hand, is only 118 ms. Even though this difference is sufficiently long to produce significant identification fluency differences

through objective measures, it may not be sufficiently long to produce a subjective experience of fluency for the participants. In a similar line of work, [Reber, Wurtz, and Zimmermann \(2004\)](#) contended that the phenomenal experience of objective fluency consists of different types of objective measures of fluency and is factored into subjective experiences at different rates. For example, objective perceptual fluency can be operationally defined as the detection of a stimulus (e.g., is there anything on the screen, regardless of what it is). Alternatively, it can also be defined as the discrimination of the stimulus from other stimuli (e.g., "What is the name of the object presented on the screen?"). For instance, they found that figure-ground contrast of the items, a factor that impacts detection speed, affects judgments of fluency independently of font type, a factor that impacts discrimination speed. [Reber et al.](#) argued that both of these factors (detection fluency and discrimination fluency) contributed a feeling of perceptual fluency separately. In all experiments that investigate metamemory judgments for perceptual fluency objectively, perceptual fluency is defined as the identification or naming latency, both of which are operational measures of discrimination. No studies have used measures to observe detection speed for the perceptual fluency hypothesis. Yet showing matching subliminal primes ([Susser et al., 2015](#)) or matching preceding contours for a very brief period of time (Experiment 3) might affect detection speed independently of discrimination speed, and the detection speed might independently contribute to metacognitive predictions. Thus, further work needs to investigate the differences between different types of objective fluency for metamemory judgments.

Some of the previous findings of perceptual fluency might also be explained through the use of this "threshold" explanation for subjective realization of experience-based processes. For example, [Sungkhasettee, Friedman, and Castel \(2011\)](#) presented participants with upright and inverted words, followed by item-by-item JOLs. The results revealed memory superiority for inverted words with no difference in JOLs across the two encoding conditions. However, relevant literature shows that inverting words produces objective fluency differences, as measured by identification latencies (e.g., [Kolers, 1968](#); [Koriat & Norman, 1985](#); [Navon & Raveh, 2004](#)). Yet there might be a threshold at which participants might be able to become conscious of different objective measures of disfluency at a subjective level. In the inverted word manipulation, the material is always intact and the disfluency is produced by the rotation. Perhaps a manipulation that only produces different rotations of the same stimulus may not be able to create differences between fluent and disfluent conditions that reach the threshold for conscious realization of difficulties. Thus, the perceptual disfluency caused by rotation fails to mediate beliefs even when it affects experience-based processes, because the objective fluency difference between upright (fluent) and inverted (disfluent) stimuli is too short for participants to realize the difference. In other words, participants might have to reach a certain threshold of objective fluency difference before being able to consider the impacts of fluency subjectively. Obviously, more research is necessary to understand at which point objective measures of fluency might be subjectively perceived to create JOL differences between encoding conditions. Nevertheless, the literature now shows that the role of experience-based and theory-based processes on the perceptual fluency effect may be a little more complicated, and that it changes depending on the type of manipulation used.

References

- Ackerman, B. P. (1981). Encoding specificity in the recall of pictures and words in children and adults. *Journal of Experimental Child Psychology*, 31, 193–211. [http://dx.doi.org/10.1016/0022-0965\(81\)90012-6](http://dx.doi.org/10.1016/0022-0965(81)90012-6)
- Alter, A. L., & Oppenheimer, D. M. (2009). Uniting the tribes of fluency to form a metacognitive nation. *Personality and Social Psychology Review*, 13, 219–235. <http://dx.doi.org/10.1177/1088868309341564>
- Amit, E., Algom, D., & Trope, Y. (2009). Distance-dependent processing of pictures and words. *Journal of Experimental Psychology: General*, 138, 400–415. <http://dx.doi.org/10.1037/a0015835>
- Begg, I., Vinski, E., Frankovich, L., & Holgate, B. (1991). Generating makes words memorable, but so does effective reading. *Memory & Cognition*, 19, 487–497. <http://dx.doi.org/10.3758/BF03199571>
- 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. <http://dx.doi.org/10.1037/0096-3445.127.1.55>
- Besken, M., & Mulligan, N. W. (2013). Easily perceived, easily remembered? Perceptual interference produces a double dissociation between metamemory and memory performance. *Memory & Cognition*, 41, 897–903. <http://dx.doi.org/10.3758/s13421-013-0307-8>
- Besken, M., & Mulligan, N. W. (2014). Perceptual fluency, auditory generation, and metamemory: Analyzing the perceptual fluency hypothesis in the auditory modality. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40, 429–440. <http://dx.doi.org/10.1037/a0034407>
- Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology*, 64, 417–444. <http://dx.doi.org/10.1146/annurev-psych-113011-143823>
- Breitmeyer, B. G. (1984). *Visual masking*. New York, NY: Oxford University Press.
- Brodeur, M. B., Dionne-Dostie, E., Montreuil, T., & Lepage, M. (2010). The Bank of Standardized Stimuli (BOSS), a new set of 480 normative photos of objects to be used as visual stimuli in cognitive research. *PLoS ONE*, 5, e10773. <http://dx.doi.org/10.1371/journal.pone.0010773>
- Brown, J. M., & Koch, C. (1993). Influences of closure, occlusion, and size on the perception of fragmented pictures. *Perception & Psychophysics*, 53, 436–442. <http://dx.doi.org/10.3758/BF03206787>
- Carpenter, S. K., & Olson, K. M. (2012). Are pictures good for learning new vocabulary in a foreign language? Only if you think they are not. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38, 92–101. <http://dx.doi.org/10.1037/a0024828>
- Castel, A. D. (2007). The adaptive and strategic use of memory by older adults: Evaluative processing and value-directed remembering. *Psychology of Learning and Motivation*, 48, 225–270. [http://dx.doi.org/10.1016/S0079-7421\(07\)48006-9](http://dx.doi.org/10.1016/S0079-7421(07)48006-9)
- Castel, A. D., Benjamin, A. S., Craik, F. I. M., & Watkins, M. J. (2002). The effects of aging on selectivity and control in short-term recall. *Memory & Cognition*, 30, 1078–1085. <http://dx.doi.org/10.3758/BF03194325>
- Dewhurst, S. A., & Conway, M. A. (1994). Pictures, images, and recollective experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 1088–1098. <http://dx.doi.org/10.1037/0278-7393.20.5.1088>
- Dunn, J. C., & Kirsner, K. (1988). Discovering functionally independent mental processes: The principle of reversed association. *Psychological Review*, 95, 91–101. <http://dx.doi.org/10.1037/0033-295X.95.1.91>
- Foley, M. A., Foley, H. J., Durso, F. T., & Smith, N. K. (1997). Investigations of closure processes: What source-monitoring judgments suggest about what is “closing.” *Memory & Cognition*, 25, 140–155. <http://dx.doi.org/10.3758/BF03201108>
- Foley, M. A., Foley, H. J., Scheye, R., & Bonacci, A. M. (2007). Remembering more than meets the eye: A study of memory confusions about incomplete visual information. *Memory*, 15, 616–633. <http://dx.doi.org/10.1080/09658210701450919>
- Glisky, E. L., & Rabinowitz, J. C. (1985). Enhancing the generation effect through repetition of operations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11, 193–205. <http://dx.doi.org/10.1037/0278-7393.11.2.193>
- Hirshman, E., Trembath, D., & Mulligan, N. W. (1994). Theoretical implications of the mnemonic benefits of perceptual interference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 608–620. <http://dx.doi.org/10.1037/0278-7393.20.3.608>
- Intraub, H. (1984). Conceptual masking: The effects of subsequent visual events on memory for pictures. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10, 115–125. <http://dx.doi.org/10.1037/0278-7393.10.1.115>
- Kim, C.-Y., & Blake, R. (2005). Psychophysical magic: Rendering the visible “invisible.” *Trends in Cognitive Sciences*, 9, 381–388. <http://dx.doi.org/10.1016/j.tics.2005.06.012>
- Kimball, S., & Mattis, P. (2013). Gimp GNU Image Manipulation Program [Computer software]. Retrieved from <http://www.gimp.org>
- Kinjo, H., & Snodgrass, J. G. (2000). Does the generation effect occur for pictures? *The American Journal of Psychology*, 113, 95–121. <http://dx.doi.org/10.2307/1423462>
- Kolers, P. A. (1968). The recognition of geometrically transformed text. *Perception & Psychophysics*, 3, 57–64. <http://dx.doi.org/10.3758/BF03212713>
- 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. <http://dx.doi.org/10.1037/0096-3445.126.4.349>
- Koriat, A., Bjork, R. A., Sheffer, L., & Bar, S. K. (2004). Predicting one’s own forgetting: The role of experience-based and theory-based processes. *Journal of Experimental Psychology: General*, 133, 643–656. <http://dx.doi.org/10.1037/0096-3445.133.4.643>
- Koriat, A., & Helstrup, T. (2007). Metacognitive aspects of memory. In S. Magnussen & T. Helstrup (Eds.), *Everyday memory* (pp. 251–274). New York, NY: Psychology Press.
- Koriat, A., & Norman, J. (1985). Reading rotated words. *Journal of Experimental Psychology: Human Perception and Performance*, 11, 490–508. <http://dx.doi.org/10.1037/0096-1523.11.4.490>
- 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. <http://dx.doi.org/10.1177/0956797611407929>
- Martinovic, J., Gruber, T., & Müller, M. M. (2008). Coding of visual object features and feature conjunctions in the human brain. *PLoS ONE*, 3, e3781. <http://dx.doi.org/10.1371/journal.pone.0003781>
- Matvey, G., Dunlosky, J., & Guttentag, R. (2001). Fluency of retrieval at study affects judgments of learning (JOLs): An analytic or nonanalytic basis for JOLs? *Memory & Cognition*, 29, 222–233. <http://dx.doi.org/10.3758/BF03194916>
- 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. <http://dx.doi.org/10.1037/a0025548>
- Miller, J. K., Lloyd, M. E., & Westerman, D. L. (2008). When does modality matter? Perceptual versus conceptual fluency-based illusions in recognition memory. *Journal of Memory and Language*, 58, 1080–1094. <http://dx.doi.org/10.1016/j.jml.2007.12.006>
- Mintzer, M. Z., & Snodgrass, J. G. (1999). The picture superiority effect: Support for the distinctiveness model. *The American Journal of Psychology*, 112, 113–146. <http://dx.doi.org/10.2307/1423627>
- Moreno-Martínez, F. J., & Montoro, P. R. (2012). An ecological alternative to Snodgrass & Vanderwart: 360 high quality colour images with norms

- for seven psycholinguistic variables. *PLoS ONE*, 7, e37527. <http://dx.doi.org/10.1371/journal.pone.0037527>
- Mueller, M. L., Dunlosky, J., & Tauber, S. K. (2015). The effect of identical word pairs on people's metamemory judgments: What are the contributions of processing fluency and beliefs about memory? *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*. <http://dx.doi.org/10.1080/17470218.2015.1058404>
- 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. <http://dx.doi.org/10.1016/j.jml.2013.09.007>
- Mueller, M. L., Tauber, S. K., & Dunlosky, J. (2013). Contributions of beliefs and processing fluency to the effect of relatedness on judgments of learning. *Psychonomic Bulletin & Review*, 20, 378–384. <http://dx.doi.org/10.3758/s13423-012-0343-6>
- Nairne, J. S., Riegler, G. L., & Serra, M. (1991). Dissociative effects of generation on item and order retention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17, 702–709. <http://dx.doi.org/10.1037/0278-7393.17.4.702>
- Navon, D., & Raveh, O. (2004). On the process of recognizing inverted words: Does it rely only on orientation-invariant cues? *Memory & Cognition*, 32, 1103–1117. <http://dx.doi.org/10.3758/BF03196885>
- Nelson, D. L., Reed, V. S., & McEvoy, C. L. (1977). Learning to order pictures and words: A model of sensory and semantic encoding. *Journal of Experimental Psychology: Human Learning and Memory*, 3, 485–497. <http://dx.doi.org/10.1037/0278-7393.3.5.485>
- Nelson, D. L., Reed, V. S., & Walling, J. R. (1976). Pictorial superiority effect. *Journal of Experimental Psychology: Human Learning and Memory*, 2, 523–528. <http://dx.doi.org/10.1037/0278-7393.2.5.523>
- Paivio, A. (1975). Perceptual comparisons through the mind's eye. *Memory & Cognition*, 3, 635–647. <http://dx.doi.org/10.3758/BF03198229>
- Paivio, A. (1986). *Mental representation: A dual coding approach*. Hillsdale, NJ: Erlbaum.
- Reber, R., Winkelman, P., & Schwarz, N. (1998). Effects of perceptual fluency on affective judgments. *Psychological Science*, 9, 45–48. <http://dx.doi.org/10.1111/1467-9280.00008>
- Reber, R., Wurtz, P., & Zimmermann, T. D. (2004). Exploring “fringe” consciousness: The subjective experience of perceptual fluency and its objective bases. *Consciousness and Cognition: An International Journal*, 13, 47–60. [http://dx.doi.org/10.1016/S1053-8100\(03\)00049-7](http://dx.doi.org/10.1016/S1053-8100(03)00049-7)
- 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. <http://dx.doi.org/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. <http://dx.doi.org/10.3758/PBR.16.3.550>
- Serra, M. J., & Dunlosky, J. (2005). Does retrieval fluency contribute to the underconfidence-with-practice effect? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31, 1258–1266. <http://dx.doi.org/10.1037/0278-7393.31.6.1258>
- 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. <http://dx.doi.org/10.3758/s13423-011-0114-9>
- Susser, J. A., Jin, A., & Mulligan, N. W. (2015). Identity priming consistently affects perceptual fluency but only affects metamemory when primes are obvious. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Advance online publication. <http://dx.doi.org/10.1037/xlm0000189>
- Susser, J. A., & Mulligan, N. W. (2015). The effect of motoric fluency on metamemory. *Psychonomic Bulletin & Review*, 22, 1014–1019. <http://dx.doi.org/10.3758/s13423-014-0768-1>
- Susser, J. A., Mulligan, N. W., & Besken, M. (2013). The effects of list composition and perceptual fluency on judgments of learning (JOLs). *Memory & Cognition*, 41, 1000–1011. <http://dx.doi.org/10.3758/s13421-013-0323-8>
- Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, 381, 520–522. <http://dx.doi.org/10.1038/381520a0>
- Undorf, M., & Erdfelder, E. (2015). The relatedness effect on judgments of learning: A closer look at the contribution of processing fluency. *Memory & Cognition*, 43, 647–658. <http://dx.doi.org/10.3758/s13421-014-0479-x>
- Vokey, J. R., Baker, J. G., Hayman, G., & Jacoby, L. L. (1986). Perceptual identification of visually degraded stimuli. *Behavior Research Methods, Instruments & Computers*, 18, 1–9. <http://dx.doi.org/10.3758/BF03200985>
- Westerman, D. L., Miller, J. K., & Lloyd, M. E. (2003). Change in perceptual form attenuates the use of the fluency heuristic in recognition. *Memory & Cognition*, 31, 619–629. <http://dx.doi.org/10.3758/BF03196102>
- 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. <http://dx.doi.org/10.3758/s13421-012-0255-8>

Received April 23, 2015

Revision received November 7, 2015

Accepted December 17, 2015 ■