Psychophysiology, 52 (2015), 754–769. Wiley Periodicals, Inc. Printed in the USA.

DOI: 10.1111/psyp.12412

The pupil as an indicator of unconscious memory: Introducing the pupil priming effect

CARLOS ALEXANDRE GOMES, DANIELA MONTALDI, AND ANDREW MAYES
Human Memory Laboratory, School of Psychological Sciences, University of Manchester, Manchester, UK

Abstract

We explored whether object behavioral priming and pupil changes occur in the absence of recognition memory. Experiment 1 found behavioral priming for unrecognized objects (Ms) regardless of whether they had been encoded perceptually or conceptually. Using the same perceptual encoding task, Experiment 2 showed greater pupil dilation for Ms than for correct rejections of unstudied objects (CRs) when reaction times were matched. In Experiment 3, there was relatively less pupil dilation for Ms than for similarly matched CRs when objects had been encoded conceptually. Mean/peak pupil dilation for CRs, but not Ms, increased in Experiment 3, in which novelty expectation was also reduced, and the pupillary time course for both Ms and CRs was distinct in the two experiments. These findings indicate that both behavioral and pupil memory occur for studied, but unrecognized stimuli, and suggest that encoding and novelty expectation modulate pupillary memory responses.

Descriptors: Priming, Novelty, Recognition memory, Pupil dilation, Implicit memory, Pupillometry

Learning, memory, and attention have mainly been investigated through the use of behavioral measures, introspective reports, and, more recently, by direct and noninvasive imaging of brain activity. However, the human eye can, in some circumstances, also be an accurate indicator of cognitive-specific processes. Indeed, over the past few decades, eye-tracking methodology (i.e., the measurement of eye movements and/or eye fixation patterns) and pupillometry (i.e., the measurement of pupil diameter) have become established and important markers of cognitive function (Laeng, Sirois, & Gredebäck, 2012; Richardson, Dale, & Spivey, 2007).

Eye-tracking methodology has been extensively used to investigate cognitive processes, including learning and memory (see Hannula et al., 2010; Rayner, 2009; Tanenhaus, Magnuson, Dahan, & Chambers, 2000, for reviews on these topics). Eye movements at encoding have been found to predict later memory (e.g., Kafkas & Montaldi, 2011) while, at retrieval, distinct eye-movement patterns for previously encountered items have been observed (e.g., Chanon & Hopfinger, 2008; Kafkas & Montaldi, 2012; Ryan, Althoff, Whitlow, & Cohen, 2000), even when explicit awareness of such pupillary behavior is ruled out (e.g., Hollingworth & Henderson, 2002; Hollingworth, Williams, & Henderson, 2001; Ryan et al., 2000; Theeuwes, Kramer, Hahn, & Irwin, 1998; Theeuwes, Kramer, Hahn, Irwin, & Zelinsky, 1999).

Pupillometry studies, although fewer in number, have also been conducted in order to identify correlates of cognitive processes involving changes in pupil diameter, such as pupil dilation. Some studies have revealed changes in pupil diameter, particularly under conditions of high cognitive load (e.g., Granholm, Asarnow, Sarkin, & Dykes, 1996; Just & Carpenter, 1993; Kahneman & Beatty, 1966; Klingner, Tversky, & Hanrahan, 2011).

Researchers have also used pupillometry to investigate memory in an attempt to identify pupil changes related to memory functions. Early studies conducted by Gardner and colleagues (e.g., Gardner, Beltramo, & Krinsky, 1975; Gardner, Mo, & Borrego, 1974; Gardner, Philp, & Radacy, 1978) showed pupillary response changes related to memory encoding and retrieval.

More recent studies have confirmed the link between pupil-related changes and memory. For instance, Kafkas and Montaldi (2011) found that, when encoding was incidental, pupil dilation decreased as a function of subsequent memory strength. Papesh, Goldinger, and Hout (2012), however, found increased pupil dilation for later successful retrieval, but their encoding task, unlike that of Kafkas and Montaldi's (2011), was a directed one, so their effects are likely to reflect encoding effort rather than an underlying memory-related autonomic response. Using the same object stimuli as in their earlier study, Kafkas and Montaldi (2012) later found that, at retrieval, the opposite relationship was apparent—with greater pupil dilation for old compared to new, and for recollected compared to familiar stimuli, even when reaction times (RTs) and presumably effort was matched. Otero, Weekes, and

---

This work was funded by the Fundação para a Ciência e a Tecnologia (SFRH/BD/41637/2007) and the University of Manchester. CAG is now at the Department of Psychology, Saarland University, Germany.

Address correspondence to: Carlos A. Gomes, Faculty of Behavioral Sciences, Department of Psychology, Saarland University, Building A 2.4, 2nd Floor, D-66123 Saarbrücken, Germany.
E-mail: carlosalexandre.gomes@uni-saarland.de

Hutton (2011), using the traditional remember/know (R/K) procedure (Tulving, 1985), also found this pupil old/new effect and less pupil dilation for familiar relative to recollected words. Kafkas and Montaldi, however, having used a modified R/K paradigm (Mayes, Montaldi, & Migo, 2007; Montaldi, Spencer, Roberts, & Mayes, 2006) and RT matching, were able to conclude that these retrieval-related recollection/familiarity pupil effects can occur independent of memory strength and, relatedly, effort. Finally, Võ et al. (2008) also illustrated a pupil old/new effect in which emotionally weighted, correctly recognized words produced more pupil dilation than correctly novel, unstudied words, although the effect found in this study was shown to be driven by the emotional valence of the stimuli.

In contrast, and perhaps surprisingly, few studies have reported pupil-related changes of long-term priming, a type of implicit memory characterized by improved processing of studied material that can occur in the absence of conscious memory (see Roediger, 1990; Tulving & Schacter, 1990, for reviews). To our knowledge, only one study (Laeng et al., 2007) has attempted to investigate long-term priming using pupillometric measures. In this study, three patients with amnesia were shown a collection of pictures of objects from different categories (e.g., animals, transports) at encoding and were required to associate the pictures with spoken semantic information. At test, the three patients were asked to provide yes/no recognition responses to each picture presented one at a time on a computer screen. Although all three patients were impaired at visual recognition memory, their pupils exhibited increased dilation for new relative to old items. Importantly, this pattern was evident even for one of the patients who responded "new" to all pictures during the recognition memory test, which suggested that explicit recognition memory was absent, but implicit recognition as measured by pupillometry was intact.

Confidence in Laeng et al.'s (2007) interesting findings is somewhat reduced by several concerns. First, the patient sample included in this study was very small (three patients). Second, the absence of a control group in their study makes it impossible to ascertain whether this phenomenon would generalize to individuals without memory disorders and, relatedly, whether unaware pupil memory was completely normal in patients. Third, pupil priming was reported as a raw score, although it is more common and reliable to report it as a deviation from a baseline. Fourth, priming was calculated as the difference between studied and unstudied pictures, regardless of whether the response was correct. In addition, of the two patients who gave new and old responses, one of them responded equally fast for old and new, whereas, for the other patient, RTs were faster for new than for old, which may suggest residual recognition memory for the old stimuli.

Given the ongoing debate about whether the memory sources driving long-term priming and recognition memory should be viewed as independent (e.g., Voss & Gonsalves, 2010; Voss, Reber, Mesulam, Parrish, & Paller, 2008) or not (e.g., Berry, Shanks, Speekenbrink, & Henson, 2012), the possibility that pupillometry can help resolve this issue merits serious consideration. For this reason, we decided to conduct two experiments (Experiment 2 and 3) to determine whether pupil changes mark object memory when there is no recognition of the studied material.

In Experiment 1, participants performed both a perceptual (dot-counting judgment) and a conceptual (animacy judgment) encoding task on different objects, whereas at test they engaged in an object decision task followed by a recognition memory test on a trial-by-trial basis. This experiment was conducted in order to ascertain whether reliable behavioral priming could be produced for unrecognized but studied stimuli. In Experiment 2, participants performed the same perceptual encoding task as in Experiment 1. At test, we administered a recognition memory task without the object decision priming task while recording pupil data simultaneously. Pupil priming[1] was defined as the difference in pupil dilation between misses (i.e., studied objects incorrectly judged "new") and correct rejections (i.e., novel objects correctly judged "new") with matching RTs. Experiment 3 was similar to Experiment 2 with the difference that a conceptual task (size classification) was used at encoding to determine whether pupil priming would be different for conceptually encoded objects than for perceptually encoded ones.

To anticipate our key findings, behavioral priming (as measured by RTs) for unrecognized objects occurred following both perceptual and conceptual encoding (Experiment 1). Pupil priming was also identified in Experiment 2 as an increase in mean/peak pupil dilation for misses (Ms) relative to RT-matched correct rejections (CRs), whereas this effect was reversed in Experiment 3. In addition, the pupillary time course differed between matched Ms and CRs as well as between experiments. Possible interpretations for these findings will be considered in the General Discussion.

Experiment 1

We initially conducted a behavioral experiment using a typical behavioral priming measure (i.e., RTs) in order to investigate whether behavioral priming could be obtained for unrecognized but studied stimuli using a novel behavioral task. In this experiment, participants performed, at study, both a shallow- and a deep-encoding task on pictures of common objects. At test, they judged whether or not each object was real and, with real objects, decided whether they had been presented during the study phase. We were particularly interested in determining whether behavioral priming occurred when studied objects were not recognized. If this was found, and participants showed different pupil sizes for Ms and CRs under very similar circumstances, it would strongly suggest that pupil priming and behavioral priming depend on common mechanisms.

All experiments reported in this article were approved by the School of Psychological Sciences Research Ethics Committee of the University of Manchester.

Method

Participants. Thirty undergraduate students at the University of Manchester were recruited. All participants had normal or corrected-to-normal vision and gave informed written consent to take part in this study. One participant was excluded from the statistical analysis due to answering correctly only once to nonobjects.

Materials. Stimuli were selected from an Internet clip art database (www.clipart.com) and consisted of 60 black and white pictures of real objects, half of which comprised animate objects (e.g., animals) whereas the other half represented inanimate objects (e.g.,

---

[1] In Experiment 1, priming was indicated directly as a kind of memory through behavior in the form of an RT speedup. Theoretically, neuroimaging methods can reveal the neural underpinnings of this kind of stimulus-specific unconscious memory. Pupil changes are neither the defining behavioral indicators of priming nor its neural underpinnings, but, presumably, are triggered specifically by the brain regions that underpin this kind of memory. So, pupil changes indirectly mark the occurrence of priming, and for brevity we refer to this as "pupil priming," although the pupillary response is not itself primed.

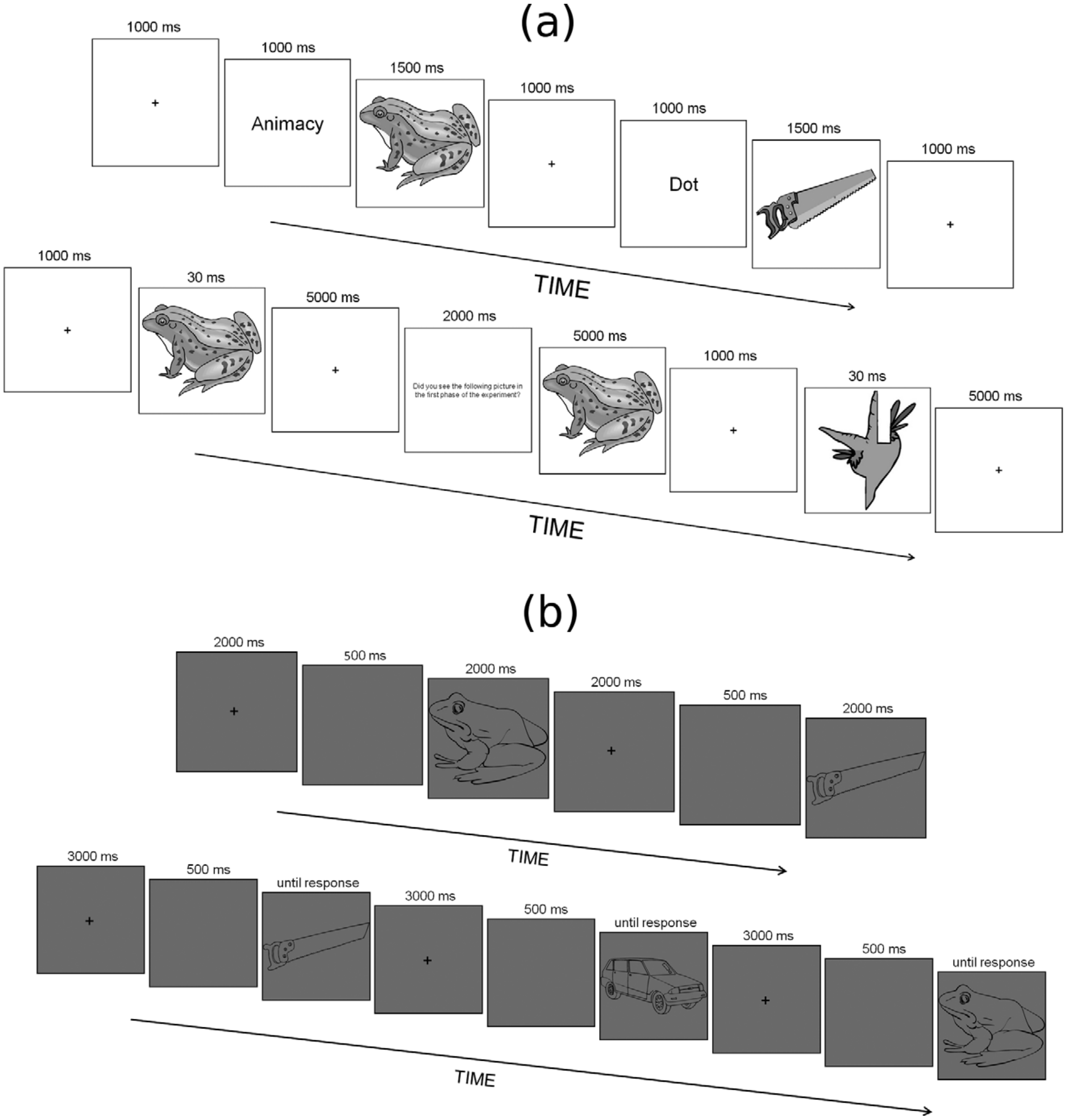

Figure 1. a: Experimental paradigm of Experiment 1. At study (top), participants either counted the number of red dots presented within the object picture (dot task) or judged whether an object was either animate or inanimate (animacy task). At test (bottom), they performed an object decision task followed by a recognition memory test. b: Experimental paradigm of Experiment 2. At study (top), participants counted the number of red dots presented along the lines of the pictures. At test (bottom), they performed a recognition memory test. The numbers above each event correspond to the duration of those events in milliseconds.

vehicles). In addition, a different set of 30 pictures depicting different objects were altered to create the nonobjects used in the test task (see Figure 1a for an example; see also Gomes & Mayes, 2014, for information on how these nonobjects were created).

Procedure. Figure 1a shows the experimental design used in this experiment. Participants engaged in two different encoding tasks that were randomly intermixed: In the dot task, participants indicated whether they had seen either one or two red dots flashed within the picture, whereas in the animacy task, they indicated whether a picture depicted an animate or inanimate object. The encoding task was randomly selected on each trial, and a cue word was always shown before the presentation of a picture to inform participants which task they were about to engage in. There were a total of 40 pictures; half of these were randomly assigned to the animacy task and the other half to the dot task, with the constraint that for each task half of the pictures depicted animate objects whereas the other half inanimate objects. Each study trial began with a fixation cross shown for 1,000 ms followed by the cue word, also presented for 1,000 ms.

A picture of an object was presented immediately after the cue for 1,500 ms. If the task corresponded to the dot task, each dot briefly flashed for 500 ms at a random location within the image, and when two dots were presented, they were shown sequentially with an interval of approximately 200 ms between them. Participants pressed one of the control keys on the keyboard for one-dot/animate decisions, and the other control key for two-dots/inanimate decisions.

For the test phase, 20 new pictures of objects as well as an additional 30 pictures of nonobjects were shown along with the 40 studied objects (half of them from the dot task and the other half from the animacy task). Participants were instructed to, first, decide as quickly as possible whether a picture represented an object (object-decision task), and, second, except when the stimulus was a nonobject, whether it had been presented at study (recognition task). Each test trial started with a fixation cross for 1,000 ms followed by a brief presentation of either an object or a nonobject for 30 ms. Participants attempted to respond within 5,000 ms and, for object stimuli, the question, “Did you see the following picture during the first phase of the experiment?” appeared for 2,000 ms. The same picture was then presented for up to 5,000 ms, within which time participants had to respond. They pressed one of the control keys for “real object” and “old object” decisions, and the other control key for “nonobject” and “new object” decisions. This and subsequent experiments were programmed and run using the MATLAB (http://www.mathworks.com) toolbox Cogent (http://www.vislab.ucl.ac.uk/cogent.php).

Design and analysis. The experimental design comprised category (hits, misses, correct rejections) as a within-subject factor. The data were analyzed using repeated measures analysis of variance (ANOVA) and paired *t* tests on RTs. A Huynh-Feldt correction was applied to the degrees of freedom of those tests for which the assumption of sphericity was violated. The alpha level was set for all statistical tests at .05, and *t* tests were two tailed.

Results

Incorrect trials, trials with absent responses, and/or with an RT more than 2.5 standard deviations above or below the mean were removed from subsequent analyses (13.5% of trials for all conditions and participants).

A single score representing the “corrected” hit rate (PR) for each condition and for each participant was calculated by subtracting the false alarm rate from the hit rate of each study condition. A paired *t* test between $\mathrm{PR}_{\mathrm{deep}}$ and $\mathrm{PR}_{\mathrm{shallow}}$ was significant, $t(28) = 6.671$, $p < .001$, $d = 2.52$, indicating better explicit memory for deeply encoded objects (0.59) than shallowly encoded objects (0.34). Notwithstanding, both $\mathrm{PR}_{\mathrm{deep}}$, $t(28) = 19.278$, $p < .001$, $d = 3.47$, and $\mathrm{PR}_{\mathrm{shallow}}$, $t(28) = 10.823$, $p < .001$, $d = 1.97$, were significantly above chance.

In order to ascertain whether priming could be observed without explicit memory, we separated object-decision trials into three categories according to subsequent recognition: a hits (Hs) category, a misses (Ms) category, and a correct rejections (CRs) category. Only participants with more than four trials per cell across all categories were included in the analysis (four participants excluded).

Figure 2 shows the RTs for Hs and Ms, separate for each encoding condition, as well as for CRs. For Ms, a single-factor repeated measures ANOVA (category: $\mathrm{Ms}_{\mathrm{deep}}$, $\mathrm{Ms}_{\mathrm{shallow}}$, CRs) yielded a significant effect, $F(2,30) = 6.550$, $p < .01$, $\eta_{\mathrm{p}}^2 = .30$. Paired *t* tests revealed that both $\mathrm{Ms}_{\mathrm{deep}}$ and $\mathrm{Ms}_{\mathrm{shallow}}$ objects were judged significantly faster than CRs ($\mathrm{Ms}_{\mathrm{deep}}$ vs. CRs: $t(15) = -3.197$, $p < .01$, $d = .80$; $\mathrm{Ms}_{\mathrm{shallow}}$ vs. CRs: $t(28) = -2.964$, $p < .01$, $d = .55$). There was no difference between $\mathrm{Ms}_{\mathrm{shallow}}$ and $\mathrm{Ms}_{\mathrm{deep}}$, $t(15) = -1.556$, $p > .10$, $d = .39$. The ANOVA on the Hs data did not approach significance, $F(1.661,46.510) = 2.141$, $p > .10$, $\eta_{\mathrm{p}}^2 = .07$. However, a paired *t* test between $\mathrm{Hs}_{\mathrm{shallow}}$ and CRs revealed a trend towards significance, $t(28) = -1.868$, $p = .07$, $d = .35$.

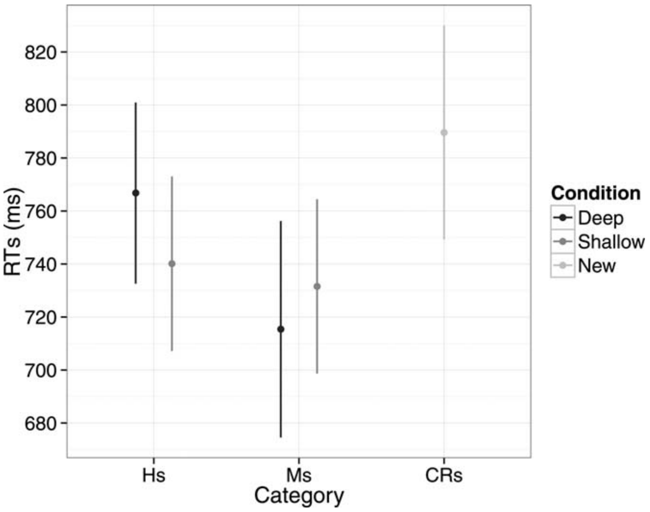

Figure 2. Mean reaction times (RTs) for Experiment 1. $\mathrm{Hs}_{\mathrm{deep}}$ = hits for objects studied under deep-encoding conditions; $\mathrm{Hs}_{\mathrm{shallow}}$ = hits for objects studied under shallow-encoding conditions; $\mathrm{Ms}_{\mathrm{deep}}$ = misses for objects studied under deep-encoding conditions; $\mathrm{Ms}_{\mathrm{shallow}}$ = misses for objects studied under shallow-encoding conditions; CRs = correct rejections. Error bars represent the standard error of the mean.

Discussion

The important finding of the present experiment was the detection of a reliable priming effect: faster RTs for Ms relative to CRs. This result provides some reassurance that any detectable difference in pupil dilation between Ms and CRs should arise from a very similar kind of memory trace to the one supporting the behavioral effect obtained in the present experiment. However, there was no reliable priming for Hs in either condition (although priming for $\mathrm{Hs}_{\mathrm{shallow}}$ approached significance). One possible explanation for this result could be that interference occurred during the priming task when participants were unable to stop themselves anticipating the recognition memory decision that was to follow. This interference was presumably greater when the objects were involuntarily recognized, which would have occurred more with stronger memories. Such active, explicit memories would have distracted attention from the priming task even more, and thus caused a bigger slowdown in the primed responses so as to minimize any speedup. In other words, during the object-decision task, participants could not stop themselves from attempting to retrieve whether a particular object had been presented earlier, and, if a conscious memory had been produced, the competing priming response would have been slowed down.

General Pupillometry Method

The pupillometry experiments reported in this paper follow a typical study-test procedure. At study, participants saw pictures of objects, while the kind of encoding was manipulated across experiments (see Method section of each experiment). According to dual-process models of recognition memory (e.g., Yonelinas, 2002), recollection occurs when a recognition test stimulus cues recall of aspects of its encoding context, whereas familiarity occurs when it feels that a recognition test stimulus has been encountered before, without any cued recall of specific associations from its encoding context. In the following two pupillometry experiments, participants were presented, at test, with a modified R/K recognition memory test used to measure the two kinds of memory that support recognition memory. In this modified familiarity-only procedure (Montaldi et al., 2006), participants are explicitly instructed to report quickly and accurately whether familiarity is present, and not to try to recollect specific contextual details about the stimulus, but to report if they occur involuntarily.

Accuracy (i.e., proportion of correct responses), RTs, and pupil data were recorded. Critically, six response categories (recollection hits, familiarity hits (Fs), misses (Ms), correct rejections (CRs), recollection false alarms, and familiarity false alarms) were extracted for offline analysis. However, given our specific interest in priming, the analyses described in this paper focused mainly on differences between Ms and CRs, although Fs were also included when found relevant. Familiarity recognition success rate (PR) was calculated assuming stochastic independence using the formula [familiarity hit rate/(1 – recollection hit rate)] – [familiarity false alarm rate/(1 – recollection false alarm rate)] (Yonelinas, 2002).

It should be noted that a pupillary response difference between Ms and CRs is not an entirely satisfactory demonstration of unconscious memory because Ms are usually found to be slower on average than CRs (as in our experiments). The slowness of Ms relative to CRs is almost certainly related to the greater difficulty in making M decisions so more effort has to be exerted, and it is well known that increased effort is a major driver of pupil dilation (e.g., Beatty & Kahneman, 1966; Kahneman & Beatty, 1966). In the recognition memory context, such increased effort and decision time for Ms rather than CRs could result from two sources. First, Ms could have needed more effort to process than CRs in nonmemory-related ways. Although selection of studied and unstudied objects was based on randomization, Ms and CRs were selected by participants, so it is conceivable, even if unlikely, that their selections were biased in terms of the relative processing difficulty of M and CR pictures. To eliminate this possibility, we used a procedure to confirm that M and CR pictures did not differ in this way (see below). Second, the greater time and effort for Ms was almost certainly related to above-chance levels of familiarity, whereas the familiarity of CRs must have been at chance. This is because familiarity strength is a continuous variable (Yonelinas, 1994), and the closer familiarity strength is to the recognition threshold, the more difficult and effortful it will be to reach F and M decisions, slowing these responses (Kafkas & Montaldi, 2014). Therefore, a straightforward comparison of Ms with CRs is confounded by the greater familiarity and, relatedly, the greater effort needed for M decisions. Consistent with this, we found that RTs were slower for Ms than for CRs, and that pupil dilation was greater for slower than for faster Ms.

To help ensure that any pupillary difference between Ms and CRs reflects completely unconscious object memory (i.e., priming) unconfounded by greater levels of above chance, but just below threshold levels of familiarity, and, relatedly, higher levels of M-related effort, Ms and CRs were RT matched. An overall RT confound matching procedure was used in both pupillometry experiments. This was developed by Spencer, Montaldi, Gong, Roberts, and Mayes (2009) as a way to control for regression towards the mean. The procedure involves the selection of pairs of individual Ms and CRs that have identical RTs. However, given that it is unlikely that a sufficient number of trials are identical at the millisecond level, we allowed some variation to occur between any two given trials, with the constraint that this variation did not exceed 100 ms on either side.[2]. The number of positive differences (i.e., Ms RT < CRs RT) was balanced with the number of negative differences (i.e., Ms RT > CRs RT) in an unbiased way, starting with the smallest possible difference between the two categories and proceeding by always alternating one negative difference with a positive one. Thus, after matching, both response categories should have identical RT distributions with equivalent means and standard deviations.

Variations in RTs for Ms are a joint function of weak familiarity-related effort and of nonmemory-related effort. In contrast, RT variations for CRs depend only on the nonmemory-related effort because CR-related familiarity levels must be at chance. Both these sources of effort need to be matched between Ms and CRs to ensure that relative pupil dilation for Ms indicates an effect of completely unconscious object memory (priming) rather than a confound due to residual familiarity. Without this double match, the matching RTs for Ms and CRs could have resulted from quicker/slower nonmemory-related effortful processing counterbalanced by slower/quicker memory-related effortful processing. With an overall RT match, if nonmemory-related effort for Ms was reduced, then any greater pupil dilation for Ms would have related to residual familiarity and greater effort. A nonmemory RT confound matching procedure was used to match Ms and CRs in the following way: we computed the mean RT for each stimulus when it featured as a CR, because, as familiarity was at chance for CRs, RTs could only have been influenced by variations in nonmemory-related factors. Then, for each participant, the M and CR RTs for each stimulus were replaced by the corresponding mean CR RT for that stimulus. For each participant, a mean CR-based M and CR score was calculated, which enabled us to check whether these RTs were matched. Such a match would indicate that nonmemory-related difficulty of the pictures comprising participants' Ms and CRs was closely similar and that, therefore, familiarity-related effort was also matched and M familiarity was very close to chance levels. It should be noted that the validity of this inference depends on there being a close relationship between RT and both picture nonmnemonic processing and how close subthresholded familiarity is to the recognition decision threshold. This, of course, depends on the correctness of our reasoning, which some may challenge.

To ensure the good quality of pupil data, only participants whose pupil discrimination allowed for a consistent recording of pupil data throughout the experimental session were included in the analyses. Also, participants who demonstrated exaggerated movement during the experiment were excluded from the analyses.

In keeping with other pupillometry studies (e.g., Kafkas & Montaldi, 2011, 2012), we also performed within-participant cleaning procedures. Artifacts such as blinks were removed from each pupillary trace, and trials with an excessive number of these artifacts (> 30%) were excluded from statistical analyses. Linear

---

[2] We identified the cutoff point of 100 ms as a reasonable compromise between closely matched RTs and number of Ms available, as shorter time windows would have greatly reduced the number of potential Ms that could be matched. In any case, choosing different time windows did not substantially affect the main effects reported.

interpolation replaced data points for a particular trace that deviated from the mean for that trace by more than ± 2 standard deviations. An unweighted 5-point moving average filter was then passed over the data to minimize the partial lid closure effect (i.e., to account for the constriction that occurs when the lid closes).

Finally, Goldinger and Papesh (2012) have recently delineated a few experimental considerations that could improve the validity of pupillometry as a means to study memory. In short, they suggested using (a) colorless, low-contrast stimuli, (b) relatively long stimulus exposure, (c) a sufficiently long interstimulus interval, (d) baseline-corrected diameter analyses, and (e) different but complementary pupillometric measures (e.g., mean and peak pupil diameter). Thus, in order to increase the reliability of our interpretations, we followed these recommendations in the present study.

Experiment 2

In Experiment 1, reliable behavioral priming (as measured by RTs) was observed for shallowly encoded and unrecognized objects. The present experiment was conducted in order to determine whether this effect could also be measured under very similar conditions but with pupillometry. More specifically, we sought to investigate whether Ms objects, for which recognition is likely to be partially or completely absent, still show a pupillary response that differs from that of CRs. To ensure that conditions were as similar as possible to the behavioral experiment, participants engaged in the same low-level study task as in Experiment 1.[3] During the test phase, a recognition memory test was administered using familiarity-only instructions (Mayes et al., 2007; Montaldi et al., 2006).

We purposefully chose a recognition task rather than the behavioral priming task during the test phase because it is likely that explicit memory contamination occurs for some trials during long-term priming paradigms including ours. Without having a measure of subjective recognition memory, any pupillary response difference between studied (primed) and unstudied (unprimed) stimuli could simply reflect recognition of some stimuli, which could inflate pupil dilation for the primed stimuli. This would be a serious confound in our experiment that we wished to avoid. Also, using exactly the same paradigm as in Experiment 1 (i.e., a behavioral priming test followed by a recognition memory test) would be problematic because a form of interference would be present during priming trials as a result of participants anticipating recognition memory judgments. An item-by-item priming-then-recognition test would seriously undermine the possibility of disentangling whether the pupillary response was due to true priming or just anticipation effects because the pupil is sensitive to anticipation (e.g., Tressoldi, Martinelli, Semenzato, & Cappato, 2011). Thus, we opted to administer an object recognition memory test on its own during the test phase on the assumption that reliable behavioral priming would automatically occur for unrecognized objects under similar encoding conditions as Experiment 1 (but see footnote 3).

Method

Participants. Thirty undergraduate students from the University of Manchester were recruited (*M* = 20.37, *SD* = 1.78). All participants had normal or corrected-to-normal vision and gave informed written consent to take part in this study. Two participants were excluded due to difficulties with discrimination of the pupil, and an additional participant was excluded due to exaggerated movement throughout the experiment.

Materials. Two hundred line-drawing images of common objects, and an additional 10 for the practice trials, were selected from an Internet database (http://www.clipart.com/). In order to ensure that luminance levels across images were constant, only pictures of basic drawings consisting of black lines were used (see Figure 1b for some examples). The pictures were carefully selected such that they contained an approximately equal number of dark pixels. Consequently, the average pixel level was well matched across pictures (mean: 121, range: 112–126). The images were then rescaled to fit in a box of 400 × 400 pixels, in order to minimize potential retinal size differences, as well as to maintain an approximate equal number of dark pixels in each image. Finally, all pictures were presented in low contrast on the center of the screen and on a dark gray background with an RGB (red/green/blue) level of 128. The baseline screen, which was shown before the presentations of each picture, comprised an empty dark gray frame with the same RGB level as the background of the picture stimuli (i.e., 128).

Procedure. Participants positioned approximately 60 cm from the computer screen performed the same dot-encoding task as in Experiment 1. Each trial started with a fixation cross presented at the center of the screen for 2,000 ms followed by the baseline screen for 500 ms. A picture of an object then appeared for 2,000 ms, with either one or two dots each briefly flashed for 500 ms (see Figure 1b). Participants pressed the left and right control keys on the keyboard for one-dot and two-dot decisions. A total of 100 pictures randomly selected from a pool of 200 pictures was presented during this phase.

Immediately after the study phase, participants were told they would be presented with pictures, some of which they had studied before, and that they were required to judge which pictures they remembered having seen in the study phase. Importantly, they were trained to discriminate between the two kinds of memory they could experience during the recognition test, namely, recollection and familiarity memory. This was done by initially having them read an instruction sheet describing the characteristics of each kind of memory (see Appendix), and then provide the experimenter with examples of each type of memory, to ensure they properly understood the difference between familiarity and recollection memory. All participants tested showed a proper understanding of the difference between the two kinds of memory. In order to reduce the occurrence of involuntary recollection memory, we instructed participants to make their decisions quickly but accurately (Mayes et al., 2007; Montaldi et al., 2006).

After a short practice session, the test phase began. Each test trial commenced with a fixation cross presented for 3,000 ms followed by the baseline screen for 500 ms. Each picture remained on the screen until participants gave a response. The 100 pictures presented at study were randomly intermixed with the remaining 100 novel pictures. Participants pressed predefined keys on the keyboard to give their recollection, familiarity, and new responses.

---

[3] For technical reasons, we used slightly different object stimuli in Experiment 1 and Experiment 2. However, care was taken in choosing pictures that were perceptually very similar between experiments (see Figure 1): both picture sets consisted of clip art line drawings, were colorless, had a similar physical size, and depicted common objects. Nevertheless, we acknowledge that these nonspecific differences may limit confidence in arguing that pupil priming in the present experiment and behavioral priming observed in Experiment 1 are very similar kinds of unconscious object-related memory.

Apparatus. The equipment used in this experiment to record eye movements and pupil data was an Applied Sciences Laboratory (www.asleyetracking.com) eye tracker (Eye-Trac 6000) using a sampling rate of 60 Hz. The eye tracker consisted of a remote camera positioned under the middle of the screen that was used to present the stimuli, at a distance of approximately 60 cm from each participant. A chin rest was used to minimize movement, and the pupil data were recorded from the left eye of each participant. Environmental light conditions were the same for all participants and were kept constant throughout the whole experiment. The presentation computer was connected via a serial connection to the interface computer containing the software that controlled the eye tracker. This connection allowed the interface computer to trigger the eye tracker to initiate the recording of data during the baseline period and presentation of pictures. In order to convert the units utilized by the eye tracker into metric units (i.e., in millimetres), a 4-mm artificial pupil was placed at the location of the participant's eye and its value recorded for later conversion.

Data cleaning, reduction, and analysis. On average, 18% of the trials were excluded using the procedures delineated in the General Pupillometry Method.

Accuracy and RTs were calculated for each response category. For the pupillometric data, both mean pupil dilation (i.e., average value of pupil dilation for each trial) and peak pupil dilation (i.e., maximum value of pupil dilation for each trial) were also calculated for each category. In addition, analyses of the pupillary behavior across time were also performed by splitting up the pupil data for each trial into six categorical variables (bins), each containing an equivalent number of pupil recordings. This was done as a means of standardizing different trials with different RTs, since participants responded at their own pace, thus producing different numbers of pupil recordings for each trial (see Kafkas & Montaldi, 2012, for a similar approach).

The data were analyzed using repeated measures ANOVAs and *t* tests on accuracy, RTs, and pupil dilation scores. The alpha level was set at .05, and statistical tests were two tailed; however, whenever we predicted the direction of the pupillary response (e.g., when testing mean pupil scores against zero, we did not expect pupil constriction), we performed one-tailed *t* tests. Effect sizes are reported in the form of Cohen's *d* and partial eta-squared (η_p^2) where appropriate.

Results

An additional 4% of trials with outlying RTs were excluded from the analyses.

Table 1 (left column) shows the proportion of trials in each response category. A one-sample *t* test was conducted on the familiarity PR scores (assuming stochastic independence) to ascertain whether participants' familiarity memory would be above-chance levels. This test revealed a significant effect, $t(26) = 11.181$, $p < .001$, $d = 4.35$, indicating that participants exhibited reasonably good familiarity memory for studied objects ($M = .25$).

Table 1 (right column) shows the mean RTs for each response category. Paired *t* tests revealed that Fs were slower than both Ms, $t(26) = 2.174$, $p < .05$, $d = .42$, and CRs, $t(26) = 4.247$, $p < .001$, $d = .82$. Importantly, Ms were slower than CRs, $t(26) = 3.376$, $p < .01$, $d = .65$, suggesting that participants had above-chance familiarity levels for Ms that made new decisions for this category more effortful, slower, and harder to make.

Next, we evaluated to what extent the effort-related difficulty in reaching a decision, as indicated by RTs, affected the pupil data for Fs, Ms, and CRs. We included Fs in this analysis to see whether difficulty in making a memory judgment affected Ms and Fs in a similar way. If RTs were found to be related positively to pupil dilation for both Fs and Ms, this would suggest that increased effort amplifies pupil dilation regardless of whether the strength of the familiarity signal is above or below the recognition threshold. For that aim, we split the RT data into six quantiles (sextiles) for each of those three categories and computed the mean pupillary response for each sextile. A linear trend analysis was conducted using a polynomial of the first degree, and the results of this analysis are illustrated on the left panel of Figure 3. Both Fs, $F(1,26) = 4.118$, $p = .05$, $\eta_p^2 = .14$, and Ms, $F(1,26) = 6.294$, $p < .05$, $\eta_p^2 = .20$, showed a linear increase in mean pupil dilation with RTs, whereas the effect was not significant for CRs, $F(1,26) = .012$, $p > .10$, $\eta_p^2 = .0004$.

Because the increase in dilation for Ms (but not for CRs) as a function of RTs, as well as the slower RTs for Ms than for CRs, could indicate memory-related effort, Ms and CRs were matched in terms of RTs. Both overall and nonmemory RT confound matching procedures were used to ensure that any M-related pupil dilation indicated unconscious object memory. To achieve the overall RT confound match, approximately 78% of Ms were matched with CRs, producing an average of 28 matched Ms per participant, and their corresponding matched CRs. After matching, neither the difference in RTs, $t(26) = -.710$, $p > .10$, $d = .14$ ($\mathrm{Ms_{matched}}$: $M = 1369$, $SD = 330$; $\mathrm{CRs_{matched}}$: $M = 1368$, $SD = 328$) nor baseline dilation levels, $t(26) = .127$, $p > .10$, $d = .02$ ($\mathrm{Ms_{matched}}$: 4.47, $\mathrm{CRs_{matched}}$: 4.44), differed significantly between these two categories.

The left panel of Figure 4 shows the mean and peak pupil dilation for the two matched categories. Paired *t* tests revealed that, relative to $\mathrm{CRs_{matched}}$, $\mathrm{Ms_{matched}}$ showed greater pupil dilation for both mean, $t(26) = 2.064$, $p < .05$, $d = .40$, and peak, $t(26) = 2.147$, $p < .05$, $d = .41$, pupil dilation.

Table 1. *Proportion of Responses (Range) and Reaction Times (Standard Error of the Mean) for Each Response Category*

Experiment 2	Proportion %	Reaction times (ms)
H	0.46 (10–76)	1,614 (89)
R	0.07 (0–52)	1,748 (121)
F	0.39 (8–62)	1,600 (88)
M	0.36 (9–56)	1,502 (80)
CR	0.64 (12–91)	1,396 (71)
FA	0.17 (1–37)	1,630 (85)
FA_R	0.008 (0–5)	1,673 (178)
FA_F	0.17 (1–35)	1,625 (87)
Experiment 3		
H	0.79 (32–96)	1,356 (55)
R	0.16 (0–49)	1,472 (66)
F	0.63 (18–94)	1,370 (57)
M	0.13 (4–32)	1,325 (50)
CR	0.79 (50–97)	1,161 (49)
FA	0.13 (1–36)	1,443 (72)
FA_R	0.006 (0–6)	1,875 (266)
FA_F	0.12 (1-36)	1,412 (71)

Note. H plus M (as well as CR plus FA) do not sum up to 1 due to "bad" trials that were excluded prior to statistical analysis (see General Pupillometry Method for a description). H = hit rate; R = recollection hit rate; F = familiarity hit rate; M = misses; CR = correct rejections; FA = false alarm rate; FA_R = recollection false alarm rate; FA_F = familiarity false alarm rate.

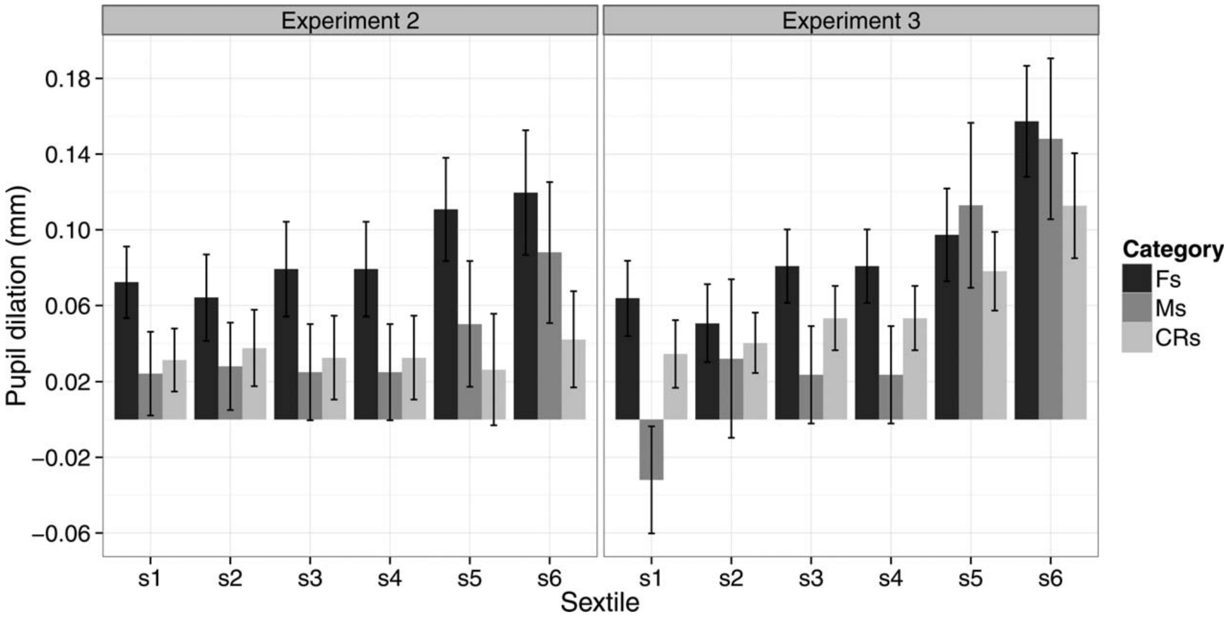

Figure 3. Mean pupil dilation for familiarity hits (Fs), misses (Ms), and correct rejections (CRs) as a function of reaction times (RTs) in Experiment 2 (left) and 3 (right). The first sextile (s1) contains the trials with fastest RTs, whereas the last sextile (s6) contains the trials with the slowest RTs. Sextiles s2–s5 contain trials with intermediate speeds. Error bars represent the standard error of the mean.

Even though RT-matched Ms showed larger pupil dilation than CRs, we used the nonmemory RT confound matching procedure (see General Pupillometry Method) to ensure that this effect reflected unconscious object memory rather than a failure to match for nonmemory-related difficulty of the pictures selected as Ms and CRs. Such a match failure would have meant that the enhanced pupil dilation for Ms, matched overall, was related to above-chance levels of familiarity. We compared a modified version of the RTs of the Ms$_{\text{matched}}$ and CRs$_{\text{matched}}$ in which, for each participant and picture by picture, the mean CR-RT score was substituted for each individual's M or CR RT score. There was no significant difference between Ms$_{\text{matched}}$ (1,382 ms) and CRs$_{\text{matched}}$ (1,379 ms) when this modified mean CR-RT measure was used, $t(26) = .191$, $p > .10$, $d = .04$, which showed that the perceptual difficulty (nonmemory-related processing speed) of Ms$_{\text{matched}}$ and CRs$_{\text{matched}}$ was also matched. Therefore, the familiarity level of Ms must have been effectively at chance and the enhanced pupil dilation of Ms must have reflected unconscious object memory of some kind rather than an effect of above chance, but below threshold levels of familiarity.

To confirm that the pupil effect was not mainly confined to slower Ms and CRs, we tested whether a similar pupillary dilation difference would occur for fast and slow Ms and CRs that were RT matched. We divided trials into a fast and slow group based on a median split and calculated the pupillary response difference between Ms$_{\text{matched}}$ and CRs$_{\text{matched}}$ for each group. Only participants with more than six trials in each condition were included in this analysis (four participants excluded). As we predicted, the difference was identical between the two speed groups (see Figure 5, left), $t(22) = -.141$, $p > 0.10$, $d = .03$.

We next analyzed the time course of the pupillary response using a 2 Category (Ms$_{\text{matched}}$, CRs$_{\text{matched}}$) × 6 Bin repeated measures ANOVA (see Method section for an explanation on how the bins were constructed). There was only a significant main effect of category, $F(1,26) = 4.193$, $p = .05$, $\eta_p^2 = .14$ (Figure 6, left),

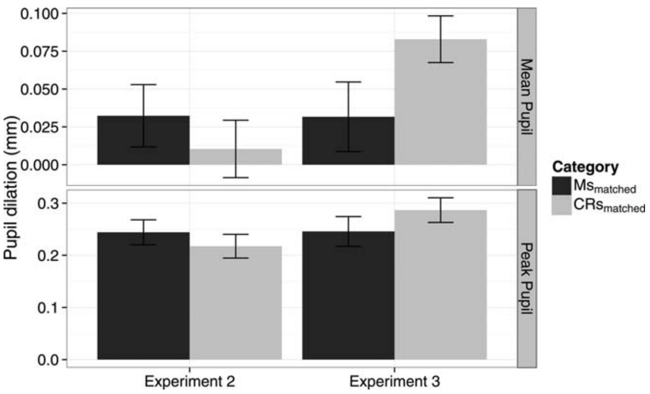

Figure 4. Mean and peak pupil dilation for misses (Ms$_{\text{matched}}$) matched in terms of reaction times with correct rejections (CRs$_{\text{matched}}$) in Experiment 2 (left) and 3 (right). Error bars represent the standard error of the mean.

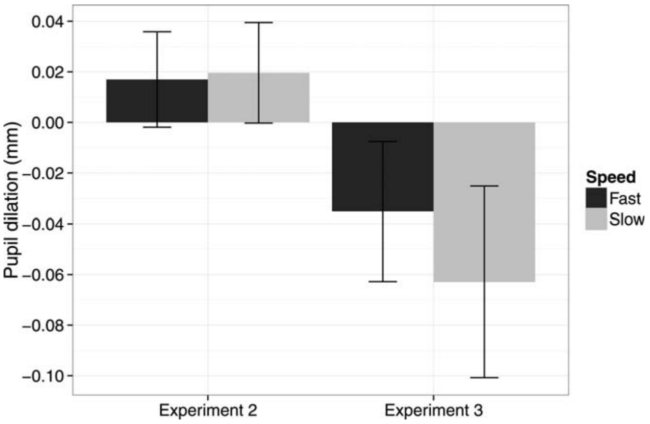

Figure 5. Mean pupil dilation difference between Ms$_{\text{matched}}$ and CRs$_{\text{matched}}$ shown separately for fast and slow responses (division based on a median split) in Experiment 2 (left) and Experiment 3 (right). Error bars represent the standard error of the mean.

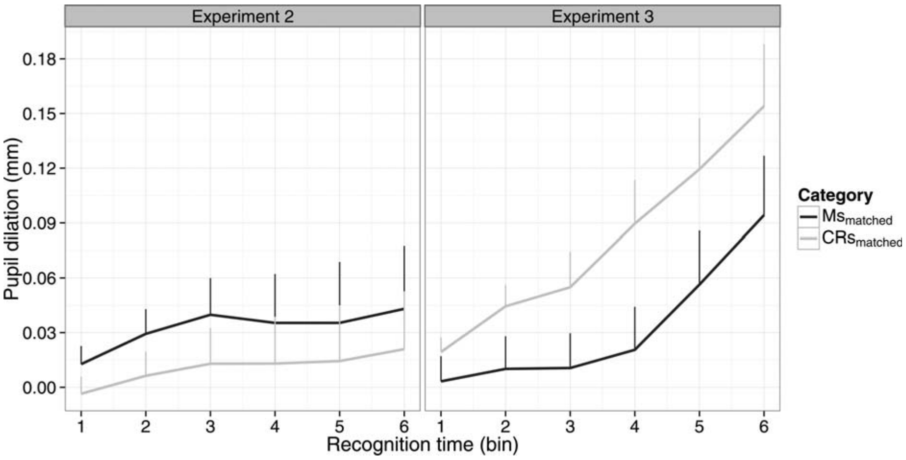

Figure 6. Time course of the pupillary response (from stimulus onset—bin 1 until participants' response—bin 6) for RT-matched misses (Ms$_{matched}$) and RT-matched correct rejections (CRs$_{matched}$) in Experiment 2 (left) and 3 (right). Error bars represent the standard error of the mean.

indicating how mean pupil dilation developed over time for Ms$_{matched}$ relative to CRs$_{matched}$. One-sample *t* tests revealed that only bins 2 and 3 were significantly above baseline for the Ms$_{matched}$ category (bin 2: $t(26) = 2.166$, $p < 0.05$, $d = .42$; bin 3: $t(26) = 1.977$, $p < 0.05$, $d = .38$), whereas pupil size did not differ from baseline at any level of bin for the CRs$_{matched}$ category (all *p*s > 0.26).

Finally, given that slower Fs and Ms produced greater pupil dilation, Fs and Ms were RT matched in order to see whether Fs produced greater pupil dilation than Ms when effort was matched. If there was a pupil dilation difference, it should have taken time to develop, so we also analyzed the temporal evolution of the pupillary response for Fs$_{matched}$ and Ms$_{matched}$ by splitting the pupil data into six bins for each category. RTs between these two categories were well matched (Ms$_{matched}$: $M = 1{,}486$ ms, $SD = 380$, Fs$_{matched}$: $M = 1{,}485$ ms, $SD = 379$; $t(26) = .458$, $p > .10$, $d = .09$). A 2 Category × 6 Bin repeated measures ANOVA revealed a main effect of category, $F(1,26) = 6.082$, $p < .05$, $\eta_p^2 = .19$, which indicated that Fs$_{matched}$ produced a greater pupil dilation than Ms$_{matched}$. The main effect of bin was also significant, $F(1.350,35.109) = 4.228$, $p < .05$, $\eta_p^2 = .14$, indicating that pupil dilation increased across time. More importantly, the interaction was highly significant, $F(1.532,39.821) = 11.584$, $p < .001$, $\eta_p^2 = .31$, which indicated that, whereas Fs$_{matched}$ showed a linear increase across recognition time bins, Ms$_{matched}$ did not (see Figure 7, left). As would be expected, the greater size of the pupillary response for Fs took time to emerge, beginning by the third bin. This greater size of the pupil for Fs following RT matching indicates that this dilation is a direct consequence of familiarity because M and F effort was matched.

Discussion

In the present experiment, pupil priming was investigated using three different measures for determining pupil size differences (mean pupil diameter, peak pupil diameter, and pupil diameter across time) between forgotten studied objects (Ms) and correctly rejected new objects (CRs) matched in terms of RTs. The results confirmed the usefulness of pupillometry as a measure of priming. The main finding was the observation of an increase in both mean and peak pupillary dilation for Ms relative to CRs that were RT matched both overall and for nonmemory-related effort. This double confound check procedure showed that the RT match was not confounded by Ms being based on pictures for which nonmemory-related processing was easier. This lack of confound strongly implied that familiarity levels of Ms matched those of CRs and must, therefore, have been at chance, provided, of course, that our arguments for the RT-matching procedures are correct. The Ms$_{matched}$ versus CRs$_{matched}$ pupil memory effect was, therefore, unlikely to have been caused by above baseline but subthreshold familiarity memory, related to the need to use greater effort in decision making. Finally, although the analysis of pupil dilation across time only revealed a main effect of category (reproducing the mean pupil effect), Ms$_{matched}$ deviated from baseline relatively early in time whereas CRs$_{matched}$ did not differ from baseline at any time bin.

Even though the pupil effect for Ms$_{matched}$ strongly suggests a completely unconscious kind of object memory, it remains possible that the memory indicated by the pupil dilation for Ms$_{matched}$ is based on the same memory storage system as familiarity memory, with the latter being supported by a stronger memory trace in the same memory storage system. This possibility is consistent with the greater pupil dilation shown with Fs$_{matched}$ than Ms$_{matched}$ and also with the emergence of this greater pupil dilation of Fs as a function of the time following stimulus presentation. This pattern of pupillary response is also compatible with M and F memories being based on distinct memory systems, so our findings are not conclusive either way (see General Discussion for further discussion of these points).

Experiment 3

In Experiment 2, the study task directed participants' attention away from the objects' meaning. It was shown that the pupil dilated more for Ms$_{matched}$ than CRs$_{matched}$ during a subsequent recognition memory test, a phenomenon we called pupil priming. The aim of Experiment 3 was to determine whether this pupil priming effect would also be found when using a different encoding task. In this experiment, participants decided, at study, whether an object was bigger than a shoebox whereas, at test, they engaged in the familiarity-only recognition memory test. Importantly, the study task of the present experiment required a different kind of study processing from that of Experiment 2: specifically, a conceptual task (size judgment) versus a perceptual task (dot counting), so processing in the present experiment was more elaborate.

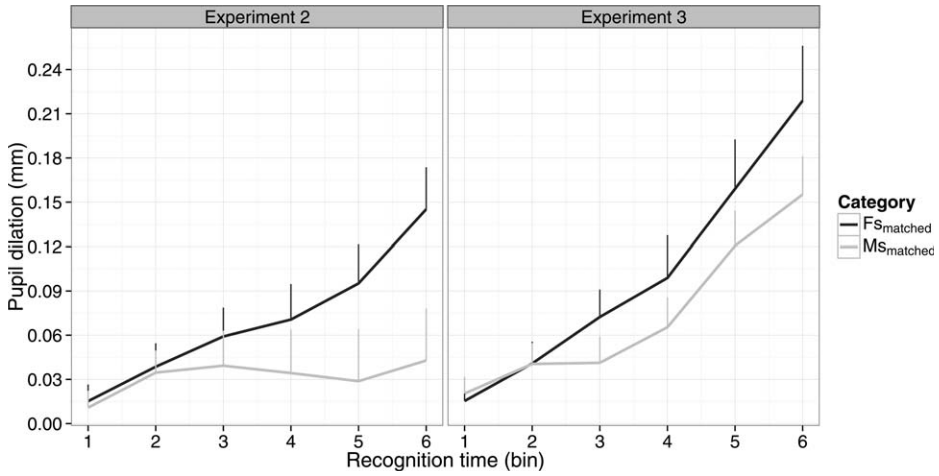

Figure 7. Time-course of the pupillary response (from stimulus onset—bin 1 until participants' response—bin 6) for RT-matched familiarity hits ($Fs_{matched}$) and RT-matched misses ($Ms_{matched}$) in Experiment 2 (left) and 3 (right). Error bars represent the standard error of the mean.

The main aim of this experiment was to ascertain whether the pupil-priming effect observed in the previous experiment would generalize to a second experiment in which only the encoding task differed. Importantly, in Experiment 1, we showed that deeply and shallowly encoded objects exhibited equivalent behavioral priming, and thus we were confident that any pupillary response difference between Ms and CRs in the present experiment would be related to that behavioral priming. It should be pointed out that an animacy task was used in Experiment 1 rather than a size-judgment task. We used a different encoding task in this experiment because we needed a large number of objects meeting strict selection criteria, with half falling into one category and half into another. It proved impossible to select 100 animate objects that matched our criteria (e.g., equal luminosity, similar line complexity) without risking overlapping conceptual information (e.g., two kinds of bird). However, participants needed to encode and extract semantic information in both tasks, so both should have produced similar behavioral priming.

Method

Participants. Thirty undergraduate students of the University of Manchester were recruited ($M = 20.48$, $SD = 5.54$). All participants had normal or corrected-to-normal vision and gave informed written consent to take part in this study. Two participants were excluded due to difficulties with discrimination of the pupil, and one additional participant was excluded due to exaggerated movement.

Materials. The materials used in this experiment were mostly taken from Experiment 2. A total of 200 objects was also used in this experiment. Half of these objects were smaller than a shoebox whereas the other half were bigger.

Procedure. At study, participants were told they were going to see a series of pictures of common objects on the screen and were required to decide whether each object was bigger than a standard shoebox in real life. Each study trial started with the presentation of a fixation cross for 2,000 ms, followed by the baseline screen for 500 ms and a picture of an object presented for 2,000 ms. Participants pressed the left and right control keys for bigger and smaller decisions. Piloting showed that recognition memory was close to ceiling if tested immediately after the size-classification task (and, consequently, there would not have been a sufficient number of $Ms_{matched}$); therefore, the study and test phases were separated by a break of approximately 20 min, which included three distractor tasks followed by the test instructions. The three distractor tasks consisted of solving arithmetic problems, writing down 15 cities in the United Kingdom, and reading an excerpt from a book, from which participants were asked to memorize as many details as possible. After the three distractor tasks were completed, participants were informed they would have to perform a recognition memory test (the instructions were identical to the ones given in Experiment 2).

Data cleaning, reduction, and analysis. On average, 8% of the trials were excluded from further analyses, using the procedures delineated in the General Pupillometry Method section. The experimental design, the statistical tests performed, and their corresponding significance level were identical to the previous experiment.

Results

An additional 4% of trials with outlying RTs were excluded from the analyses.

Table 1 (left column) shows the proportion of trials for each response category during this experiment. In order to ascertain whether participants' familiarity memory was above chance, a one-sample t test was conducted on the familiarity PR scores (assuming stochastic independence). This test revealed a significant effect, $t(26) = 18.000$, $p < .001$, $d = 7.06$, indicating that familiarity memory was relatively high in this experiment ($M = .63$), and, indeed, significantly higher than familiarity memory in Experiment 2 ($M = .25$), $t(44.593) = -9.015$, $p < .001$, $d = 2.70$.

Table 1 (right column) shows the mean RTs for each category. Paired t tests revealed Fs were slower than both CRs and Ms, although the effect only reached significance for the Fs versus CRs comparison, $t(25) = 5.180$, $p < .001$, $d = 1.01$. Importantly, Ms were slower than CRs, $t(26) = 5.574$, $p < .001$, $d = 1.07$, which, as with the previous experiment, could suggest some degree of subthreshold, but above-chance familiarity memory for the former category.

We then tested for a linear trend between pupil dilation and RTs for Fs, Ms, and CRs (see Figure 3, right). Fs: $F(1,26) = 9.743$,

$p < .01$, $\eta_p^2 = .27$; Ms: $F(1,23) = 11.458$, $p < .01$, $\eta_p^2 = .33$; and CRs: $F(1,26) = 6.874$, $p < .01$, $\eta_p^2 = .21$, each showed a linear increase in mean pupil dilation as RTs became slower.

As with Experiment 2, Ms and CRs were RT matched using the overall and nonmemory RT confound matching procedures. Given that participants in the current experiment produced, as expected, fewer Ms than in the previous experiment, only those who had six or more $\text{Ms}_{\text{matched}}$ were included in the matching analysis (five participants excluded). On average, approximately 90% of Ms were possible to match with CRs using the overall RT confound matching procedure. After the matching procedure, neither the difference in RTs, $t(21) = -1.146$, $p > .10$, $d = .24$ ($\text{Ms}_{\text{matched}}$: $M = 1228$, $SD = 239$, $\text{CRs}_{\text{matched}}$: $M = 1230$, $SD = 241$) nor baseline dilation levels, $t(21) = -.312$, $p > .10$, $d = .06$ ($\text{Ms}_{\text{matched}}$: $M = 4.34$, $\text{CRs}_{\text{matched}}$: $M = 4.33$), differed significantly between these two categories.

Figure 4 (right) shows the results from the main analysis. Paired *t* tests revealed a significant result for the mean pupil dilation, $t(21) = -2.102$, $p < .05$, $d = .45$, with greater pupil dilation for $\text{CRs}_{\text{matched}}$ than $\text{Ms}_{\text{matched}}$. For the peak pupil dilation data, although greater dilation was also observed for $\text{Ms}_{\text{matched}}$ relative to $\text{CRs}_{\text{matched}}$, this effect did not approach significance, $t(21) = -1.555$, $p > .10$, $d = .33$, presumably due to the peak pupil data being slightly noisier than the mean pupil data. The nonmemory RT confound check procedure found no significant difference between $\text{Ms}_{\text{matched}}$ (1,129 ms) and $\text{CRs}_{\text{matched}}$ (1,138 ms), $t(21) = -.717$, $p > .10$, $d = .15$, indicating that nonmnemonic processing speed of $\text{Ms}_{\text{matched}}$ and $\text{CRs}_{\text{matched}}$ was well matched. It can, therefore, be inferred that familiarity was at chance for the $\text{Ms}_{\text{matched}}$. The pupil effect was also independent of response speed because the difference between $\text{Ms}_{\text{matched}}$ and $\text{CRs}_{\text{matched}}$ was equivalent for fast and slow trials, $t(26) = -.700$, $p > .10$, $d = .15$ (see Figure 5, right).

Next, the analysis of the pupil time course was conducted on the matched M and CR categories. A 2 Category ($\text{Ms}_{\text{matched}}$, $\text{CRs}_{\text{matched}}$) × 6 Bin repeated measures ANOVA revealed a significant main effect of category, $F(1,21) = 4.428$, $p < .05$, $\eta_p^2 = .17$, and bin, $F(1.460,30.661) = 13.903$, $p < .001$, $\eta_p^2 = .40$, as well as a significant interaction, $F(3.219,67.601) = 2.695$, $p < .05$, $\eta_p^2 = .11$, which resulted from a delay in the time course of pupil dilation for $\text{Ms}_{\text{matched}}$ relative to $\text{CRs}_{\text{matched}}$ (see Figure 6, right). Indeed, one-sample *t* tests showed that, whereas every bin in $\text{CRs}_{\text{matched}}$ was above chance (all $p < 0.01$), only bin 5 and 6 were significantly greater than zero for $\text{Ms}_{\text{matched}}$ (bin 5: $p < 0.05$, bin 6: $p < 0.01$). Thus, the analysis of the pupillary time course in the current experiment revealed a distinct pattern of pupil dilation across time between $\text{Ms}_{\text{matched}}$ and $\text{CRs}_{\text{matched}}$, with the latter showing an earlier and consistent increase in pupil dilation relative to the former.

Finally, the pupillary time course of RT-matched Fs and Ms was examined. RTs were well matched between these two categories ($\text{Ms}_{\text{matched}}$: $M = 1{,}327$ ms, $SD = 251$; $\text{Fs}_{\text{matched}}$: $M = 1{,}328$ ms, $SD = 251$; $t(26) = -.793$, $p > 0.10$, $d = .15$). A 2 Category × 6 Bin repeated measures ANOVA revealed a main effect of bin, $F(1.678,2.815) = 24.560$, $p < .001$, $\eta_p^2 = .54$, as well as a significant interaction, $F(2.815,59.110) = 3.495$, $p < .05$, $\eta_p^2 = .14$, indicating that the temporal evolution of the pupillary response for $\text{Fs}_{\text{matched}}$ and $\text{Ms}_{\text{matched}}$ was distinct (see Figure 7, right). F-related pupil dilation occurred following RT matching, thus indicating a direct effect of familiarity that was independent of effort.

Interexperimental Analyses

Surprisingly, the two experiments showed different patterns of pupillary responses for $\text{Ms}_{\text{matched}}$ and $\text{CRs}_{\text{matched}}$, with greater pupil dilation for $\text{Ms}_{\text{matched}}$ relative to $\text{CRs}_{\text{matched}}$ in Experiment 2 but the reverse pattern in Experiment 3. To check the statistical validity of this difference, a 2 Experiment (Experiment 2, Experiment 3) × 2 Category ($\text{Ms}_{\text{matched}}$, $\text{CRs}_{\text{matched}}$) mixed repeated measures ANOVA was conducted separately for mean and peak pupil scores, using experiment as a between-subject factor.

Although the main effect of category did not approach significance for either mean or peak pupil dilation (mean pupil: $F(1,47) = 1.385$, $p > .10$, $\eta_p^2 = .03$; peak pupil: $F(1,47) = .268$, $p > .10$, $\eta_p^2 = .01$), the interaction term was significant for both pupillometric measures (mean pupil: $F(1,47) = 8.619$, $p < .01$, $\eta_p^2 = .16$; peak pupil: $F(1,47) = 6.085$, $p < .05$, $\eta_p^2 = .12$), confirming the directional difference in pupil scores across experiments (see Figure 4). Importantly, when the same ANOVA was conducted on the baseline pupil dilation levels, neither the main effect of category, $F(1,47) = .373$, $p > .10$, $\eta_p^2 = .01$, nor the interaction, $F(1,47) = .194$, $p > .10$, $\eta_p^2 = .004$, was significant, suggesting that the difference in the direction of the mean/peak pupillary response cannot be accounted for by differential pupil baseline levels.

For the pupillary time-course analysis, a 2 Experiment × 2 Category × 6 Bin mixed repeated measures ANOVA revealed a main effect of bin, $F(1.407,66.106) = 7.865$, $p < .01$, $\eta_p^2 = .14$, reflecting the gradual increase in dilation across time. The interaction between experiment and category was also significant, $F(1,47) = 8.598$, $p < .01$, $\eta_p^2 = .16$, replicating the ANOVA conducted above with the mean pupil. The interaction between experiment and bin was also significant, $F(1.407,66.106) = 4.136$, $p < .05$, $\eta_p^2 = .08$, with greater increase in pupil dilation across time in Experiment 3 than in Experiment 2. More importantly, the three-way interaction approached significance, $F(2.568,120.708) = 2.227$, $p = .098$, $\eta_p^2 = .05$, confirming our previous observation of distinct pupillary time courses for $\text{Ms}_{\text{matched}}$ and $\text{CRs}_{\text{matched}}$ across experiments.

Discussion

Surprisingly, in the current experiment, pupil priming was identified as a decrease in pupillary response for $\text{Ms}_{\text{matched}}$ relative to $\text{CRs}_{\text{matched}}$. Although this finding is in marked contrast to the greater dilation for $\text{Ms}_{\text{matched}}$ than for $\text{CRs}_{\text{matched}}$ responses found in Experiment 2, it is, nevertheless, consistent with Laeng et al.'s (2007) finding of decreased pupil dilation for Ms relative CRs, when a conceptual task was also used at encoding.

It should be noted that, apart from the study–test lag and different kinds of encoding, the study and test conditions were purposefully made similar across experiments (e.g., same timing parameters, same kind of stimuli, same testing room, equipment, and experimenter). The study–test lag difference between the experiments (approximately 5 min in Experiment 2 and 20 min in Experiment 3) was necessary because had we not introduced a delay between study and test phases in Experiment 3, we would not have had enough $\text{Ms}_{\text{matched}}$ to analyze. However, we find it unlikely that a slight temporal difference in study–test lag would lead to such a dramatic reversal in the pupil effects, especially considering that the same number of intervening stimuli were presented between study and test phases in both Experiment 2 and 3. It is more likely that the direction of the pupillary response was influenced by other factors, as will be considered in the General Discussion.

General Discussion

This study was conducted in order to evaluate whether pupillometry could be used as a technique to detect priming without any recognition being present. In Experiment 1, participants engaged in both a perceptual (dot-counting) and a conceptual (animacy) encoding task. Then, at test, they performed an object decision task followed by a recognition memory test on a trial-by-trial basis. Reliable behavioral priming was observed for both perceptually and conceptually studied objects that were not recognized (i.e., Ms). In Experiment 2, we recorded pupil data while participants performed a recognition memory test after having engaged in the same perceptual task as in Experiment 1. We observed a pupil priming effect, which was indicated by an increase in dilation for Ms relative to correctly rejected novel objects (CRs), when these had been matched, using two procedures, in terms of RTs. In Experiment 3, participants engaged in a recognition memory test, after having classified objects with respect to their size at study, and pupil priming was identified as a decrease in pupil dilation for $\text{Ms}_{\text{matched}}$ relative to $\text{CRs}_{\text{matched}}$. To our knowledge, our study was the first to successfully identify changes in pupil diameter that relate to long-term priming in control participants and that occur when recognition memory and familiarity in particular are completely at chance.

A possible explanation for the different direction of pupillary responses between the two experiments is that pupil priming depended on the different kinds of processing at test that were triggered by processing differences during encoding, determining the pupillary response for $\text{Ms}_{\text{matched}}$. Indeed, priming has often been attributed to the reinstatement, during retrieval, of either perceptual or conceptual study processes, with this processing overlap between study and test being the source of increased memory performance (e.g., Blaxton, 1989; Roediger, 1990). Given that the study task of Experiment 2 and Experiment 3 emphasized perceptual and conceptual processing, respectively, the differential pupillary response between $\text{Ms}_{\text{matched}}$ and $\text{CRs}_{\text{matched}}$ could have reflected this difference in processing type during the categorization of $\text{Ms}_{\text{matched}}$ objects.

It is important to note, however, that the above explanation assumes that the pupillary behavior for $\text{Ms}_{\text{matched}}$ changed across experiments, since $\text{Ms}_{\text{matched}}$ but not $\text{CRs}_{\text{matched}}$ objects had been processed at study. In stark contrast to this, $\text{Ms}_{\text{matched}}$ exhibited virtually the same amount of mean/peak dilation in both experiments (see Figure 4). Thus, pupil priming did not result from $\text{Ms}_{\text{matched}}$ exhibiting less dilation in Experiment 3 than in Experiment 2, but rather from greater pupillary response to $\text{CRs}_{\text{matched}}$ in Experiment 3 than in Experiment 2. A simple interpretation of these effects based on processing type is therefore insufficient, since it does not account for the pupillary response changes of $\text{CRs}_{\text{matched}}$ between the two experiments.

A plausible explanation for the differential pupillary response for $\text{CRs}_{\text{matched}}$ across experiments could relate to expectation of novelty. Weiskrantz (1998, p. 325) argued that, since the pupil is under the control of the autonomic nervous system, the pupillary response "is generated by a primitive, adaptive system that is especially tuned to the detection of sudden, novel occurrences." There is considerable evidence that these novelty-related pupil dilation responses increase when likelihood and expectancy of novel events is lower (e.g., Friedman, Hakerem, Sutton, & Fleiss, 1973; Qiyuan, Richer, Wagoner, & Beatty, 1985). Thus, the different dilation levels for $\text{CRs}_{\text{matched}}$ across experiments could relate to how weak the expectation of detecting novel objects is and, relatedly, how much attention is paid to them as a result of their perceived significance. Participants were not informed about the ratio of old to novel test stimuli, so the expectation about the frequency of novel objects likely developed differently between experiments as a function of the proportion of judgments that were new. Indeed, the difference between old and new responses, regardless of study status, in Experiment 2 (76 old vs. 124 new) far exceeded that in Experiment 3 (99 old vs. 101 new). We calculated the expectation of novelty using the formula [(CRs + Ms)/(CRs + Ms + Hs)] and compared the scores across experiments. The difference between the scores was highly significant, $t(47) = 6.696$, $p < .001$, $d = 1.95$, suggesting a much higher expectation of novel objects in Experiment 2 (.69) than in Experiment 3 (.55), and, consequently, greater pupil dilation in the latter experiment. Importantly, expectancy seems to influence not only mean and peak pupil diameter but also the evolution of the response across time: whereas $\text{CRs}_{\text{matched}}$ in Experiment 2 exhibited a relatively flat pupillary response across time (in fact, it did not differ from zero), in Experiment 3 $\text{CRs}_{\text{matched}}$ evoked a large and significant response quite early in time, with a constant increase in dilation throughout the entire time window (see Figure 6). This finding, in particular, is consistent with the notion that the perceived significance of CRs was augmented in Experiment 3 as the result of lower expectation of novel objects.[4] Relative to the shallow-encoding task of Experiment 2, the deep-encoding task used in Experiment 3 most probably led to the creation of a stronger memory trace for old, studied stimuli. This may have had the additional effect of enhancing the salience of novel stimuli, leading to a high engagement in the detection of novelty during the recognition memory test.

However, the mean/peak pupil diameter for Ms did not seem to change between Experiments 2 and 3 as a function of expectancy. Assuming that the matching procedure has eliminated any residual familiarity contribution to the pupillary responses for Ms, it is plausible to argue that the brain has detected that, with Ms, the objects have been studied, but that the identifying brain region(s) is unable to produce a feeling of aware memory for such objects. However, activity in the "unconscious oldness-detecting" region(s) inhibits any enlargement of the pupil's mean/peak diameter that would otherwise have been caused by low novelty expectation. In other words, the failure of the mean/peak pupil diameter for Ms to change as a function of probability of detecting novelty shows that oldness is being indicated unconsciously by the different pupillary response patterns of Ms and CRs in Experiments 2 and 3.

The expectancy view proposes that mean/peak pupil dilation is greater when novelty expectation is lower (as in Experiment 3), but that this response is inhibited for stimuli that the brain unconsciously identifies as having been studied. This view implies that the mean/peak amplitude of the CRs' pupillary response will increase according to an unknown function as expectancy decreases. Therefore, with high expectation of novelty, CRs' pupillary responses will be smaller than Ms' pupillary responses and, with low expectation of

---

[4] It should be noted that we are not arguing that novel objects were somehow unexpected in Experiment 3-we still observed the typical pupil old/new effect (i.e., greater pupil dilation for Fs than CRs; data not shown), suggesting that the levels of novelty expectation did not exceed the levels of expectation of familiar stimuli. The proposal of lower levels of novelty expectation in Experiment 3 relative to Experiment 2 simply attempts to explain the differential pupillary response for CRs across experiments.

novelty, they will be larger than Ms' pupillary responses, but at some point in between, the pupil diameter for CRs and Ms should not differ. This needs to be demonstrated without altering encoding and other variables. If true, it still indicates that Ms' and CRs' mean/peak pupillary responses differ because they respond differently to the expectancy variable. Therefore, to confirm that mean/peak pupil diameter for appropriately matched Ms and CRs indicate unconscious stimulus memory, it will be necessary to compare the pupillary responses over several different novelty expectancy levels, while keeping other factors as constant as possible.

Also, it should be noted that expectancy as calculated above is based on conscious detection of novelty because it includes both CRs and Ms. However, it is possible that unconscious detection of oldness disrupts the pupil response to novelty. It may be, therefore, that expectancy is better indicated as the ratio of new judgments for which oldness has been neither consciously nor unconsciously detected to old judgments as well as new judgments where there is neither conscious nor unconscious detection of oldness. Future work will be needed to determine whether this kind of expectancy predicts pupillary responses to CRs better than the formula that was used above.

Our general finding that the pupillary response distinguishes properly RT-matched Ms and CRs is consistent with the findings of other work using ERPs and fMRI. Several studies have found qualitatively different ERP signatures between Ms and CRs, particularly in the time window range 200–500 ms (e.g., Friedman, 2007; Rugg et al., 1998; Walla, Endl, Lindinger, Deecke, & Lang, 1999). Similarly, some neuroimaging research indicates that activity in brain structures involved in recognition memory, such as the medial temporal lobes (the hippocampus and parahippocampal gyrus), can distinguish between old and new items irrespectively of participants' subjective recognition memory judgments (e.g., Daselaar, Fleck, & Cabeza, 2006; Kirwan, Shrager, & Squire, 2009). Our study is consistent with these findings, and suggests that pupillometry can also be used to determine the actual old/new status of test items.

Although the evidence considered so far does not convincingly show that perceptual versus conceptual processing of stimuli differentially influences the pattern of Ms and CRs mean/peak pupil diameter if other factors, such as novelty expectancy, are kept constant, other ERP evidence supports the idea that kind of stimulus processing may differentially influence the temporal course of development of the $\mathrm{Ms}_{\mathrm{matched}}$ pupillary response. Several studies have identified a positive-going peak between 120–200 ms (the P150 effect), which is believed to reflect perceptual priming (e.g., Harris, Cutmore, O'Gorman, Finnigan, & Shum, 2009; Schendan & Kutas, 2003; Voss & Paller, 2010; Voss, Schendan, & Paller, 2010). A close inspection of the pupillary time course of Experiment 2 indicated that $\mathrm{Ms}_{\mathrm{matched}}$ deviated from baseline during bins 2 and 3—the period corresponding to approximately 450 ms and 680 ms poststimulus (see Figure 7). This period corresponds to the first half of the often present "double bump" in pupillometric studies, which has been associated with stimulus-driven processing (e.g., Laeng, Ørbo, Holmlund, & Miozzo, 2011; Murphy, Robertson, Balsters, & O'Connell, 2011). To the extent that certain ERP amplitudes covary with the pupil dilation response, which suggests that both measures are indexing similar kinds of processing (e.g., Friedman et al., 1973; Kuipers & Thierry, 2011; Nieuwenhuis, De Geus, & Aston-Jones, 2011), this early pupil effect may reflect access to, or retrieval of, a perceptual representation.

In contrast, in Experiment 3, $\mathrm{Ms}_{\mathrm{matched}}$ only started to show a noticeable increase in dilation (above baseline) roughly between the fourth and fifth bin (between 800–1,000 ms poststimulus), after which a steep increase in dilation was observed. This time signature may correspond to the FN400 component (between 300–500 ms) in ERP studies, which some researchers have argued reflects a form of conceptual priming (e.g., Voss, Lucas, & Paller, 2010; Voss & Paller, 2006, 2007). The conceptual study task used in Experiment 3 required participants to think about the concept of each object, so retrieval of encoded conceptual information about studied objects was much more likely to occur in Experiment 3 than in Experiment 2.

The slightly later onset of the pupil priming effect (approximately 500 and 900 ms for Experiment 2 and 3, respectively) than for their putative ERP correlates (P150 and FN400 for Experiment 2 and 3, respectively) is unsurprising and has been noticed elsewhere in different experimental contexts (e.g., Kuchinke, Võ, Hofmann, & Jacobs, 2007; Kuipers & Thierry, 2011; Laeng et al., 2011). Contrary to the almost instantaneous detection of brain activity changes allowed by ERPs, the pupil dilation response usually has a latency of around 400 ms after the onset of a cognitive demand (e.g., Kuchinke et al., 2007; Nieuwenhuis et al., 2011).

Although we have argued that behavioral priming and the possibly closely related pupil memory can be found without any recognition memory, there is still an unresolved debate over whether priming and recognition memory (particularly familiarity) depend on a common memory storage system or two distinct, if overlapping, memory storage systems. There are possible ways in which even these kinds of familiarity and priming could both rely on the same memory storage system rather than distinct systems, even when familiarity is not triggered by a priming-related fluency attribution, which is surely common. First, and contrary to our argument, our RT match of Ms and CRs may not have succeeded, so that familiarity strength was still above chance for Ms. If this was so, pupil priming for $\mathrm{Ms}_{\mathrm{matched}}$ could have been based on subthreshold but above-chance familiarity. We think it very unlikely that this happened considering how well matched Ms and CRs were so as to equate memory and nonmemory-related effort, which implies that familiarity should have been at chance. However, it must be acknowledged that yet unidentified factors could be confounding our comparisons so as to make our RT-matching procedures insufficiently effective.

Second, even if there was no residual familiarity for $\mathrm{Ms}_{\mathrm{matched}}$ in the critical comparison, it could be that familiarity falls to chance at higher levels of trace strength of the same storage system that supports priming. This trace strength level might be sufficient to support robust unconscious memory, which only falls to chance at lower trace strength levels. As Duss et al. (2014) recently found in support of this single system view for a relational kind of associative priming, both recognition and priming activated the hippocampus (a likely storage region), although much more activity was found for recognition, which was more sensitive to small hippocampal lesions, than priming. It is less clear, however, whether the kind of item priming identified here corresponds to a weaker level of the same storage system that supports corresponding item familiarity memory. Future work with imaging and lesion studies is needed to address this issue.

This study has several limitations. First, we only analyzed pupil data up to participants' response, so we are presently unable to determine whether any significant changes in pupil dilation occur beyond this time period. Extending the interstimulus interval to allow a longer recording window would have reduced the quality of the pupil data by prolonging a demanding, repetitive task with many trials, set in a dimly lit room and using a chin rest. Perhaps future experiments could use a response deadline procedure and examine the effects of postdecision pupillary effects. Second, a potential problem with the current experiments may be the relatively low number of $\mathrm{Ms}_{\mathrm{matched}}$, particularly in Experiment 3. As increasing the number of stimuli so as to obtain more $\mathrm{Ms}_{\mathrm{matched}}$ may introduce other difficulties (e.g., stimulus overlap, fatigue), using a different design (e.g., containing several breaks and multiple runs) could be a suitable alternative. Third, it will be important that future studies show a more direct relationship between behavioral and pupil priming, perhaps by recording pupillary responses during speeded priming tasks that are not likely to lead to high levels of recognition memory. Finally, it will be also interesting to see whether pupillometry could also indicate other kinds of long-term priming such as negative priming (e.g., Grison, Tipper, & Hewitt, 2005), stimulus-response learning (e.g., Dennis & Perfect, 2013; Horner & Henson, 2009, 2011, 2012), or priming for novel associations (e.g., Gomes & Mayes, 2014).

Conclusion

Pupil priming was identified in Experiment 2 as an increase in pupil dilation for $\mathrm{Ms}_{\mathrm{matched}}$ relative to $\mathrm{CRs}_{\mathrm{matched}}$ after participants engaged in a perceptual encoding task, whereas following a conceptual encoding task pupil priming was identified as decreased pupil dilation for $\mathrm{Ms}_{\mathrm{matched}}$ relative to $\mathrm{CRs}_{\mathrm{matched}}$ in Experiment 3. The pupil priming effect resulted from different cognitive processes operating independently for $\mathrm{Ms}_{\mathrm{matched}}$ and $\mathrm{CRs}_{\mathrm{matched}}$. Specifically, task-evoked dilation for $\mathrm{CRs}_{\mathrm{matched}}$ may have been produced by low novelty expectancy, whereas completely unconscious object memory probably drove the pupillary behavior of $\mathrm{Ms}_{\mathrm{matched}}$. It is yet to be determined whether object familiarity and unconscious object memory are supported by a common or distinct storage system. Future studies will need to employ neuroimaging and pupillometry with healthy and lesioned participants to address this issue.

References

Beatty, J., & Kahneman, D. (1966). Pupillary changes in two memory tasks. *Psychonomic Science*, *5* 371–372.

Berry, C. J., Shanks, D. R., Speekenbrink, M., & Henson, R. N. A. (2012). Models of recognition, repetition priming, and fluency: Exploring a new framework. *Psychological Review*, *119*, 40–79. doi: 10.1037/a0025464

Blaxton, T. A. (1989). Investigating dissociations among memory measures—Support for a transfer-appropriate processing framework. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 657–668.

Chanon, V. W., & Hopfinger, J. B. (2008). Memory's grip on attention: The influence of item memory on the allocation of attention. *Visual Cognition*, *16*, 325–340.

Daselaar, S. M., Fleck, M. S., & Cabeza, R. (2006). Triple dissociation in the medial temporal lobes: Recollection, familiarity, and novelty. *Journal of Neurophysiology*, *96*, 1902–1911. doi:10.1152/jn.01029.2005

Dennis, I., & Perfect, T. J. (2013). Do stimulus–action associations contribute to repetition priming? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*, 85–95. doi:10.1037/a0028479

Duss, S. B., Reber, T. P., Hänggi, J., Schwab, S., Wiest, R., Müri, R. M., … Henke, K. (2014). Unconscious relational encoding depends on hippocampus. *Brain*, *137*, 3355–3370.

Friedman, D. (2007). ERP studies of recognition memory: Differential effects of familiarity, recollection, and episodic priming. In M.-K. Sun (Ed.), *New research in cognitive sciences* (4th ed., p. 188). Hauppauge, NY: NOVA Science Publishers Inc.

Friedman, D., Hakerem, G., Sutton, S., & Fleiss, J. L. (1973). Effect of stimulus uncertainty on the pupillary dilation response and the vertex evoked potential. *Electroencephalography and Clinical Neurophysiology*, *34*, 475–484.

Gardner, R. M., Beltramo, J. S., & Krinsky, R. (1975). Pupillary changes during encoding, storage, and retrieval of information. *Perceptual and Motor Skills*, *41*, 951–955.

Gardner, R. M., Mo, S. S., & Borrego, R. (1974). Inhibition of pupillary orienting reflex by novelty in conjunction with recognition memory. *Bulletin of the Psychonomic Society*, *3*, 237–238.

Gardner, R. M., Philp, P., & Radacy, S. (1978). Pupillary changes during recall in children. *Journal of Experimental Child Psychology*, *25*, 168–172.

Goldinger, S. D., & Papesh, M. H. (2012). Pupil dilation reflects the creation and retrieval of memories. *Current Directions in Psychological Science*, *21*, 90–95. doi:10.1177/0963721412436811

Gomes, C. A., & Mayes, A. (2014). The kinds of information that support novel associative object priming and how these differ from those that support item priming. *Memory*. Advanced online publication. doi:10.1080/09658211.2014.937722

Granholm, E., Asarnow, R. F., Sarkin, A. J., & Dykes, K. L. (1996). Pupillary responses index cognitive resource limitations. *Psychophysiology*, *33*, 457–461.

Grison, S., Tipper, S. P., & Hewitt, O. (2005). Long-term negative priming: Support for retrieval of prior attentional processes. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *58A*, 1199–1224. doi:10.1080/02724980443000557

Hannula, D. E., Althoff, R. R., Warren, D. E., Riggs, L., Cohen, N. J., & Ryan, J. D. (2010). Worth a glance: Using eye movements to investigate the cognitive neuroscience of memory. *Frontiers in Human Neuroscience*, *4*, 166. doi:10.3389/fnhum.2010.00166

Harris, J. D., Cutmore, T. R. H., O'Gorman, J., Finnigan, S., & Shum, D. (2009). Neurophysiological indices of perceptual object priming in the absence of explicit recognition memory. *International Journal of Psychophysiology*, *71*, 132–141. doi:10.1016/j.ijpsycho.2008.08.005

Hollingworth, A., & Henderson, J. M. (2002). Accurate visual memory for previously attended objects in natural scenes. *Journal of Experimental Psychology: Human Perception and Performance*, *28*, 113–136.

Hollingworth, A., Williams, C. C., & Henderson, J. M. (2001). To see and remember: Visually specific information is retained in memory from previously attended objects in natural scenes. *Psychonomic Bulletin & Review*, *8*, 761–768.

Horner, A. J., & Henson, R. N. (2009). Bindings between stimuli and multiple response codes dominate long-lag repetition priming in speeded classification tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 757–779. doi:10.1037/a0015262

Horner, A. J., & Henson, R. N. (2011). Stimulus-response bindings code both abstract and specific representations of stimuli: Evidence from a classification priming design that reverses multiple levels of response representation. *Memory & Cognition*, *39*, 1457–1471. doi: 10.3758/s13421-011-0118-8

Horner, A. J., & Henson, R. N. (2012). Incongruent abstract stimulus-response bindings result in response interference: FMRI and EEG evidence from visual object classification priming. *Journal of Cognitive Neuroscience*, *24*, 760–773. doi:10.1162/jocn_a_00163

Just, M. A., & Carpenter, P. A. (1993). The intensity dimension of thought: Pupillometric indices of sentence processing. *Canadian Journal of Experimental Psychology*, *47*, 310–339.

Kafkas, A., & Montaldi, D. (2011). Recognition memory strength is predicted by pupillary responses at encoding while fixation patterns distinguish recollection from familiarity. *Quarterly Journal of Experimental Psychology*, *64*, 1971–1989. doi:10.1080/17470218.2011.588335

Kafkas, A., & Montaldi, D. (2012). Familiarity and recollection produce distinct eye movement, pupil and medial temporal lobe responses when memory strength is matched. *Neuropsychologia*, *50*, 3080–3093. doi:10.1016/j.neuropsychologia.2012.08.001

Kafkas, A., & Montaldi, D. (2014). Two separate, but interacting, neural systems for familiarity and novelty detection: A dual-route mechanism. *Hippocampus*, *24*, 516–527. doi:10.1002/hipo.22241

Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science*, *154*, 1583–1585.

Kirwan, C. B., Shrager, Y., & Squire, L. R. (2009). Medial temporal lobe activity can distinguish between old and new stimuli independently of overt behavioral choice. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 14617–14621. doi:10.1073/pnas.0907624106

Klingner, J., Tversky, B., & Hanrahan, P. (2011). Effects of visual and verbal presentation on cognitive load in vigilance, memory, and arithmetic tasks. *Psychophysiology*, *48*, 323–332. doi:10.1111/j.1469-8986.2010.01069.x

Kuchinke, L., Võ, M. L.-H., Hofmann, M., & Jacobs, A. M. (2007). Pupillary responses during lexical decisions vary with word frequency but not emotional valence. *International Journal of Psychophysiology*, *65*, 132–140. doi:10.1016/j.ijpsycho.2007.04.004

Kuipers, J. R., & Thierry, G. (2011). N400 amplitude reduction correlates with an increase in pupil size. *Frontiers in Human Neuroscience*, *5*, 61. doi:10.3389/fnhum.2011.00061

Laeng, B., Ørbo, M., Holmlund, T., & Miozzo, M. (2011). Pupillary Stroop effects. *Cognitive Processing*, *12*, 13–21. doi:10.1007/s10339-010-0370-z

Laeng, B., Sirois, S., & Gredebäck, G. (2012). Pupillometry: A window to the preconscious? *Perspectives on Psychological Science*, *7*, 18–27.

Laeng, B., Waterloo, K., Johnsen, S. H., Bakke, S. J., Låg, T., Simonsen, S. S., & Høgsaet, J. (2007). The eyes remember it: Oculography and pupillometry during recollection in three amnesic patients. *Journal of Cognitive Neuroscience*, *19*, 1888–1904. doi:10.1162/jocn.2007.19.11.1888

Mayes, A., Montaldi, D., & Migo, E. (2007). Associative memory and the medial temporal lobes. *Trends in Cognitive Sciences*, *11*, 126–135. doi:10.1016/j.tics.2006.12.003

Montaldi, D., Spencer, T. J., Roberts, N., & Mayes, A. R. (2006). The neural system that mediates familiarity memory. *Hippocampus*, *16*, 504–520. doi:10.1002/hipo.20178

Murphy, P. R., Robertson, I. H., Balsters, J. H., & O'Connell, R. G. (2011). Pupillometry and P3 index the locus coeruleus-noradrenergic arousal function in humans. *Psychophysiology*, *48*, 1532–1543. doi:10.1111/j.1469-8986.2011.01226.x

Nieuwenhuis, S., De Geus, E. J., & Aston-Jones, G. (2011). The anatomical and functional relationship between the P3 and autonomic components of the orienting response. *Psychophysiology*, *48*, 162–175. doi:10.1111/j.1469-8986.2010.01057.x

Otero, S. C., Weekes, B. S., & Hutton, S. B. (2011). Pupil size changes during recognition memory. *Psychophysiology*, *48*, 1346–1353. doi:10.1111/j.1469-8986.2011.01217.x

Papesh, M. H., Goldinger, S. D., & Hout, M. C. (2012). Memory strength and specificity revealed by pupillometry. *International Journal of Psychophysiology*, *83*, 56–64. doi:10.1016/j.ijpsycho.2011.10.002

Qiyuan, J., Richer, F., Wagoner, B. L., & Beatty, J. (1985). The pupil and stimulus probability. *Psychophysiology*, *22*, 530–534.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, *62*, 1457–1506. doi:10.1080/17470210902816461

Richardson, D. C., Dale, R., & Spivey, M. J. (2007). Eye movements in language and cognition: A brief introduction. In M. Gonzalez-Marquez, I. Mittelberg, S. Coulson, & M. J. Spivey (Eds.), *Methods in cognitive linguistics* (pp. 323–344). Amsterdam, The Netherlands: John Benjamins.

Roediger, H. L. (1990). Implicit memory. Retention without remembering. *American Psychologist*, *45*, 1043–1056.

Rugg, M. D., Mark, R. E., Walla, P., Schloerscheidt, A. M., Birch, C. S., & Allan, K. (1998). Dissociation of the neural correlates of implicit and explicit memory. *Nature*, *392*, 595–598. doi:10.1038/33396

Ryan, J. D., Althoff, R. R., Whitlow, S., & Cohen, N. J. (2000). Amnesia is a deficit in relational memory. *Psychological Science*, *11*, 454–461.

Schendan, H. E., & Kutas, M. (2003). Time course of processes and representations supporting visual object identification and memory. *Journal of Cognitive Neuroscience*, *15*, 111–135. doi:10.1162/089892903321107864

Spencer, T. J., Montaldi, D., Gong, Q.-Y., Roberts, N., & Mayes, A. R. (2009). Object priming and recognition memory: Dissociable effects in left frontal cortex at encoding. *Neuropsychologia*, *47*, 2942–2947. doi:10.1016/j.neuropsychologia.2009.06.022

Tanenhaus, M. K., Magnuson, J. S., Dahan, D., & Chambers, C. (2000). Eye movements and lexical access in spoken-language comprehension: Evaluating a linking hypothesis between fixations and linguistic processing. *Journal of Psycholinguistic Research*, *29*, 557–580.

Theeuwes, J., Kramer, A. F., Hahn, S., & Irwin, D. E. (1998). Our eyes do not always go where we want them to go: Capture of the eyes by new objects. *Psychological Science*, *9*, 379–385.

Theeuwes, J., Kramer, A. F., Hahn, S., Irwin, D. E., & Zelinsky, G. J. (1999). Influence of attentional capture on oculomotor control. *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 1595–1608.

Tressoldi, P. E., Martinelli, M., Semenzato, L., & Cappato, S. (2011). Let your eyes predict: Prediction accuracy of pupillary responses to random alerting and neutral sounds. *SAGE Open*. doi:10.1177/2158244011420451

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology*, *26*, 1–12.

Tulving, E., & Schacter, D. L. (1990). Priming and human memory systems. *Science*, *247*, 301–306.

Võ, M. L.-H., Jacobs, A. M., Kuchinke, L., Hofmann, M., Conrad, M., Schacht, A., & Hutzler, F. (2008). The coupling of emotion and cognition in the eye: Introducing the pupil old/new effect. *Psychophysiology*, *45*, 130–140. doi:10.1111/j.1469-8986.2007.00606.x

Voss, J. L., & Gonsalves, B. D. (2010). Time to go our separate ways: Opposite effects of study duration on priming and recognition reveal distinct neural substrates. *Frontiers in Human Neuroscience*, *4*, 227. doi:10.3389/fnhum.2010.00227

Voss, J. L., Lucas, H. D., & Paller, K. A. (2010). Conceptual priming and familiarity: Different expressions of memory during recognition testing with distinct neurophysiological correlates. *Journal of Cognitive Neuroscience*, *22*, 2638–2651. doi:10.1162/jocn.2009.21341

Voss, J. L., & Paller, K. A. (2006). Fluent conceptual processing and explicit memory for faces are electrophysiologically distinct. *Journal of Neuroscience*, *26*, 926–933. doi:10.1523/JNEUROSCI.3931-05.2006

Voss, J. L., & Paller, K. A. (2007). Neural correlates of conceptual implicit memory and their contamination of putative neural correlates of explicit memory. *Learning & Memory*, *14*, 259–267. doi:10.1101/lm.529807

Voss, J. L., & Paller, K. A. (2010). Real-time neural signals of perceptual priming with unfamiliar geometric shapes. *Journal of Neuroscience*, *30*, 9181–9188. doi:10.1523/JNEUROSCI.0403-10.2010

Voss, J. L., Reber, P. J., Mesulam, M.-M., Parrish, T. B., & Paller, K. A. (2008). Familiarity and conceptual priming engage distinct cortical networks. *Cerebral Cortex*, *18*, 1712–1719. doi:10.1093/cercor/bhm200

Voss, J. L., Schendan, H. E., & Paller, K. A. (2010). Finding meaning in novel geometric shapes influences electrophysiological correlates of repetition and dissociates perceptual and conceptual priming. *NeuroImage*, *49*, 2879–2889. doi:10.1016/j.neuroimage.2009.09.012

Walla, P., Endl, W., Lindinger, G., Deecke , & Lang, W. (1999). Implicit memory within a word recognition task: An event-related potential study in human subjects. *Neuroscience Letters*, *269*, 129–132.

Weiskrantz, L. (1998). Pupillary responses with and without awareness in blindsight. *Consciousness and Cognition*, *7*, 324–326. doi:10.1006/ccog.1998.0360

Yonelinas, A. P. (1994). Receiver-operating characteristics in recognition memory: Evidence for a dual-process model. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 1341–1354. doi:10.1037/0278-7393.20.6.1341

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, *46* 441–517.

(Received July 7, 2014; Accepted December 13, 2014)

Appendix

Recognition Memory Test Instructions

Recollection memory

Sometimes when you are shown previously studied stimuli, you will be able to recollect something specific about encountering them in the study session. In order to be recollecting, you must be retrieving something that isn't the stimulus that is in front of you or even part of it. Rather, you must be recollecting something additional to the stimulus itself, and this involves bringing to mind information that is not currently in front of you. The type of information you might recollect includes what you were thinking when you encountered the stimulus during the study session. For example, you might have thought that a studied face resembled someone you know or that an object was particularly beautiful, and this comes back to you when you see the face or object again during the memory test. You might also recollect information that you might not have explicitly thought about during study, but also isn't part of the stimulus. For example, you might recall that a stimulus came very early in the list or that you sneezed when the stimulus was originally presented.

However, if when your memory is tested you just remember seeing a stimulus or even part of a stimulus in the study session, but there is no other more specific extra information that you recall about the study session, then you should say that the stimulus is familiar (because you are recalling nothing about it that is specific to the study episode).

Familiarity memory

Everyone has a pretty good idea of what it is like to find a stimulus familiar without being able to recall anything about it. For example, we have all been in the situation where we have met someone whom we recognize as familiar, although we can't recall anything about them such as their name, where we last saw them, or why we know them.

You can have the same feeling of being sure that you encountered a stimulus in a specific context (such as a study session in a testing room) although you are unable to recall anything extra about what you thought or what else happened in the room when you encountered the stimulus. Sometimes the stimulus can feel very familiar from the context and sometimes only weakly familiar.

We want you to be very clear about this idea of familiarity so that you will be able to decide which stimuli you have studied before because they feel familiar to you even though you have recalled nothing specific about them. The distinction between familiarity and recollection is very important in this experiment, and it should not be confused with your feelings of confidence. You can be very confident that you have encountered something before, independent of whether you find it familiar or you recollect something about it.

In this experiment, I would like you to try to only use familiarity and not to try to recollect any specific details about stimuli that you think may have been presented earlier in the study. If you inadvertently recollect some specific details about the previous item, then indicate this using the appropriate key.