BRIEF REPORT

The road to long-term memory: Top-down attention is more effective than bottom-up attention for forming long-term memories

Edyta Sasin [1] · Daryl Fougnie [1]

Abstract

Does the strength of representations in long-term memory (LTM) depend on which type of attention is engaged? We tested participants' memory for objects seen during visual search. We compared implicit memory for two types of objects—related-context nontargets that grabbed attention because they matched the target defining feature (i.e., color; top-down attention) and salient distractors that captured attention only because they were perceptually distracting (bottom-up attention). In Experiment 1, the salient distractor flickered, while in Experiment 2, the luminance of the salient distractor was alternated. Critically, salient and related-context nontargets produced equivalent attentional capture, yet related-context nontargets were remembered far better than salient distractors (and salient distractors were not remembered better than unrelated distractors). These results suggest that LTM depends not only on the amount of attention but also on the type of attention. Specifically, top-down attention is more effective in promoting the formation of memory traces than bottom-up attention.

Keywords Top-down attention · Bottom-up attention · Attentional capture · Long-term memory

We encounter millions of objects every day. While our ability to retain some of these objects in visual long-term memory (VLTM; a large-capacity, passive storage system for visual episodic memories) is surprisingly high and detailed (Brady, Konkle, Alvarez, & Oliva, 2008; Konkle, Brady, Alvarez, & Oliva, 2010; Shepard, 1967; Standing, Conezio, & Haber, 1970; Vogt & Magnussen, 2007), many of these objects are either not encoded or forgotten from memory (Lew, Pashler, & Vul, 2016; Mercer & Jones, 2019). What factors determine whether an item will or will not be stored in memory? On the one hand, previous research has identified many factors that influence whether something will be encoded in long-term memory (LTM). For example, stimuli are more likely to be remembered if they are repeated (Williams, 2010b), are processed deeply (Craik & Lockhart, 1972), carry emotional or personal importance (Hamann, 2001; Kensinger, Garoff-Eaton, & Schacter, 2007; Loftus, Loftus, & Messo, 1987),

appear together with a target of another task (attentional boost effect; Swallow & Jiang, 2010, 2013), or are salient (Celikkale, Erdem, & Erdem, 2015). On the other hand, attentional lapses during encoding (deBettencourt, Norman, & Turk-Browne, 2018) or demanding concurrent tasks (Evans & Baddeley, 2018) can lead to worse memory. However, there is little research exploring how important is visual attention for successful encoding into VLTM. In one study, it was found that targets of a visual search task showed better memory performance than distractors in a surprise recognition test, suggesting that the increased attention given to targets was important for successful memory encoding (Williams, Henderson, & Zacks, 2005). However, the influence of attention on VLTM is certainly more complex than that captured by these findings.

The paucity of work relating VLTM and attention is quite surprising, given the extensive work aimed towards understanding attention and WM. Visual attention is known to have a critical role for encoding into visual WM. For example, stimuli that appear outside of attention often go undetected, as illustrated by research on change detection (Hollingworth, 2004; Rensink, 2002; Simons & Rensink, 2005), the attentional blink (Raymond, Shapiro, & Arnell, 1992; Vogel, Luck, & Shapiro, 1998), or

✉ Edyta Sasin
edytasasin@gmail.com

[1] Department of Psychology, New York University Abu Dhabi, Abu Dhabi, United Arab Emirates

inattentional blindness (Nakayama, Deutsch, & Nakayama, 1999; Simons & Chabris, 1999).

The goal of the present work is to further understand the role of attention in successful VLTM performance. In particular, attention is not a unitary construct. There is clear behavioral and neural evidence for separate attentional systems for purposefully attending to something (top-down) versus attending to salient parts of the environment (bottom-up) (Awh, Belopolsky, & Theeuwes, 2012; Connor, Egeth, & Yantis, 2004; Corbetta & Shulman, 2002; Pinto, Leij, Sligte, Lamme, & Scholte, 2013; Theeuwes, 2010). Top-down attention is shifted voluntarily, according to the current goals of the observer. Bottom-up attention, on the other hand, is captured in a stimulus-driven manner, by stimuli that differ significantly from surrounding inputs (Awh et al., 2012; Corbetta & Shulman, 2002; Egeth & Yantis, 1997; Theeuwes, 2010).

Given that attention is composed of two (at least partially) distinct mechanisms, does the benefit of attention for memory depend on which form is engaged? Here we examined the implicit memory of objects presented during a visual search task while manipulating the type of attention. More specifically, we compared the implicit memory of two types of objects—related-context nontargets that grabbed attention because they matched a target feature (top-down attention) and salient distractors that captured attention only because they were perceptually distracting (bottom-up attention). Note that capture by an object that shares a feature with a target held in memory is operationalized as top-down capture rather than priming effect (i.e., the facilitation of the processing of a stimulus due to the prior presentation of a stimulus that is perceptually or semantically related; Kristjánsson & Campana, 2010). Such a distinction is also consistent with studies showing that recent exposure to an object is insufficient to elicit capture by matching distractors, and that only representations held in WM might guide attention. (Olivers, Meijer, & Theeuwes, 2006; Soto, Heinke, Humphreys, & Blanco, 2005; Soto, Humphreys, & Rotshtein, 2007).

The magnitude of attentional capture was estimated from reaction times in the search task. The amount of capture is typically used to infer the amount of attention allocated to the distractors (Folk & Remington, 2008; Olivers, 2009; Olivers et al., 2006; Posner, 1980; van Moorselaar, Battistoni, Theeuwes, & Olivers, 2015; Yantis & Hillstrom, 1994). It should be pointed out that the amount of attention might be a sum of two components of attentional capture: the time of attentional focus on an object and the number of shifts of attention towards the object. Better VLTM performance for one of these distractors would suggest that encoding into VLTM depends on the type of attention that is engaged.

Experiment 1

Method

Participants According to pilot studies, in which we found an effect size of $\eta_p^2 = .35$, $\alpha = .05$, and power = 0.95, the sample size of a minimum 17 was required to find a significant effect in the memory performance with 95% probability, if the effect exists. Seventeen students and staff of the New York University of Abu Dhabi (12 males; $M = 26$ years, $SD = 7.27$) participated in the experiment in exchange for course credit, or alternatively received a subsistence allowance of 50 AED per hour. All participants had normal or corrected-to-normal visual acuity and gave informed consent. The experiments were approved by the New York University Abu Dhabi Institutional Review Board.

Apparatus and stimuli Stimuli were presented using Psychtoolbox for MATLAB (Brainard, 1997), and the experiments were run on computers that were fitted with a 22-inch BenQ XL2411 monitor (144 Hz refresh rate, 1,920 × 1,080 pixels). All stimuli were presented on a black background at a viewing distance of 57 cm. The stimulus set consisted of 540 pictures of categorically distinct objects pulled from the Brady, Konkle, Gill, Oliva, and Alvarez (2013) data set. Twenty-four of these pictures were used only in the practice block. Three hundred sixty pictures were used in a search task (90 as targets, 30 as salient distractors, 30 as related-context nontargets, and 210 as distractors), and 90 pictures were used only in the surprise memory test as the novel objects. The pictures were assigned to conditions randomly on a per participant basis. Each picture was fit to a square of 100 × 100 pixels (2.92° × 2.92°). Importantly, each image had an object defined by a single, dominant color (e.g., a *blue* couch). Color was used to define the search target. At the beginning of each trial, participants were given a cue—a centrally presented colored circle (radius 0.90° visual angle) to indicate the target color. The target colors were selected randomly from a set of four possible colors that were created by manipulating the color of the target image. Specifically, the color was the target image's dominant color shifted by 0°, 90°, 180° or 270° (Brady et al., 2013) in hue space using the LAB circular color space. Critically, there were three types of trials in the search task. On *neutral* trials, the search display contained the target and three distractors. On *salient distractor* trials, the search display contained the target, a salient distractor, and two distractors. A salient distractor is defined as a distractor with a color unrelated to the target item, but which has increased bottom-up salience because it flickered on the screen (the other items were presented without flicker). The salient distractor flickered rapidly (appear and disappear) at frequency rates chosen randomly during a trial from the frequencies between 0.3 and 0.9 Hz. On *related-context nontarget trials*,

the search display contained the target, related-context nontarget, and two distractors. A related-context nontarget is one whose color was similar to, but not exactly the same as, the target color (i.e., shifted by 30° in hue space from the hue of the target). Note that this distractor was related to the target by color, the defining feature of the search task, but the object identity and locations were completely independent. The colors of the unrelated distractors or salient distractor were chosen randomly from the set of four colors, excluding the target color (e.g., if the target image's color was shifted by 90° in hue, other images could have colors shifted by 0°, 180°, or 270° in hue). Both salient distractors and related-context nontargets were never targets. Critically, the related-context nontarget and salient distractor conditions were found to be equally distracting in pilot experiments.

The search display was composed of the four different objects located equidistantly on an imaginary circle of radius 4.38° around fixation, with locations determined randomly on a per trial basis. The items were at 45°, 135°, 225°, and 315° positions. The boundaries of each object were separated by at least 1.46° of visual angle.

Procedure The experimental procedure is illustrated in Fig. 1. At the beginning of each trial, participants were presented with a target color for 1,000 ms, which was followed by a 500-ms blank interval. Subsequently, the search display appeared on the screen. The task participants were given was to localize the object with a color matching the target color.

To respond, participants indicated the search target location by pressing one of the four keys ("A," "K," "Z," "M") that corresponded to the location on the screen. Participants were instructed to make the search responses as quickly as possible. The search display remained on the screen until a response was made or until the maximum presentation time of 2 s was reached. (only 0.4 % of displays reached the 2-s presentation time). Participants were required to make a response to continue to a new trial even on trials in which when the search display was removed after 2 s. All pictures (90 targets, 30 salient distractors, 30 related-context nontargets, and 210 as unrelated distractors, chosen randomly for each participant) were repeated four times during the search session. Thus, there were 120 trials for each search condition, resulting in 360 trials in total. The conditions were intermixed and were
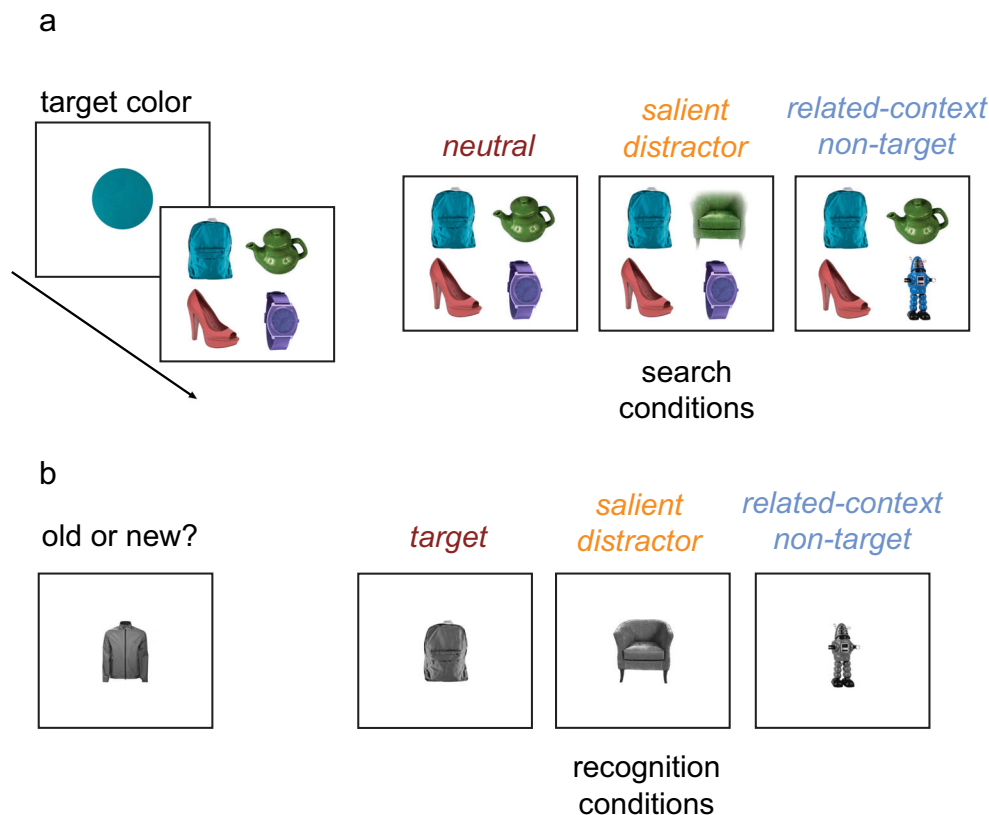

Fig. 1 a The structure of the search task. Participants searched for an object of a specific color that changed on a per-trial basis. In *neutral* conditions, the search display contained the target and unrelated distractors. In *related-context nontarget* conditions, one of the distractors has a color similar to, but not the same as, the target color (in this example, it is a blue robot). In *salient distractor* conditions, one of the distractors flickered rapidly at a random frequency (in this example, it is a green armchair). b The structure of the recognition test. Participants were asked to indicate whether an object on the screen had been shown earlier on in the experiment. Memory was separately assessed for search targets, salient distractors, and related-context nontargets

presented in random order. The experiment was preceded by six practice trials to give participants familiarity with the task. The experimental phase was followed by an unexpected recognition test that required participants to indicate whether an object on the screen had been shown earlier on in the experiment. The recognition test included 30 targets from the neutral search condition, 30 salient distractors, 30 related-context nontargets, and 90 novel objects. Importantly, the novel objects were not presented during the experiment. Participants were instructed to respond by pressing "Z" when an object was identified as "old" (this was considered correct for the target and distractors) and "M" when the object was identified as "new" (this was considered correct for novel objects). These objects were shown randomly, one at a time. The color of each object presented in the recognition test was converted to grayscale. Participants were told to respond accurately (speed was not stressed, and the object remained in view until the judgment was made).

Results

Correct search trials made up 88% of the data (91% in both the neutral condition, 92% in the salient distractor condition, and 82% in the related-context nontarget condition). However, due to human error, the accuracy of two trials from each condition could not be determined, which lead to an overall slightly lower accuracy than would be expected in this task. Before analyzing the reaction times (RTs) for the search task, we excluded the trials with incorrect responses in the search task. Next, we excluded trials with search-RTs shorter than 150 ms or longer than 3,000 ms and trials with search RTs above a cutoff value of three standard deviations from the mean. This procedure resulted in a loss of 2.59% of the data points. Importantly, none of the qualitative conclusions are altered by excluding the aforementioned trials. Search reaction times and sensitivity are illustrated in Fig. 2.

We performed a repeated-measures analysis of variance (ANOVA) on search RTs with search condition as a factor (*neutral* vs. *salient distractor* vs. *related-context nontarget*). This showed a significant effect of search type, $F(2, 32) = 43.63$, $p < .001$, $\eta_p^2 = .73$. Planned contrasts revealed that salient distractor trials (695 ms, 95% CI [684, 705]) resulted in slower RTs relative to neutral trials (602 ms, 95% CI [594, 611]), $t(16) = 12.67$, $p < .001$, $d = 3.07$. Similarly, related-context nontarget trials (708 ms, 95% CI [695, 721]) resulted in slower RTs relative to neutral trials, $t(16) = 8.58$, $p < .001$, $d = 2.08$. There was no significant difference between salient distractor trials and related-context nontarget trials, $t(16) = 0.85$, $p = .406$. Further, there is no evidence that capture by the salient distractor was initially strong and then had a reduced effect in later trials (see Supplemental Material).

The current results demonstrate not only that both types of distractors captured attention but also suggests that the amount of attention afforded to each distractor was equivalent for the two conditions. Given the evidence of *equal* attentional capture for both distractor conditions, we can examine performance on the recognition task as a function of the *type* of attention.

Performance on the surprise recognition task was also examined. To ensure that analysis of memory was done only on trials where participants successfully identified the target, the recognition test analyses were restricted to stimuli from trials with correct search responses. To measure the degree to which objects were encoded into memory, we calculated a sensitivity index (d'), a signal detection measure conceptualized as the distance between signal and noise distributions[1] (Stanislaw & Todorov, 1999). A one-tailed one-sample *t* test was performed on sensitivity indices d' for each memory condition: targets, salient distractors, and related-context nontargets. The one-tailed *t*-test evaluations showed that in each memory condition, d' was significantly above zero (all ts > 4.13, all ps < .001), indicating that participants were able to remember something about the items even though they were not informed about the test in advance. Next, sensitivity indices d' were entered in an ANOVA, with memory condition as a factor. This analysis revealed significant effect of memory condition, $F(2, 32) = 10.81$, $p < .001$, $\eta_p^2 = .40$. A series of post hoc comparisons were then performed using the Holm–Bonferroni correction. These comparisons showed that the memory performance was better for the targets ($d' = 1.00$, 95% CI [0.78, 1.22]) than the salient distractors ($d' = 0.37$, 95% CI [0.16, 0.59]), $t(16) = 4.08$, $p = .003$, $d = 0.99$. The difference in memory performance between the targets and the related-context nontargets ($d' = 0.75$, 95% CI [0.58, 0.92]) did not quite reach significance, $t(16) = 1.95$, $p = .07$, $d = 0.47$. Importantly, the memory performance was better for the related-context nontargets than for the salient distractors, $t(16) = 3.07$, $p = .015$, $d = 0.75$, suggesting that the attentional capture by those distractors was more successful in encoding/storage into LTM.[2]

First, it is important to highlight that analysis of memory performance in Experiment 1 showed that search targets, related-context nontargets, and salient distractors were indeed

---

[1] Specifically, d' was calculated as a difference between *z*-transforms of hit rate (i.e. the proportion of "old" trials to which subject responded "old") and false alarm rate (i.e., the proportion of "new" trials to which subject responded "old"), where *z*-transforms are inverse of the normal distribution cumulative density function. Note that the false-alarm rate for each participant was calculated based on incorrect responses to all "new" objects, and thus it was the same for all conditions.

[2] Note that this analysis assumes that differences in hit rate arose because of differences in memory signal, not decision criterion, which is reasonable given the intermixed nature of our design. However, to further rule out differences in decision criterion, we conducted an additional analysis that fixed criterion across all memory conditions (see Supplemental Material).
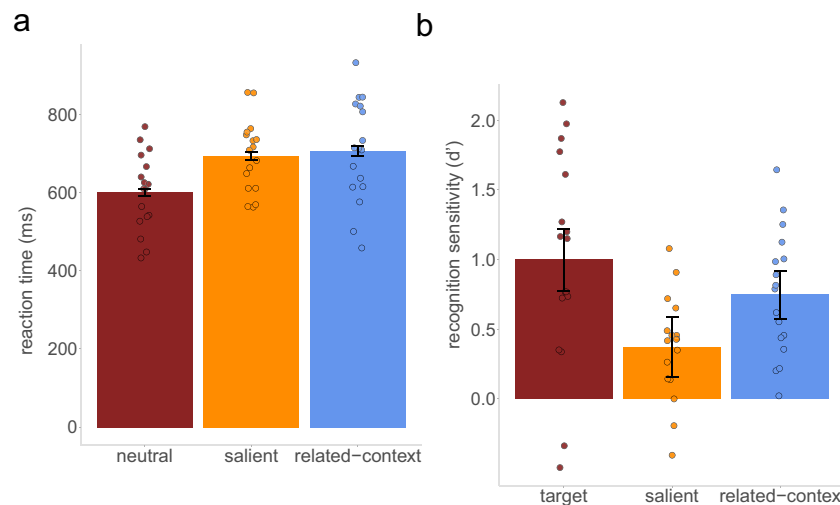
Fig. 2 The results of Experiment 1. **a** Mean RTs (ms) in the search task as a function of the search condition. **b** Mean sensitivity index (d') as a function of recognition conditions. Error bars reflect 95% within-subjects confidence intervals (Cousineau, 2005; Morey, 2008)

encoded into LTM even if participants were not asked to memorize these objects. Such incidental encoding led to overall small sensitivity indices, which was expected, taking into account that executing the search task does not require any identification of the presented objects. Moreover, objects in the recognition test were grayscale versions of colored objects presented during the search task, which might also contribute to overall small sensitivity indices, according to encoding specificity (Tulving & Thomson, 1973). Crucially, the results showed the same magnitude of attentional capture produced by related-context nontargets (top-down capture) and salient distractors (bottom-up capture). However, the surprise recognition test revealed that the memory performance was much better for the related-context nontargets than the salient distractors.

Search targets were better remembered than the other stimuli (although the difference in memory between targets and related-context nontargets did not reach significance level, perhaps due to insufficient statistical power). This replicates results from previous studies (Tatler & Tatler, 2013; Williams, 2010a, 2010b; Williams et al., 2005) and suggests that objects that are encoded as targets are better encoded into LTM.

Experiment 2

The critical finding of Experiment 1 is that salient distractors were remembered less than related-context nontargets, suggesting that bottom-up attention is less efficient as a means to VLTM encoding compared with top-down attention. A possible criticism of the current design is that our method of inducing bottom-up saliency involved flickering the item, which means that the item was on the screen for a reduced amount of time. The goal of Experiment 2 is to replicate the

findings of Experiment 1 and to test whether the differences in memory performance that we observed between salient distractors and related-context nontargets could be due to the nature of the flicker rather than the difference in the type of attention that is engaged. To control for this possibility, in Experiment 2, we used a different form of increasing stimulus salience. Specifically, we alternated stimulus luminance of the salient distractor. An additional goal of Experiment 2 was to measure memory performance of the unrelated distractor items to provide a baseline of distractor memory to compare to salient distractor memory.

Method

Participants Seventeen students of the New York University of Abu Dhabi (eight males; $M = 20.3$ years, $SD = 1.28$) participated in the experiment in exchange for course credit or received a subsistence allowance of 50 AED per hour. All participants had normal or corrected-to-normal visual acuity and gave informed consent. The experiments were approved by the New York University Abu Dhabi Institutional Review Board.

Apparatus and stimuli The stimuli were identical to the ones used in Experiment 1, except for the following changes. All stimuli were presented on a white background. The luminance of the salient distractor was changed during a trial at a random frequency between 0.3 and 0.9 Hz. The luminance of the salient distractor alternated between original luminance and increased luminance (specifically, the L value of the distractor color, in CIE LAB space was increased by 60). The recognition test additionally included 30 old or new judgments on images that were used as unrelated distractors. There was no
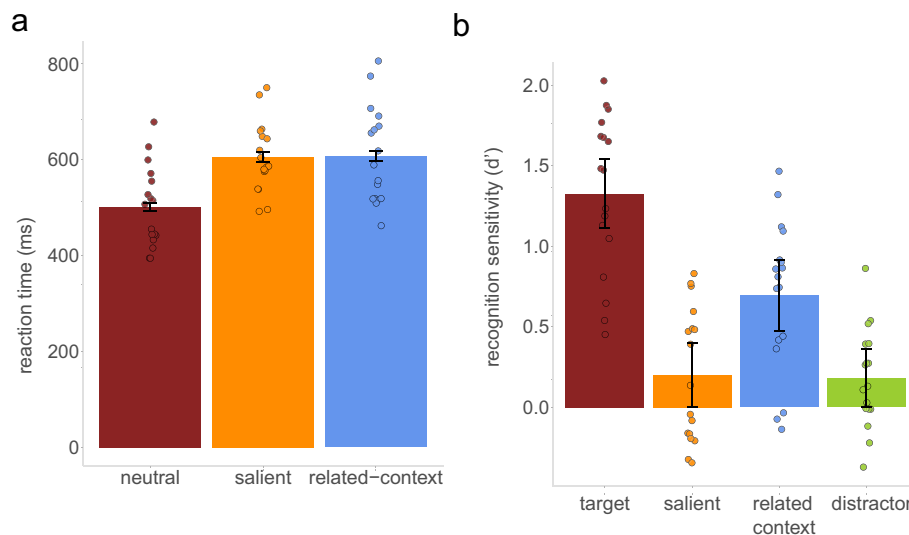
a

b

Fig. 3 The results of Experiment 2. **a** Mean RTs (ms) in the search task as a function of the search condition. **b** Mean sensitivity index (*d*′) as a function of recognition conditions. Error bars reflect 95% within-subjects confidence intervals (Cousineau, 2005; Morey, 2008)

increase in the number of "new" images in the recognition test.

Results

Participant's performance on the search task was 93% (95% in both the neutral and the salient distractor conditions and 87% in the related-context nontarget condition). Before analyzing RTs for the search task, we excluded trials with incorrect search responses, RTs less than 150 ms or greater than 3,000 ms, or RTs greater than three standard deviations from that participant mean (resulting in a loss of 2.42% of data points). Importantly, the qualitative conclusions remain the same if the aforementioned trials are not excluded. An ANOVA was performed on search RTs, with search condition as a factor (*neutral* vs. *salient distractor* vs. *related-context nontarget*). The results, illustrated in Fig. 3, showed significant effect of search type, $F(2, 32) = 45.20$, $p < .001$, $\eta_p^2 = .74$. We found that RTs were longer for both the salient distractor condition (610 ms, 95% CI [599, 621]), $t(16) = 9.98$, $p < .001$, $d = 2.42$, and the related-context nontarget condition (612 ms, 95% CI [601, 23]), $t(16) = 7.87$, $p < .001$, $d = 1.91$, relative to the neutral condition (505 ms, 95% CI [497, 513]). There was no difference in RTs between the salient distractor condition and the related-context nontarget condition, $t(16) = 0.11$, $p = .914$. Again, the results demonstrate that there was no noticeable difference in the amount of captured attention between the distractors.

We also analyzed recognition test data, restricted to stimuli from trials where participants successfully identified the target. A one-tailed one-sample *t* test on sensitivity indices *d*′ again showed that in each memory condition *d*′ was significantly above zero (all *t*s > 1.96, all *p*s < .032), indicating that

memory performances were significantly above chance, even for the unrelated distractors. The ANOVA on sensitivity indices *d*′ further revealed significant effect of memory condition, $F(3, 48) = 31.64$, $p < .001$, $\eta_p^2 = .66$. The post hoc comparisons with Holm–Bonferroni correction showed that the targets (*d*′ = 1.33, 95% CI [1.11, 1.54]) were remembered better than the salient distractors (*d*′ = 0.20, 95% CI [0.00, 0.40]), $t(16) = 8.62$, $p < .001$, $d = 2.09$, the related-context nontargets (*d*′ = 0.69, 95% CI [0.47, 0.91]), $t(16) = 4.61$, $p = .001$, $d = 1.12$, and unrelated distractors (*d*′ = 0.18, 95% CI [0.00, 0.36]), $t(14) = 7.68$, $p < .001$, $d = 1.86$. There was not a significant difference in the memory performance between the salient distractors and unrelated distractors, $t(16) = 0.19$, $p = .849$. Most critically, we replicated the finding that the related-context nontargets were remembered better than the salient distractors, $t(16) = 3.11$, $p = .013$, $d = 0.76$. This again demonstrates that all forms of attention are not equal in terms of successful LTM performance.[3]

In Experiment 2, we replicated the pattern of the results showing that, despite the same costs in the search performance engendered by related-context nontargets and salient distractors, the latter one resulted in much weaker long-term memories. In fact, the memory for the salient distractors did not differ from the memory for other distractors.

General discussion

The human sensory system is constantly bombarded by an enormous amount of information from the external world.

---

[3] We again conducted an additional analysis that fixed criterion across all memory conditions (see Supplemental Material), to ensure that our findings are not affected by differences in bias.

Which pieces of this information will be retained in memory, and why? While this topic has received considerable focus, one area that is poorly understood is the role that visual attention plays in successful VLTM. To shed light on this question, the present work tested memory for distractors seen during a visual search task, while searching for a target object of a particular color. In one condition (related-context nontarget), we made a distractor item likely to receive top-down attention by making it have a feature (i.e., color) similar to the to-be-searched feature. In another condition (salient distractor), we introduced an item that was irrelevant to the search task, but that would attract bottom-up attention due to its stimulus salience. In Experiment 1, the distracting stimulus was flickering at a rapid rate during the search task. In Experiment 2, the flickering was replaced by alternations in luminance to generalize over distinct methods of introducing stimulus salience. As expected, the presentation of both the related-context nontarget and salient distractor led to slower search compared with the baseline condition, in which none of these distractors was presented. Additionally, the amount of distraction did not differ between the distraction conditions. Importantly, even though both distractors produced the same magnitude of attentional capture, the related-context nontargets were remembered better than the salient distractors, according to a surprise VLTM test administered at the end of the study. This provides the first direct evidence that the objects that grab top-down attention are more likely to be encoded in VLTM than the objects that grab bottom-up attention.

What is the mechanism by which purposeful attention leads to more effective encoding into VLTM? It is possible that different types of attention lead to the attention being paid to different features. Perhaps in the case of bottom-up capture, attention was more focused on the distracting property itself (e.g., flicker on Experiment 1) rather than on the identity of the salient object. Indeed, recent studies have shown that when attention is focused on the task-relevant attribute of the object, the other attribute of this object is not necessarily consolidated into memory (Chen, Swan, & Wyble, 2016; Chen & Wyble, 2015). Similarly, it is likely that attention was more focused on color when the related-context nontarget was presented. But perhaps when attention was focused on a salient feature, there is less attention left over to process other image attributes (e.g., identity) then when attention was being focused by color. Importantly, it should be noted that the task-relevant feature, color, was equally irrelevant for the LTM test (which was conducted on grayscale images) as the saliency manipulations.

Another possible explanation is that purposeful attention is useful for VLTM encoding because attending is more effortful when it occurs via top-down means than when it occurs due to bottom-up salience. One of the most influential findings in LTM research is the levels of processing effect (Craik & Lockhart, 1972; Schulman, 1971), which demonstrates that effortful processing of to-be-remembered items leads to better LTM encoding than shallow processing. This framework defines "effortful" as requiring semantic processing. Although our search task did not require any conceptual assessment, it is possible that the attention grabbed by distractors that share similar traits with the target, while equal with salient distractors in terms of negative effects on the search task, leads to something akin to more "effortful" processing.

The work also has practical relevance, particularly for advertisers and anyone interested in creating a lasting impression on the human mind. It appears that the effects of salience, while strong, are short-lived and do not lead to strong encoding in VLTM. In contrast, when attention is purposefully shifted towards information, it is retained longer, even in cases where there is no explicit requirement to memorize that information. Perhaps this is good news for those of us who are sick of flashing and distracting images on the television or websites. This strategy may be useful for catching our attention in the moment, but may not ultimately be an effective advertising technique if capturing our attention in this way does not lead to encoding into VLTM. Indeed, there is evidence showing that animated advertisements attract less attention than static ones (Lee & Ahn, 2012).

In conclusion, the current study suggests that the formation of visual long-term memories not only depends on the amount of attention but also the type of attention that is engaged. Specifically, even though the magnitude of bottom-up capture by salient distractors was the same as top-down capture by related-context nontargets, the related-context nontargets were remembered better than the salient distractors. While future studies are needed to clarify how exactly attention is distributed between targets and different type of distractors, the present data provide evidence that the road to long-term memory may take many paths, but that purposeful attention provides the quickest route.

References

Awh, E., Belopolsky, A. V., & Theeuwes, J. (2012). Top-down versus bottom-up attentional control: A failed theoretical dichotomy. *Trends in Cognitive Sciences*, *16*(8), 437–443. https://doi.org/10.1016/j.tics.2012.06.010

Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(38), 14325–14329. https://doi.org/10.1073/PNAS.0803390105

Brady, T. F., Konkle, T., Gill, J., Oliva, A., & Alvarez, G. A. (2013). Visual long-term memory has the same limit on fidelity as visual working memory. *Psychological Science*, *24*(6), 981–990. https://doi.org/10.1177/0956797612465439

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10,* 433-436.

Celikkale, B., Erdem, A., & Erdem, E. (2015). Predicting memorability of images using attention-driven spatial pooling and image semantics. *Image and Vision Computing*, *42*, 35–46. https://doi.org/10.1016/J.IMAVIS.2015.07.001

Chen, H., Swan, G., & Wyble, B. (2016). *Prolonged focal attention without binding: Tracking a ball for half a minute without remembering its color. Cognition, 147*. https://doi.org/10.1016/j.cognition.2015.11.014

Chen, H., & Wyble, B. (2015). Amnesia for object attributes: Failure to report attended information that had just reached conscious awareness. *Psychological Science*, *26*(2), 203–210. https://doi.org/10.1177/0956797614560648

Connor, C. E., Egeth, H. E., & Yantis, S. (2004). Visual attention: Bottom-up versus top-down. *Current Biology*, *14*(19), R850–R852. https://doi.org/10.1016/J.CUB.2004.09.041

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, *3*(3), 215–229. https://doi.org/10.1038/nrn755

Cousineau, D. (2005). Confidence intervals in within-subject designs : A simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology*, (February), 1–5. https://doi.org/10.20982/tqmp.01.1.p042

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*(6), 671–684. https://doi.org/10.1016/S0022-5371(72)80001-X

deBettencourt, M. T., Norman, K. A., & Turk-Browne, N. B. (2018). Forgetting from lapses of sustained attention. *Psychonomic Bulletin & Review*, *25*(2), 605–611. https://doi.org/10.3758/s13423-017-1309-5

Egeth, H. E., & Yantis, S. (1997). Visual attention: Control, representation, and time course. *Annual Review of Psychology*, *48*(1), 269–297. https://doi.org/10.1146/annurev.psych.48.1.269

Evans, K. K., & Baddeley, A. (2018). Intention, attention and long-term memory for visual scenes: It all depends on the scenes. *Cognition*, *180*, 24–37. https://doi.org/10.1016/J.COGNITION.2018.06.022

Folk, C. L., & Remington, R. W. (2008). Bottom-up priming of top-down attentional control settings control settings. *Visual Cognition, 16*(2/3), 215–231. https://doi.org/10.1080/13506280701458804

Hamann, S. (2001). Cognitive and neural mechanisms of emotional memory. *Trends in Cognitive Sciences*, *5*(9), 394–400. https://doi.org/10.1016/S1364-6613(00)01707-1

Hollingworth, A. (2004). Constructing visual representations of natural scenes: The roles of short- and long-term visual memory. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(3), 519–537. https://doi.org/10.1037/0096-1523.30.3.519

Kensinger, E. A., Garoff-Eaton, R. J., & Schacter, D. L. (2007). Effects of emotion on memory specificity: Memory trade-offs elicited by negative visually arousing stimuli. *Journal of Memory and Language*, *56*(4), 575–591. https://doi.org/10.1016/J.JML.2006.05.004

Konkle, T., Brady, T. F., Alvarez, G. A., & Oliva, A. (2010). Scene memory is more detailed than you think: The role of categories in visual long-term memory. *Psychological Science*, *21*(11), 1551–1556. https://doi.org/10.1177/0956797610385359

Kristjánsson, Á., & Campana, G. (2010). Where perception meets memory: A review of repetition priming in visual search tasks. *Attention, Perception, & Psychophysics*, *72*(1), 5–18. https://doi.org/10.3758/APP.72.1.5

Lee, J., & Ahn, J. H. (2012). Attention to banner ads and their effectiveness: An eye-tracking approach. *International Journal of Electronic Commerce, 17*(1), 119-137. https://doi.org/10.2753/JEC1086-4415170105

Lew, T. F., Pashler, H. E., & Vul, E. (2016). Fragile associations coexist with robust memories for precise details in long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(3), 379–393. https://doi.org/10.1037/xlm0000178

Loftus, E. F., Loftus, G. R., & Messo, J. (1987). Some facts about "weapon focus." *Law and Human Behavior*, *11*(1), 55–62. https://doi.org/10.1007/BF01044839

Mercer, T., & Jones, G. A. (2019). Time-dependent forgetting and retrieval practice effects in detailed visual long-term memory. *Quarterly Journal of Experimental Psychology*, *72*(6), 1561–1577. https://doi.org/10.1177/1747021818799697

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorial in Quantitative Methods for Psychology*, *4*(2), 61–64.

Nakayama, K., Deutsch, D., & Nakayama, K. (1999). Inattentional blindness. *Trends in Cognitive Sciences*, *3*(1), 39. https://doi.org/10.1016/S1364-6613(98)01244-3

Olivers, C. N. L. (2009). What drives memory-driven attentional capture? The effects of memory type, display type, and search type. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(5), 1275–1291. https://doi.org/10.1037/a0013896

Olivers, C. N. L., Meijer, F., & Theeuwes, J. (2006). Feature-based memory-driven attentional capture: Visual working memory content affects visual attention. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(5), 1243–1265. https://doi.org/10.1037/0096-1523.32.5.1243

Pinto, Y., Leij, A. R. Van Der, Sligte, I. G., Lamme, V. A. F., & Scholte, H. S. (2013). Bottom-up and top-down attention are independent. *Journal of Vision*, *13*(3), 1–14. https://doi.org/10.1167/13.3.16.doi

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*(1), 3–25. https://doi.org/10.1080/00335558008248231

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? *Journal of Experimental Psychology: Human Perception and Performance*, *18*(3), 849–860. https://doi.org/10.1037/0096-1523.18.3.849

Rensink, R. (2002). Change detection. *Annual Review of Psychology*, *53*, 245–277. https://doi.org/10.1146/annurev.psych.53.100901.135125

Schulman, A. I. (1971). Recognition memory for targets from a scanned word list. *British Journal of Psychology*, *62*(3), 335–346. https://doi.org/10.1111/j.2044-8295.1971.tb02044.x

Shepard, R. N. (1967). Recognition memory for words, sentences, and pictures. *Journal of Verbal Learning and Verbal Behavior*, *6*(1), 156–163. https://doi.org/10.1016/S0022-5371(67)80067-7

Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception*, *28*(9), 1059–1074. https://doi.org/10.1068/p281059

Simons, D. J., & Rensink, R. A. (2005). Change blindness: Past, present, and future. *Trends in Cognitive Sciences, 9*(1), 16–20. https://doi.org/10.1016/j.tics.2004.11.006

Soto, D., Heinke, D., Humphreys, G. W., & Blanco, M. J. (2005). Early, involuntary top-down guidance of attention from working memory. *Journal of Experimental Psychology: Human Perception and Performance, 31*(2), 248–261. https://doi.org/10.1037/0096-1523.31.2.248

Soto, D., Humphreys, G. W., & Rotshtein, P. (2007). Dissociating the neural mechanisms of memory-based guidance of visual selection. *Proceedings of the National Academy of Sciences of the United States of America, 104*(43), 17186–17191. https://doi.org/10.1073/pnas.0703706104

Standing, L., Conezio, J., & Haber, R. N. (1970). Perception and memory for pictures: Single-trial learning of 2500 visual stimuli. *Psychonomic Science, 19*(2), 73–74. https://doi.org/10.3758/BF03337426

Stanislaw, H., & Todorov, N. (1999). Calculating of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers, 31*(1), 137–149. https://doi.org/10.3758/BF03207704

Swallow, K. M., & Jiang, Y. V. (2010). The attentional boost effect: Transient increases in attention to one task enhance performance in a second task. *Cognition, 115*(1), 118–132. https://doi.org/10.1016/j.cognition.2009.12.003

Swallow, K. M., & Jiang, Y. V. (2013, May 20). Attentional load and attentional boost: A review of data and theory. *Frontiers in Psychology.* https://doi.org/10.3389/fpsyg.2013.00274

Tatler, B. W., & Tatler, S. L. (2013). The influence of instructions on object memory in a realworld setting. *Journal of Vision, 13*(2), 5–5. https://doi.org/10.1167/13.2.5

Theeuwes, J. (2010). Top-down and bottom-up control of visual selection. *Acta Psychologica, 135*(2), 77–99. https://doi.org/10.1016/j.actpsy.2010.02.006

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80*(5), 352–373. https://doi.org/10.1037/h0020071

van Moorselaar, D., Battistoni, E., Theeuwes, J., & Olivers, C. N. L. (2015). Rapid influences of cued visual memories on attentional guidance. *Annals of the New York Academy of Sciences, 1339*(1), 1–10. https://doi.org/10.1111/nyas.12574

Vogel, E., Luck, S., & Shapiro, K. (1998). Electrophysiological evidence for a postperceptual locus of suppression during the attentional blink. *Journal of Experimental Psychology: Human Perception and Performance, 24*(6), 1656–1674. https://doi.org/10.1037/0096-1523.24.6.1656. Retrieved from http://psycnet.apa.org/journals/xhp/24/6/1656/

Vogt, S., & Magnussen, S. (2007). Long-term memory for 400 pictures on a common theme. *Experimental Psychology, 54*(4), 298–303. https://doi.org/10.1027/1618-3169.54.4.298

Williams, C. C. (2010a). Incidental and intentional visual memory: What memories are and are not affected by encoding tasks? *Visual Cognition, 18*(9), 1348–1367. https://doi.org/10.1080/13506285.2010.486280

Williams, C. C. (2010b). Not all visual memories are created equal. *Visual Cognition, 18*(2), 201–228. https://doi.org/10.1080/13506280802664482

Williams, C. C., Henderson, J. M., & Zacks, R. T. (2005). Incidental visual memory for targets and distractors in visual search. *Perception and Psychophysics, 67*(5), 816–827. https://doi.org/10.3758/BF03193535

Yantis, S., & Hillstrom, A. P. (1994). Stimulus-driven attentional capture: Evidence from equiluminant visual objects. *Journal of Experimental Psychology: Human Perception and Performance, 20*(1), 95–107. https://doi.org/10.1037/0096-1523.20.1.95