2016

Constructing the Search Template: Episodic and Semantic Influences on Categorical Template Formation

Juan D. Guevara Pinto
Louisiana State University and Agricultural and Mechanical College, jgueva3@lsu.edu

CONSTRUCTING THE SEARCH TEMPLATE:
EPISODIC AND SEMANTIC INFLUENCES ON CATEGORICAL TEMPLATE
FORMATION

A Thesis

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
Requirements for the degree of
Master of Arts

in

The Department of Psychology

by
Juan D. Guevara Pinto
B.S., Georgia Southern University, 2014
May 2017

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

ABSTRACT

Search efficiency is usually improved by presenting observers with highly detailed target cues (e.g., pictures). However, in the absence of accurate target cues, observers must rely only on categorical information to find targets. Models of visual search suggest that guidance in a categorical search results from matching categorically-diagnostic target features in the search display to a top-down attentional set (i.e., the search template), but the mechanisms by which such attentional set is constructed have not been specified. The present investigation examined the influences of both semantic and episodic memory on search template formation. More precisely, the present study tested whether observers incorporated a recent experience with a target-category exemplar into their search template, instead of relying on long-term learned regularities about object categories (Experiment 1) or on the semantic context of the search display (Experiment 2). In both experiments participants completed a categorical search task (75% of trials) in conjunction with a dot-probe response task (25% of trials). The dot-probe response task assessed the contents of the search template by capturing spatial attention if the dot-probe was presented at an inconsistent location relative to objects matching the search template. In Experiment 1 it was shown that observers include recently encoded objects into their search templates, when given the opportunity to do so. Experiment 2, however, showed that observers rely on context semantics to construct categorical search templates, and they continue to do so in the presence of repeated target cues related to different contexts. These results suggest that observers can, and will, rely on episodic representations to construct categorical search templates when such representations are available, but only if no external cues (i.e., scene semantics) are present to identify criterial target feature.

INTRODUCTION

People frequently conduct some sort of visual search as part of their daily routines. From finding small items (e.g., looking for car keys) to locating much larger objects (e.g., searching for a car in a parking lot), visual search is an integral part of different activities. In laboratory visual search tasks, observers are shown a search target, which they store in visual working memory (VWM), in order to find that target later among an item array, or in a scene, as quickly as possible. Researchers often manipulate the precision of the target cue to understand how observers conduct efficient searches. The present thesis focused on how observers use episodic and semantic information to form representations of search targets when only categorical information about targets is available. More precisely, the contents of VWM in preparation for a search task was assessed using an attention-capture paradigm to examine whether observers would include features from a single, recent experience with a target category in their target representation or if they would rely on a gist-level representation constructed from semantic memory.

Visual search is typically studied by presenting observers with a template cue (i.e., picture of the to-be-found target) or a categorical cue (i.e., word describing the target) and then having them locate the target among a set of distractor items as quickly as possible (see Figure 1). Researchers usually measure the accuracy of the search (i.e., whether or not the observer found the target) as well as the efficiency of the search (i.e., how quickly was the target found). Moreover, if eye-movements are recorded, the search process can also be divided into three stages, search initiation, guidance, and target verification (Castelhano, Pollastek, & Cave, 2008; Malcolm & Henderson, 2009; Spotorno, Malcolm, Tatler, 2014). The initiation of a search is measured by the latency and direction of the first saccade (i.e., a quick movement of both eyes

between two fixation points) away from a starting location, and is assumed to reflect pre-search preparatory processes (Malcolm & Henderson, 2009; Spotorno, Malcolm, Tatler, 2014). Search guidance, defined as the latency from the first saccade to the first fixation on the target, involves the selection of specific items in the array for further processing, as dictated by bottom-up and top-down control mechanisms (Wolfe, Cave, & Franzel, 1989). These attentional-control mechanisms allow observers to ignore irrelevant and distracting items, making the search process more efficient. Lastly, target verification results from fixating on an object and deciding that it is the target. This last stage is influenced by the target information stored in VWM (Castelhano, Pollastek, & Cave, 2008; Malcolm & Henderson, 2009).

Figure 1. Example of a categorical visual search experiment trial.

In search tasks, the number of distractor items in the display can vary from trial to trial, impairing the first two search stages when many distractors are presented (Duncan & Humphreys, 1989; Wolfe, Cave, & Franzel, 1989). Moreover, distinguishing task-relevant from task-irrelevant features is not a trivial process, as items in the search array may share several features with each other and with the target. Consequently, the search cue plays an important role, by informing observers which features should be maintained in VWM, as those features will facilitate the search process. Research demonstrates that template cues aid search guidance more efficiently than categorical cues, allowing observers to locate targets much faster (Yang & Zelnisky, 2009; Vickery, King, & Jiang, 2005). Yet, in many real-life situations, such as TSA

agents searching for illegal items in passengers' luggage or oncologists searching for nodules in x-ray scans, observers do not always have access to template representations, and must instead rely on categorical information to find targets.

Visual search models suggest that, when observers search for categorically-defined targets, top-down attentional sets guide the search toward task-relevant (or categorically-diagnostic) object features (Duncan & Humphreys, 1989; Wolfe, Cave, & Franzel, 1989). This attentional set, or search template, contains a visual representation of the target item and its features, and it is compared to potential target items in the search display. In basic research, this is commonly studied using simple target items with a controlled number of distinctive features. For example, if an observer is searching for a small, red, vertical line among different distractors, her attention may be directed towards the size, color, or orientation of the items in the search display. The closer in physical appearance the search template is with the target in the display, the stronger the attentional capture of the latter, and consequently the faster observers will detect the target (Bravo & Farid, 2009, 2012; Duncan & Humphreys, 1989; Wolfe & Horowitz, 2004). In a series of experiments using real-world objects as template cues, Hout and Goldinger (2014) manipulated the precision of search templates by varying template-to-target similarity. For example, if the search cue was a picture of an empty coffee mug, the target was sometimes that specific coffee mug, and other times it would be an inaccurate depiction of the mug (e.g., same coffee mug filled with coffee). They found that, although participants knew that targets might differ from the search cues, overall search times, scan-path ratios (i.e., ratio of objects fixated prior to the target), and verification times increased as the similarity between the expected target and the encountered target decreased. These results suggest that accurate search templates are crucial for quickly finding target items, as they facilitate search guidance and target verification.

3

In categorically-defined searches, the information available to construct accurate search templates is frequently very vague or limited (Malcolm & Henderson, 2009; Schmidt & Zelinsky, 2009). To enhance search efficiency, categorical search templates should be constructed with features that generalize to most exemplars of the target category, but remain specific enough that is possible to discriminate task-irrelevant features (Bravo & Farid, 2012; Castelhano & Heaven, 2010; Ullman, Vidal-Naquet, & Sali, 2002). For instance, Schmidt and Zelinsky (2009) observed that search times were reduced when observers were presented with detailed categorical cues (e.g., boots) relative to abstract cues (e.g., footwear). Moreover, search initiation and guidance, measured as the ratio of initial saccades to target and scan-path ratios, respectively, improved as the cues became more specific (e.g., boots vs. brown boots). These results suggest that observers can extract specific information from available categorical cues to form search templates, and that task-relevant features can be emphasized when constructing categorical search templates.

An important aspect of categorical search is how categorical knowledge is organized in the observer's semantic network. If category exemplars are organized at different levels or ranks within a given category, it is possible that some exemplars may be included in categorical search templates for easily than others Early studies demonstrated that objects are organized in a categorical hierarchy (Murphy & Brownell, 1985; Rosch et al., 1976). For example, the same object (e.g., taxi) can be categorized at a superordinate-level (vehicle), at a basic-level (car), or at a subordinate-level (sedan). Research has shown that objects are categorized fastest at the basic-level (Murphy & Brownell, 1985), and that, when individuals encode either a superordinate or a subordinate word, they tend to mistakenly retrieve its basic-level counterpart (Pansky & Koriat, 2004). In categorical search, Maxfield and Zelinsky (2012) observed reduced search times when

observers searched for subordinate-level cues relative to basic-level and superordinate-level cues. However, a basic-level advantage was observed during the verification stage of the search, suggesting that categorical search templates that generalize to most category exemplars also benefit search efficiency. Maxfield and Zelinsky (2012) explained this basic-level advantage in terms of categorical specificity, the degree of visual detailsprovided at each hierarchical-level, and distinctiveness, how similar (or dissimilar) category exemplars are to other exemplars at each hierarchical-level. For instance, subordinate-level cues are very specific, but they are not very distinctive as they share features across exemplars (e.g., taxi vs. police car). Visual features of superordinate-level cues on the other hand are very distinctive (e.g., vehicle vs. furniture), but they do not provide specific details to select target-relevant features. As such, basic-level cues provide the most satisfying balance between specificity and distinctiveness, resulting in faster target verification times.

Categorically-Diagnostic Target Features

Given that object categories can include several exemplars, accurate categorical cues are ideal for observers to quickly locate targets. Even when searching for subordinate-level cues, observers are faced with a range of features and exemplars within a single category. Therefore, one unanswered question in the visual search literature is what determines the most representative features of an object category. For instance, when observers are searching for the categorical cue helmet, do they search for features akin to a football helmet or to Roman soldier helmet? It has been suggested that the most criterial object category features are generalized object shapes, rather than specific, simple features (e.g., color or orientation), as simple features can easily vary from one category exemplar to the next (Evans & Triesman, 2005; Levin et al., 2001; Reeder & Peelen, 2013; Ullman, Vidal-Naquet, & Sali, 2002). In some cases, however,

generalized object shapes can also vary from one category exemplar to another (e.g., a birthday cake may have a different shape than a wedding cake). Theories of visual search, meanwhile, operate on the assumption that observers are aware of the categorically-diagnostic features for any object category (Duncan & Humphreys, 1989; Wolfe, Cave, & Franzel, 1989). In his latest version of Guided Search, Wolfe describes top-down guidance as "based on the match between a stimulus and the desired properties of the target…" (p.105, Wolfe, 2007), but does not discuss how observers decide on the desired target properties.

Figure 2. Example of a search target.

Figure 1 depicts a possible search target for the category car. According to Guided Search (depicted in Figure 3), object features in the display are processed in parallel by the visual system from early perceptual stages, allowing the observer to perceive the different feature-configurations in the search environment. Features such as color and orientation are initially processed across all objects in the display, along with many other different feature-channels (e.g., round shapes for wheels, rectangular shapes for windows). However, a selective attention mechanism that matches object features to the search template functions as a filter for further processing. In the target example in Figure 2, colors would be matched to red and orientations would be matched to leftwards. This feature filtering results in an attentional bottleneck, allowing only one template-matching feature to pass the bottleneck at a time, while irrelevant features (e.g., green, rightwards) are not selected for further processing. Features that match the

search template begin to accumulate past the bottleneck, activating different thresholds for each individual feature-channel, including specific, task-relevant feature-channels (e.g., black, round shapes for wheels, grayish, rectangular shapes for windows as in Figure 2). Once the activation of the task-relevant feature-channels reach a threshold, the target is identified.
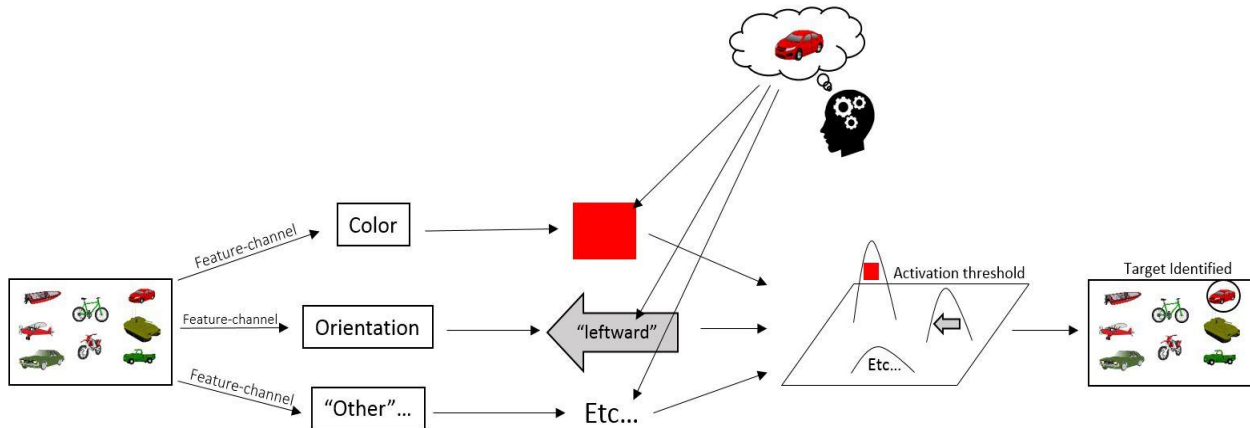
Figure 3. Simplified Guided Search model.

Duncan and Humphreys (1989) also proposed a model in which items in the display are processed in parallel during early perceptual processing, allocating a limited amount of attentional resources to each item. Items are simultaneously grouped by the visual system based on feature similarity, with each item receiving a particular weight, or bias, of resources according to goal-driven control settings determined by the search template. Because resources are limited, greater resource allocation to a particular item decreases the amount resources directed to other items, increasing the chances of template-matching items being preferentially selected (as they would receive the greatest bias for resource allocation). Given the example in Figure 2, when the observer views the display, all items are processed and similar items are grouped together (e.g., same-color items are grouped together). Bias for resource allocation would be greater for red items, relative to any other color, increasing the chances of the car in Figure 2 being selected. It

7

is important to mention that not all features are processed equally, as bottom-up mechanisms influence the allocation of resources in favor of the most salient futures (i.e., color is preferentially selected over orientation).

Guided Search (Wolfe, Cave, & Franzel, 1989) and Duncan's and Humphreys' (1989) model describe the mechanisms involved in several search phenomena, including target-distractor similarity and distractor heterogeneity. For example, it is harder to search for targets that are very similar to the distractors in the display, because many object features pass the attentional bottleneck (Wolfe et al., 1989), or because too many objects receive a bias for resource allocation, significantly reducing the amount of resources directed to the target object (Duncan & Humphreys, 1989). However, these models neither describe by which mechanism(s) observers determine the task-relevant features of categorically-defined targets. For instance, if an observer is presented with the categorical cue car, it is assumed that he will automatically include the features in Figure 2 (e.g., red, leftwards) in his search template. Levin and colleagues (2001) first suggested that search models need to include an explanation of how features are grouped to form targets and target categories. In a series of experiments, they asked observers to categorically search for either an animal among man-made distractors (i.e., artifacts), or an artifact among animal distractors. They showed that observers quickly located both types of categorical targets, but that search was more efficient (i.e., shorter search slopes) when searching for an artifact among animals than vice-versa. In their second experiment they showed this effect persisted when the features of the target and distractor items were jumbled up. They further examined this effect by calculating how much of a rectangular or curvilinear contour each item has (i.e., rectilinearity). They proposed that observers may rely on long-term knowledge about animals and artifacts to identify categorically-diagnostic target features, namely that some

artifacts can have curvilinear shape but no animal has a rectangular shape. In their third experiment, a negative correlation between search slopes and rectilinearity was observed for the artifacts, but not for animals, suggesting that individuals were using contour shape to locate the target (this correlation was later replicated in their fourth experiment). The results from this study suggest that observers can rely on long-term knowledge to identify criterial target features to attend to when searching for categorical targets (in this case, knowledge that all animals share curvilinear shape contours). Levin and colleagues (2001) concluded that existing theories of visual search are helpful for understanding and explaining the categorical search data, but they have to be expanded to account for observers' abilities to quickly identify features that specify object categories.

Target Typicality and Visual Search

There may be several potential sources of information on which observers can rely to identify the distinctive properties of categorical targets. One possibility, as suggested by Levin et al. (2001), is that observers rely on long-term stored associations between objects and their features. For example, when searching for a categorical target, observers might simply search for the most stereotypical object in the target category. In fact, research has shown that observers do tend to rely on long-term learned regularities about object categories and their features to form search templates. This is evident from the effects of target typicality (how "common" the target is relative to other same-category objects) on search guidance (Maxfield, Stalder, & Zelinsky, 2014; Robbins & Hout, 2015), and on target verification times (Castelhano, Pollastek, & Cave, 2008). These effects show that observers are faster at finding "typical" category exemplars than they are at finding atypical exemplars. For example, if an observer searches for "lamp," she will be faster at finding and verifying the target when it is a nightstand lamp than when it is a gas

lamp, assuming that nightstand lamps are more frequently encountered than gas lamps in everyday life. Similarly, an observer's search will be significantly slowed down if the target is an atypical member of the target category (e.g., lava lamp).

Although typicality is a viable mechanism that observers use to identify categorically-diagnostic target features, guidance for typical targets seems to be dependent on the distractors in the display. When searching through distractors of the same superordinate category as the target, guidance is inhibited by distractor-target similarities (Castelhano et al., 2008; Robbins & Hout, 2015), but when distractors represent different superordinate categories, typical target objects are located faster than atypical target objects (Maxfield et al., 2014). Verification, on the other hand, is influenced by typicality because word cues bias the activation of typical category-features when constructing categorical search templates (Maxfield et al., 2014). This results in faster verification of typical target objects, because they are more likely to share features with the search template (regardless of the idiosyncratic details of the category exemplar serving as the target object). For example, observers' mental representations of lamp will be biased toward the common features of table lamps, making it easier to identify the target when it is typical (e.g., a table lamp) than when it is atypical (e.g., a gas lamp), regardless of which specific nightstand lamp is used as the target.

Research on target typicality has typically examined effects in isolation from other potential mechanisms involved in constructing search templates (e.g., episodic memory). It is possible that when an alternative source of information is available to inform template creation (e.g., recency of experience with a specific object), observers may adopt it over typicality to identify the categorically-diagnostic features of a categorical target. Typicality-based templates arise from a combination of individual episodic representations of the target category, resulting

10

in a gist representation without including criterial details. On the other hand, "purely" episodic-based templates result from a single episodic experience, allowing the observer to match, as closely as possible, their top-down representation with the search target. Still, it is not known, however, to what degree observers can (or will) rely on recent experience with specific objects over learned object-regularities to construct search templates.

Semantic Context of the Search Display

Although object typicality can be a readily available source of information to build search templates, observers may also rely on external, rather than internal, cues to achieve the same goal. When the search display contains a specific semantic structure, either by object-to-object relations or by the inclusion of a scenic background, attention can be guided by the observer's semantic network (Malcolm & Henderson, 2009; Torralba, Oliva, Castelhano, & Henderson, 2006; Võ & Wolfe, 2012, 2013). For example, when observers are searching for a pen in an office scene, attention will be directed towards the desk (even if the specific office scene is novel), because pens are typically located on desks. In visual search through real-world scenes, the semantic content of the scene interacts with target information to efficiently guide the search (Spotorno, Malcolm, & Tatler, 2014, 2015). However, the semantic content of the scene needs to be clear and distinguishable. For instance, Castelhano and Heaven (2010) found that simply knowing the gist of the scene prior to search does not improve search performance. Participants in their study had previous knowledge about the search context prior to the search task (either by previewing the scene or by being cued with a word describing the gist of the scene). Relative to a control condition (no preview nor word cue), previewing the search scene improved both search initiation and guidance (but not verification). The gist word cue, on the other hand, had no significant effect on any of the three search stages (Castelhano & Heaven, 2010).

Although prior knowledge of scene gist does not seem to improve search efficiency by itself, it is still unknown whether the search context can influence the formation of search templates. In categorically-defined search, knowing the gist of the search context prior to search may allow observers to incorporate gist-related features in their search templates, resulting in more accurate search templates. For example, if observers know they will search for the categorical target "hat" within a beach scene, he may adopt a search template that resembles a beach hat, instead of a search template resembling any non-specific hat, or a specific, but context-irrelevant hat (e.g., a hard hat). As such, it is possible that the semantic context of the search display can also define the content of categorical search templates.

Episodic Memory and Target-Template Formation

The role of episodic memory in visual search, especially in search guidance, has been a topic of recent debate (Horowitz & Wolfe, 1998; Hollingworth, 2012; Peterson et al., 2001; Võ & Wolfe, 2012, 2013, 2015; Wolfe, Kemplen, Dahlen, 2000). Studies of contextual cuing and repeated visual search have shown that subsequent searches are benefited when preceded by repeated exposures to the item arrangement, as well as to the target features and identities, in the search display (Chun & Jiang, 1998; Hout & Goldinger, 2010; Huang, Holcombe, & Pashler, 2004). However, Võ and Wolfe (2013) demonstrated that the utility of episodic memory in guiding search depends on the availability and reliability of other sources of guidance (e.g., the target information and the scene context). In the absence of an accurate search template, and in the presence of conflicting semantic information (i.e., improbable object placement within a scene), individuals will rely on episodic memory traces from previous searches to find the target. In their study, Võ and Wolfe (2013) had participants search for categorical targets in scenes with consistent or inconsistent object arrangements. Participants searched through these displays

across two blocks of trials, with shorter search times observed in the second block relative to the first block. More importantly, the benefits from repeated search were greater for the conditions with improbable object arrangements. These results suggest that episodic memory can reliably guide search, and that reliance on episodic guidance is crucial when semantic information becomes unreliable.

Whereas previous research has shown that episodic memory can successfully guide search (Chun & Jiang, 1998; Võ & Wolfe, 2012, 2013, 2015), less is known about the role episodic memory plays in the formation of target templates. Hollingworth and Henderson (2002) showed that observers retain very detailed representations in memory of objects fixated in scenes. Participants passively studied real-world scenes in preparation for a long-term memory (LTM) test. In some trials, a target object changed in identity (e.g., a notebook to a different notebook) or in category (e.g., a notebook to a floppy disk) once the observer's fixation left the region surrounding that object. This procedure allowed for object-changes to occur outside the observer's focus of attention, with observers indicating if a change occurred at the end of each trial. The results showed that change-detection performance for both change-types was higher than the false-alarm rate for no-change trials. Observers later selected one of two scenes, which were visually identical except for the target object (i.e., one was the fixated object and the other was the change object). LTM performance for the fixated object was above chance performance, regardless of whether the object changed in identity or category (similar results have been observed in incidental memory for distractor objects in visual search tasks; e.g., Hout & Goldinger, 2010; Williams, Henderson, & Zacks, 2005). These results suggest that once objects are fixated, observers are able to stored representations for such objects with high fidelity, which later could be used to identify diagnostic target features in a categorical search. For instance,

when searching for lamp, observers can incorporate features from previous experience with the target into the search template to narrow down search. In fact, when participants are trained to associate one specific categorical exemplar to a name cue (e.g., an image of fish with its species name), they use very detailed search templates to find the target (Bravo & Farid, 2009). However, if the target was not identical to the trained exemplar, this highly detailed template became harmful (e.g., different fish of the same species), as reflected by increased search times. In a subsequent study, Bravo and Farid (2012) demonstrated that when participants are trained on multiple exemplars of single category, they learned to use a more "loose" search template, facilitating search when the target varied across trials. However, in many search tasks (both lab-based and applied), observers usually do not receive ample object category training, which raises the question on how extensively experienced do observers' have to be with an object category and its exemplars to incorporate such exemplars in their search templates. It is possible that single and specific episodic representations of object categories will suffice to identify target features. For instance, if observers were to encode the car exemplar in Figure 1, they may adopt a red sedan as their search template when given the categorical cue car, rather than rely on learned regularities to form categorical search templates for car.

Unlike episodic guidance (e.g., Võ & Wolfe, 2013), it is possible that the semantic context of the search display may be irrelevant for observers to include episodic representations in the search template. For instance, evidence from research on selection history has shown that individuals prioritize the selection of information that has been previously attended (Awh, Belopolsky, & Theeuwes, 2012; Huang, Holcombe, & Pashler, 2004; Maljkovic & Nakarama, 1994; Wolfe et al., 2013). If observers are able to adopt previous instances of attendedobjects as the search templates, then they would use such templates regardless of the semantic congruency

of the search template and the search display. Therefore, if observers attend to a specific exemplar of an object category, they may prioritize the selection of features from that exemplar to be included in the subsequent categorical search, regardless of the semantic meaning of the display.

One or Two Search Templates?

As discussed above, it is possible for observers to rely on either semantic or episodic traces to form categorical search templates. However, could it be possible for both sources to simultaneously guide search, each by contributing to its own independent search template? Some evidence suggests that two different templates can concurrently guide the search. For instance, Beck, Hollingworth, & Luck (2012) showed that observers successfully maintained two representations (e.g., two colors) during visual search and still found the target. When instructed to attend only to one color at a time, observers searched sequentially-exhaustive, meaning that they fixated exclusively on items sharing one of the target colors before switching to the other color if the target was not found. This type of search yielded significant time costs when observers switched from one color to the other. However, when observers were instructed to search for both colors simultaneously, observers fixated back and forth on items from both colors almost at no costs, indicating that multiple search templates can guide the search when instructed to do so.

Despite the evidence in favor of multiple search templates, some researchers assert that only one template can ever guide attention on any given time (e.g., van Moorselar, Theeuwes, & Olivers, 2014). Evidence suggests that individuals can hold approximately four independent representations in VWM (Luck & Vogel, 1997), and Olivers, Peters, Houtkamp, and Roelfsema (2011) proposed a functional division between such representations. Accessory representations

are actively maintained in VWM, but cannot access attentional-control processes nor are they affected by task-irrelevant information. These are contrasted with attentional template representation, which is a single representation located in the focus of attention. Although the attentional template can be influenced by task-irrelevant information, it can bias attention towards goal-directed items. The attentional template can continuously alternate between the multiple accessory representations stored in VWM, however, when one representation is selected, the others are denied access to visual input. To test this, van Moorselaar et al. (2014) presented observers with arrays of colored squares that had to be remembered for a change-detection task. Following the offset of the change-array, a search array appeared and observers searched for a diamond among circles. Importantly, one of the search distractor circles could be colored with 1) the same color as one of the squares in the change-array, 2) a different color than those in the change-array, or 3) no color at all. Across several experiments, they showed that when only one color square was presented in the change-detection array, search times were significantly slowed if the distractor in the search display shared that same color. This suggested that the color in the change-detection array occupied the attentional template status. However, increasing the number of colored squares in the change-array did not impact search times across the different search trial types. These results were taken as support of the functional division account of VWM, as colors in trials where the change-detection array size was greater than 1 were allegedly stored as accessory representations, and the competition between those colors for the attentional template status was not resolved by the time observers located the target.

The literature provides mixed evidence for single and multiple template accounts of search guidance. However, it is important to note that in the studies that have examined both of these possibilities, the diagnostic target features were readily available to the observers (e.g., two

colors or a diamond), which is not the case in a categorical search. Furthermore, even if two independent categorical representations are able to simultaneously guide attention, it is still a matter of empirical inquiry to investigate by which mechanism those representations are created.

Contingent Attentional Capture

The precision of the search template clearly impacts visual search (e.g., Bravo & Farid, 2009, 2012; Hout & Goldinger, 2014), but the distractor objects through which one searches also have an impact on search performance. For example, the level of target-distractor similarity in search displays strongly influences the speed with which observers find targets, such that high levels of target-distractor similarity increase search times (Alexander & Zelinsky, 2012; Duncan & Humphreys, 1989). Even in target absent trials, Alexander and Zelinksy (2012) found that participants terminated search more quickly when searching through low-similarity distractors, relative to high-similarity distractors. When displays contained high- and low-similarity distractors, the percentage of first fixations directed to high-similarity distractors was greater than the percentage directed to medium- or low-similarity distractors (Alexander & Zelnisky, 2012, Exp. 2). The slower responses and high percentage of first fixations directed to high-similarity distractors in target-absent trials were interpreted as reflecting automatic capture of attention by distractors matching the search template.

As discussed, visual search models (Wolfe et al., 1989; Duncan & Humphreys, 1989) suggest that distractors matching the search template slow search because their features break through the attentional bottleneck or because resource allocation is biased towards them, delaying the processing of the target (if present). In a recent study, Sun and colleagues (2015) employed a similar procedure as van Moorselaar et al. (2014), and showed that distractors can delay target identification if their features match information being rehearsed in working

memory (WM), even if such information is irrelevant to the target item. Participants kept in memory either one colored circle (e.g., red) or one word related to that color (e.g., rose) which would be later matched to a probe. After encoding the item, they searched for a gray target disk containing a tilted line among several gray distractor disks (which contained Xs). In half of the trials, one of the distractor disks matched the color of the item in WM, and in the other half, one of the distractors was of an unrelated color (e.g., green). Search times were significantly slower when participants held a color, or a word related to a color (e.g., rose), in memory while searching through displays containing a distractor of the same color, but not when one of the distractors was of an unrelated color. These results suggest that, during visual search, spatial attention can be captured by features rehearsed in WM, even if such features are irrelevant to the search task. Additionally, these results confirm earlier reports (e.g., Olivers et al., 2011) that only a single item in VWM can guide attention at one time.

The allocation of visual selective attention is contingent on the relationship between the characteristics of the stimuli and the observer's attentional-control settings defined by the ongoing task (Bacon & Egeth, 1994; Folk, Leber, & Egeth, 2002). Theories of attentional control have typically dissociated between bottom-up and top-down attentional control settings, although some have advocated for the inclusion of selection history as its own mode of attentional control (see Awh, Belopolsky, & Theewes, 2012). Bottom-up attentional control involves the observer's ability to direct attention to physical properties of the stimuli (i.e., external to the observer), such as stimulus salience. Top-down attentional control, on the other hand, involves the observer's ability to direct attention according to task-related goals (i.e., internal to the observer), such as the search template. Task-demands define an observer's attentional control setting, making the observer more sensitive to bottom-up (stimulus-driven) selection, top-down (goal-driven)

selection, or a combination of both. For example, attentional control settings will differ for tasks

that can be accomplish by relying on stimulus-driven selection (e.g., change-detection) and those

that rely on goal-driven selection (e.g., visual search). Observers' attentional control settings then

bias selection towards objects or features matching their settings, eliciting both voluntary and

involuntary shifts in attention (Bacon & Egeth, 1994; Folk, Leber, & Egeth, 2002).

In a series of modern classic experiments, Folk, Leber, and Egeth (2002) presented

participants with rapid serial visual presentation (RSVP) streams of colored letters, and

participants reported the identity of a target letter of a pre-specified color (e.g., red). In addition

to the central stream of letters, four peripheral distractors appeared at different item lags before

the target letter. They found that target identification was significantly impaired when one of the

distractors shared the same color as the target letter, but not when the distractors were of any

other color. This effect suggests that involuntary shifts in spatial attention are due to the selection

bias towards specific, task-relevant features by the attentional system. Moreover, it suggests that

visual attention can be captured and directed towards task-irrelevant locations, as participants

knew that the target would never be found in the periphery. Bias toward task-relevant properties

is the core of the contingent involuntary orienting hypothesis, which suggests that stimulus

features will capture attention only if they match top-down attentional control settings (Folk,

Remington, & Johnston, 1992). In a similar study, Wyble, Folk, and Potter (2013) showed that

contingent attentional capture is not limited to only distractors sharing the same low-level visual

features as the target, but that capture also occurs when targets and distractors share the same

semantic information. They presented participants with RSVP streams of real-world images and

participants reported the identity of categorically-defined targets. Performance was impaired

when one of the peripheral distractors (e.g., Ferris wheel) was an exemplar of the same

superordinate-level category (e.g., amusement park) as the target image (e.g., bumper cars), suggesting that information at the categorical level is enough to automatically capture an individual's attention.

If involuntary shifts of attention occur when visual features match top-down attentional control settings, then any features matching the search template will capture attention, regardless if a search is conducted or not. Reeder and Peelen (2013) examined the contents of the search template in terms of differences in attentional capture in an unexpected dot-probe task. In their study, participants were cued to search for a car or a person in real-world scenes, indicating whether the target was present or absent. However, in a critical subset of the trials (25% of trials), instead of conducting visual search, participants were briefly presented (67 ms) with silhouette primes of a person and of a car at opposite sides of a central fixation cross. After the offset of the silhouettes, a dot flashed for 100 ms at one of these two locations, and participants indicated via button press on which side the dot appeared. Participants were slower to respond to the dot-probe whenever it appeared at the location opposite the search target silhouette. This effect suggests that spatial attention is quickly and involuntarily captured by stimuli that match the contents of the search template. Therefore, it is possible to examine the contributions of episodic and semantic memory in constructing categorical search templates by measuring their individual spatial attention capture. For instance, if observers are given the opportunity to use episodic, rather than semantic, information to construct a search template, spatial attention should be captured by objects sharing episodic features. The present investigation studied the influences of episodic and semantic memory in search template formation when recent experiences with category exemplars are available to the observer. Given that observers prioritize the selection of previously attended information (e.g., Awh et al., 2012), and that such

20

information is retained with high visual detail (e.g., Hollingworth & Henderson, 2002), it is possible for observers to rely on explicit memory traces regarding the target category, rather than on long-term representations, to form search templates.

The Present Investigation

In general, the present study assessed the contents of categorically-defined search templates. More specifically, the main goals of this Master's thesis were 1) to examine if observers can rely on a different mechanism other than semantic memory to form search templates (e.g., episodic memory) when only categorical target information is available, and 2) to explain how these two information sources interact with each other to construct an efficient categorical search template. Episodic and semantic memory mechanisms were investigated in two experimental designs, addressing the use of internal (Experiment 1) and external (Experiment 2) cues in search template creation. In addition, similar to the effects observed on search guidance (Võ & Wolfe, 2013), it is possible for episodic and semantic memory to compete for prioritization to be included in the target template. Therefore, the relative influence of each information source in selecting specific target-features was assessed with within-experiment manipulations, typicality versus encoded object features in Experiment 1, and semantic context versus target repetition in Experiment 2.

An unanswered question in the visual search literature relates to how observers identify categorically-diagnostic target features when searching for categorical targets (Levin et al., 2001). Although target typicality has been shown to be frequently used by observers to form search templates (Castelhano et al., 2008; Maxfield et al., 2014; Robbins & Hout, 2015), it has often been studied in isolation from other potential mechanisms involved target-template creation. Previous research suggests that observers are able to store representations of objects

prior to the search as a result of explicit training (Bravo & Farid, 2009, 2012) and after fixating on objects during the search task (Hollingworth & Henderson, 2002). If observers can store detailed representations of target objects, it is possible that they may rely on these recent experiences with the target category to construct search templates in subsequent categorical searches. Also, it is possible for the observer to rely on the semantic context of the search display to identify categorically-diagnostic target features. However, whether or not an observer will choose to rely on episodic, rather than semantic, information to decide on categorical target features remains an unanswered theoretical question.

In order to measure the content of search templates, I adopted a contingent attentional capture paradigm similar to that of Reeder and Peelen (2013), in which participants formed categorical search templates in preparation for a search task. In a subset of trials, the contents of the search template were assessed using the dot-probe method, allowing the comparison of attention capture for episodic- versus semantic-based object features. According to the contingent involuntary orienting hypothesis, if the focus of attention is determined by a top-down attentional set (i.e., the search template), then spatial attention should be captured by any stimulus sharing visual features with that attentional set. The strength of the attentional capture was measured by the reaction latency to the onset of the dot-probe, as well as by the number of errors made when reacting to the dot-probe. The slower the reaction (or the higher the error-rate), the stronger the capture of attention by the object to the opposite side of the dot-probe. This would suggest that observers are holding the features of that object (or similar ones) in VWM. Therefore, if observers are holding a semantically-constructed template in VWM, attention capture should be greater for objects sharing semantic features, relative to objects sharing episodic features, that appear opposite the dot-probe. Likewise, if observers are holding an

22

episodically-constructed template in VWM, automatic attention capture should be greater for objects sharing those same episodic-features, relative to objects sharing semantically-based features.

The attentional capture of objects sharing episodic features should be stronger than that of objects sharing semantic features, as search templates constructed from semantic information should not contain the same degree of precision and specificity as those formed from episodic memory (Bravo & Farid, 2009, 2012; Hollingworth & Henderson, 2002). In addition, if episodic information is adapted into the search template, then attentional capture and search performance should remain constant regardless of the congruency between the target object and the search display's semantic context. In two experiments, participants completed visual search for categorically-defined targets. The contents of the search template were investigated by having a critical subset of dot-probe trials interleaved with search trials. The comparison prime in dot-probe trials depended on the feature-selecting mechanisms under study. Experiment 1 compared encoded objects and object typicality, while Experiment 2 compared target repetition and the semantic context.

EXPERIMENT 1

Experiment 1 consisted of two phases, an encoding phase followed by a search phase. Each phase was repeated across six blocks. During the encoding phase, participants committed one object from twenty different object categories into memory. During the search phase, participants conducted multiple categorical searches. A multidimensional scaling (MDS) databased was used to assess the similarity relationships between exemplar objects of the same category (from Hout, Goldinger, & Brady, 2014). MDS provides a "map" of spatial relationships among groups of objects, such that similar objects are located near each other and dissimilar objects are located further apart (see Hout, Papesh, & Goldinger, 2012, for a review). I used existing measurements of similarity to operationalize the number of features shared by any two exemplars of any given category: The closer in MDS space two exemplars were located, the greater the number of features they shared (see Figure 4).
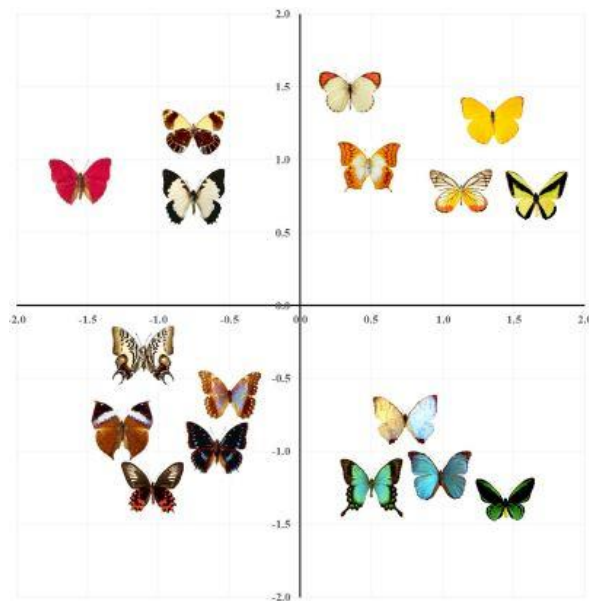

Figure 4. Example of multidimensional scaling (MDS) object similarity relantionship (Figure 3 from Hout et al., 2014) for the object category "butterfly."

Additionally, the MDS database also provides ratings of prototypicality for each object within each object category, such that the exemplar that shares the most features with every other within-category exemplar (i.e., central position in MDS space) is operationalized as the most "typical" object for that object category.[1] As such, if an observer relies on semantic memory to form the search template, then the features of the most typical (prototypical) object should be activated when participants are presented with a categorical target cue. This would be reflected in greater attention capture by prototypical category objects relative to other less typical objects of the same category. Conversely, if the contents of the search template reflect the activation of object representations stored in episodic memory, then attention capture should be greater for objects sharing those specific, activated features, regardless of the object's typicality. This would be reflected by greater attentional capture for objects stored in memory during the encoding phase relative to novel objects of the same category. Additionally, if observers adopt the objects from the encoding phase as their search template, then greater attention capture for this objects should be observed even if those objects are non-prototypical exemplars of their respective category.

Participants. A power analysis (1-β = .95; α = .05; one-tail) conducted on pairwise effect size of dot-probe response times observed in experiment 1 of Reeder, Zoest, & Peelen (2015; N = 12, d = -0.917) suggested a sample size of 15 participants. However, with the goal of increasing the number of observations given the Dot-Probe trial breakdown in the present study (e.g., Studied Prototypical-Inconsistent, Novel NonTypical-Consistent), 64 undergraduate Psychology students participated in exchange for partial course credit. All participants (M_{age} = 18.9 years, SD = 1.7;

---

[1] I also informally verified that the prototypical objects selected by the MDS database were compatible to what observers would expect from reading the category name by comparing the selected object to Google stock images.

36 females) were native-English speakers with normal or normal-to-corrected vision, and normal self-reported color vision. All participants provided signed informed consent before participating in the experiment.

Stimuli and Apparatus. All stimuli was presented on a 21.5-inch Dell-monitor with a screen resolution of 1920 x 1080 and a sampling rate at 60Hz, with all experimental procedures handled by E-Prime 2.0 software (Psychology Software Tools, 2006). A pool of over 3800 real-world object images from 240 distinct object categories, with pre-analyzed MDS distance data for every exemplar pair in each object category (Hout, Goldinger, & Brady, 2014), was used. Across participants, 120 randomly selected categories were used as studied targets (encoding phase), 96 categories were used as unstudied targets (not encoded during encoding phase), and the remaining 24 categories served as distractors during search trials. Target categories were never repeated across trials, however, distractor categories did repeat across trials. For each target category, both studied and unstudied, one typical (the prototypical) and one less typical exemplar were selected. The five-dimensional (5D) classification of the MDS database was used to assess the typicality of each category exemplar. Prototypical and non-typical objects were selected based on their ranking on the 5D centrality ratings for each category. The exemplar ranked the highest (i.e., position 1) was used as the prototypical exemplar for each category.

Figure 5. Prototypical (left) and non-typical (right) exemplars for the category "guitar" (positions 1 and 9 on 5D centrality ranking, respectively).

Similarly, an exemplar ranked in a mid-distance position (e.g., position 9-10) was selected as the non-prototypical object (see Figure 5). The use of mid-distance positions as selection criterion for non-prototypical exemplars safeguards that the object was less typical but not too atypical, as this might capture attention regardless of whether or not participants are using that object as a search template. Additionally, two exemplars from 10 categories (distinct from those in the MDS database) were collected from the internet to be used as targets for practice trials.

Procedure and Design. There were two independent experimental designs for search trials and dot-probe trials. For search trials, the design consisted of a 2 (Target Type: Studied, Novel) x 2 (Typicality: Prototypical, Non-typical) within-subjects design. The experimental design for dot-probe trials consisted of a 2 (Target Type: Studied, Novel) x 2 (Typicality: Prototypical, Non-Typical) x 2 (Dot Location: Congruent, Incongruent) within-subjects design. The experiment consisted of six blocks of trials, each starting with an encoding phase followed by a search phase. The search phase was comprised of 24 search trials and 8 dot-probe trials, with presentation order randomized (across the entire experiment, there was 144 search trials, 48 probe trials). At the start of every block of trials, participants committed to memory a set of 20 objects[2] from distinct categories (half of the objects were prototypical, half were non-typical). Objects were presented in random order for 1000 ms each, followed by a 1000 ms inter-stimulus interval (ISI). Each object was repeated three times during the study phase. After encoding the memory set, participants completed a 2-alternative forced choice (2-AFC) memory test to ensure that the

---

[2] Out of the 20 objects encoded in each block, only 14 were used in the experimental trials, 6 were used as memory-targets in search trials, while the other 8 were used in dot-probe trials. The category of the 6 remaining objects were used as the search cue during half of the target-absent search trials.

objects were successfully stored in memory. The foil objects during the 2-AFC were rotated versions of the encoded exemplars. If rotation was not possible (e.g., circular objects) then the color of the exemplar was changed. Participants were instructed that they had to obtain more than 80% hits on the memory test, and if they failed to do so they had to notify the researcher immediately (the researcher marked the block number that the participant failed before allowing him or her to advance to the search phase). Moreover, participants were advised that the memory and search tasks were independent and unrelated from each other, and that they could complete the search task using any strategy of their preference. After the encoding phase, participants searched for target objects through an array of 16 real-world objects. At the beginning of each search trial, a categorical target cue appeared at the center of the screen. Following a 2500 ms interval, a prompted appear at the bottom-center of the screen asking participants to press SPACE to initiate the trial. This 2500 ms interval was imposed to avoid motor errors of participants engaging in repeated button presses and mistakenly starting the trial without reading the target cue. Moreover, previous research suggest that categorical-target cues require more time (800-1000 ms) to be set in VWM relative to template-target cues (Vickery, King, & Jiang, 2005; Wolfe, Butcher, Lee, & Hyle, 2004). As such, a 2500 ms interval should encourage the formation of an accurate categorical search template. In half of the trials (16 trials) the target cue was a studied category (i.e., the category of an object studied during the encoding phase) and in the other half it was a novel category. After participants initiated the trial, a central fixation cross appeared for 500 ms, followed by the presentation of either a search array (75% of trials; 24 trials) or two object primes from the target category, each at 8.5º of visual angle from either side of the fixation cross (25% of trials; 8 trials). If presented with a search array, participants searched the array, as quickly as possible, until they self-terminated the search by pressing

SPACE. Afterward, the 16 objects were replaced with number digits (1-16) and participants used

the computer mouse to click on the number in the location of the target, terminating the trial. To

encourage effort, a 4 second time-penalty was imposed if participants selected the incorrect

target (see Figure 6). In 25% of the trials (6 trials) the target was a studied-nontypical object; in

another 25% of trials the target was a studied-prototypical object; in another 25% of trials the

target was a non-typical exemplar form a novel category (i.e., novel-nontypical); and in the last

25% of trials the target was the prototypical exemplar from a novel category (i.e., novel-

prototypical; see Appendix A for a trial breakdown). All search trials were target-present trials
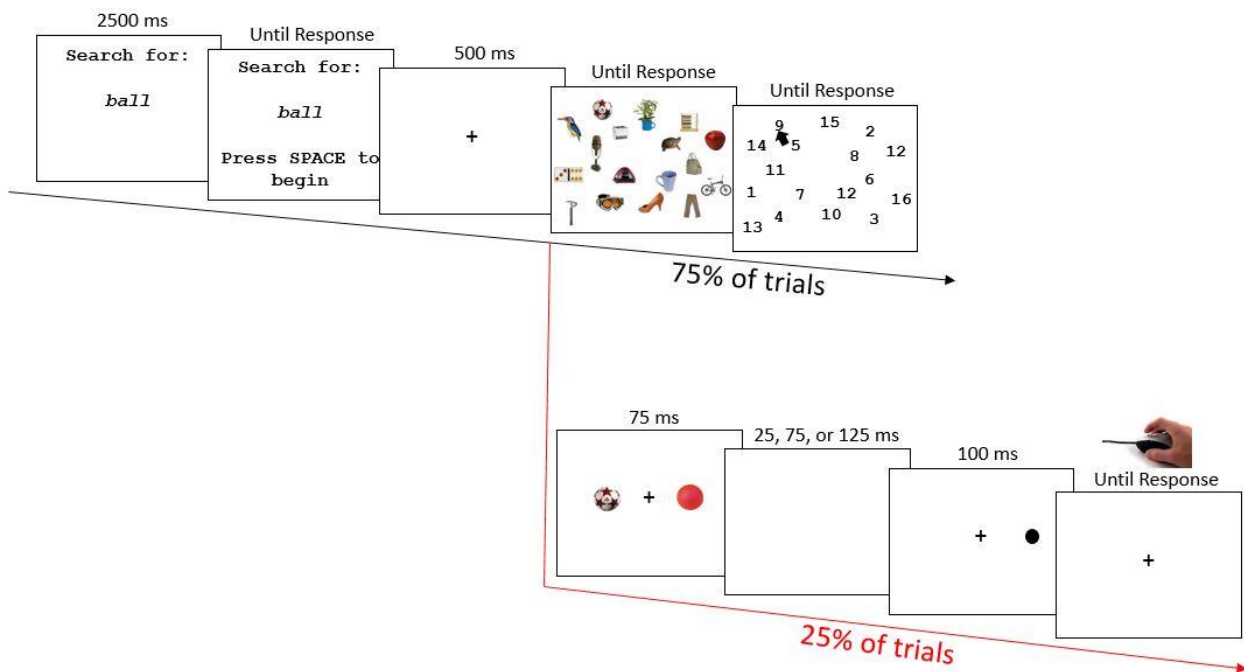
and they were presented in random order.

Figure 6. Example of a search and a dot-probe trial for Experiment 1.

If presented with the two object primes, participants were instructed to disregard these two images, which disappeared after 75 ms. A black dot-probe appeared 25, 75, or 125 ms following the offset of the two primes at one side of the fixation cross for 100 ms. Participants were asked to indicate the location of the dot-probe (left or right) as quickly as possible by clicking the right or left mouse button. In half of the trials (novel trials; 4 trials), both primes were exemplars of a novel category (one non-typical and the prototypical exemplar). Semantic trials were always preceded by a novel categorical target cue. In the other half of the trials (studied trials; 4 trials), one of the primes was a studied object and the other was a novel object of the same category (studied-prototypical and studied-nontypical primes were used an equal number of trials). Studied trials were always preceded by a studied categorical target cue. Additionally, in half of novel trials (2 trials) the dot-probe appeared on the opposite location as the novel-prototypical prime (novel-incongruent trials) while in the other half it appeared in the same location as the novel-prototypical prime (novel-congruent trials). Similarly, in half of the studied trials (2 trials) the dot-probe appeared in the opposite location as the studied prime (studied-incongruent trials) while in the other half it appeared in the same location as the studied prime (studied-congruent trials). A four second time-penalty followed an incorrect response, just like in the search trials (no feedback was provided following correct responses). Moreover, if participants took 1500 ms or more to respond, a prompt appeared on the screen (regardless of response accuracy) reminding participants that they should respond as quickly as possible. Participants were familiarized with the procedure at the beginning of the experiment by completing a shorter memory phase (only 5 objects were studied and tested) as well as 5 search trials and 5 dot-probe trials, both randomly intermixed.

Results and Discussion

The significance level for all analyses was set at .05, and all multiple comparisons were subjected to Bonferroni corrections. There were no significant differences in memory performance across blocks, $F(5, 270) = .61$, $p > .05$ (see Appendix C), however, trial blocks in which participants did not achieve 80% accuracy on the 2-AFC test were not included in any subsequent analyses (13 blocks across 9 participants). Prior to data analysis, outlier trials were filtered out, defined as 2.5 standard deviations above each individual subject mean for Search Times (Search outliers) and Dot-Probe Response Times (Dot-Probe outliers). This resulted in 3% of Search trials (271 trials) and 3.8% of Dot-Probe trials (117 trials) being dropped across all participants, with the remaining 11,601 Search trials and 2,851 Dot-Probe trials included in the analyses.

Trial Initiation. In order to understand how efficiently were observers in creating categorical search templates, trial initiation times for all trials were examined. This was operationalized as the latency of participants' SPACE press after the onset of the prompt asking them to begin each trials. Trial initiation times were collapsed across the six trial blocks. The mean trial initiation times for each participant were analyzed in a 2 (Target Type: Studied, Novel) x 2 (Typicality: Prototypical, Non-typical) repeated-measures ANOVA. A main effect of Target Type was observed, $F(1, 63) = 7.52$, $p < .01$, $\eta^2_p = .11$, in which observers were faster at initiating search trials when target cues were from a studied object ($M = 517$ ms, $SE = 20$ ms) relative to novel objects ($M = 538$ ms, $SE = 22$ ms), suggesting that observers produced categorical search templates faster when episodic representations are available. The main effect of Typicality and the interaction were non-significant.

Search Trials. The variable of interest for search trials was the overall search times (STs),

defined as the latency of participants' target detection response after the onset of the search array

(before the objects are replaced with digits). Only correct trials were analyzed, collapsed across

the six trial blocks[3]. The mean STs for each participant were analyzed in a 2 (Target Type:

Studied, Novel) x 2 (Typicality: Prototypical, Non-typical) repeated-measures ANOVA. Main

effects of Target Type, $F(1, 63) = 196.67$, $p < .00$, $\eta^2_p = .77$, and Typicality, $F(1, 63) = 109$, $p <$

$.01$, $\eta^2_p = .63$, were observed, as well as a significant interaction between the two factors, $F(1,$

$63) = 9.01$, $p < .01$, $\eta^2_p = .13$. Pairwise comparisons revealed that observers' STs were faster for

targets they previously studied ($M = 1323$ ms, $SE = 38$ ms) relative to novel targets ($M = 1635$

ms, $SE = 43$ ms), suggesting that the encoded objects were successfully incorporated into the

search templates. Moreover, observers' STs were also faster for prototypical targets ($M = 1402$

ms, $SE = 38$ ms) than for non-typical targets ($M = 1556$ ms, $SE = 41$ ms), replicating previous

research on typicality effects in visual search (e.g., Maxfield, Stalder, & Zelinsky, 2014). The

interaction was characterized by a significantly shorter STs for Novel-Prototypical targets ($M =$

1525 ms, $SE = 44$ ms) than for Novel-NonTypical targets ($M = 1745$ ms, $SE = 45$ ms), $t(63) =$

8.63, $p < .01$, Cohen's $d = .63$, as well as a significant, but smaller, difference in STs between

Studied-Prototypical targets ($M = 1280$ ms, $SE = 37$ ms), relative to Studied-NonTypical targets

($M = 1367$ ms, $SE = 44$ ms), $t(63) = 3.18$ $p < .05$, Cohen's $d = .27$ (see Figure 7).

---

[3] A significant effect of Block was observed in which both, trial initiation and search times, overall, were reduced in the later blocks (Blocks 5 and 6) relative to earlier blocks (Blocks 1 and 2).
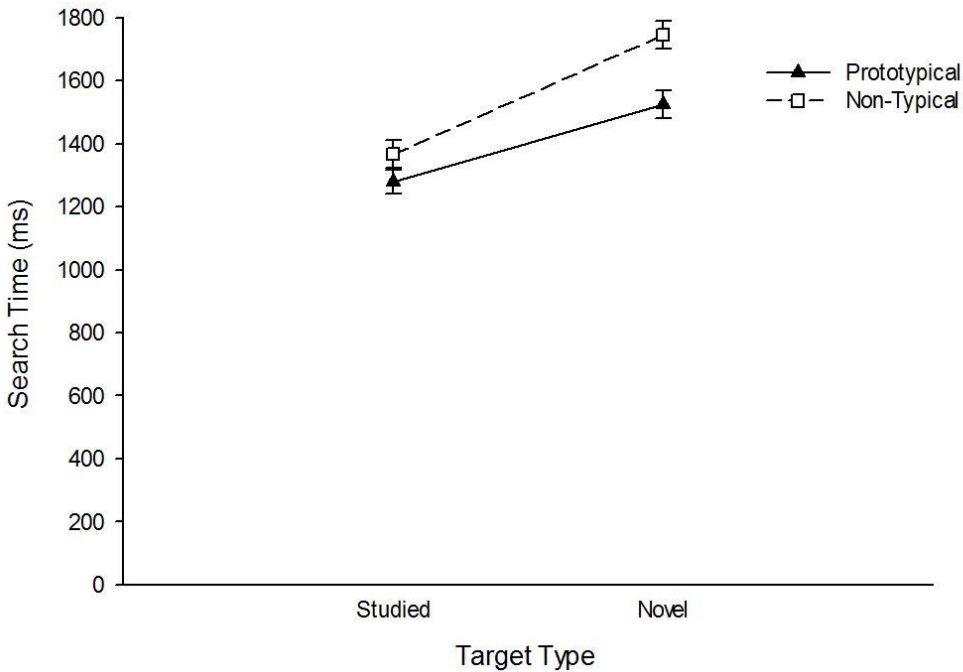
Figure 7. Search times for Exp. 1. Error bars represent standard errors.

It is important to note that overall STs can only provide partial evidence suggesting the prioritization of either episodic or semantic memory for constructing search templates (e.g., without eye-tracking it is not possible to observe which search stage is affected by the experimental manipulation). Definitive conclusions about the mechanism used to build categorical search templates cannot be drawn from STs without converging support from alternative quantifications of the search template. Such specific analyses of the contents of the search template came from comparisons conducted on performance during the dot-probe trials. For dot-probe trials, the variables of interest were participants' accuracy (error rate) and response times (RTs), defined as the latency of participants' responses to the location of the dot-probe after its onset.

Dot-Probe Accuracy. Dot-probe trials were collapsed across the 25, 75, and 125 ms ISIs prior the onset of the dot-probe. Overall, participants completed this task with high accuracy for both Studied targets (M = 82.78%, SD = 20.55%) and Novel targets (M = 85.27%, SD = 22.36%). However, given the emphasis of making these responses as quickly as possible, it was expected that participants would make more mistakes when the dot-probe appears at the opposite side of the screen relative to the object matching their search template (incongruent trials). Thus, differences in error-rates for incongruent, relative to congruent, trials for studied objects would provide evidence for the integration of episodic information in the search template. For Studied targets, the mean error rate for each participant was entered in a 2 (Typicality: Prototypical, Non-typical) x 2 (Dot Location: Congruent, Incongruent) repeated measures ANOVA. The results showed that the main effect of Dot Location was significant, $F(1, 63) = 23.12$, $p < .01$, $\eta^2_p = .25$, in which Incongruent trials produced higher error rates (M = .22, SE = .03) than Congruent trials (M = .13, SE = .02). Additionally, both the main effect of typicality and the Typicality x Dot Location interaction were non-significant (both ps > .05), suggesting that attention was equally captured by the studied objects, regardless of the object's typicality (see Figure 8).
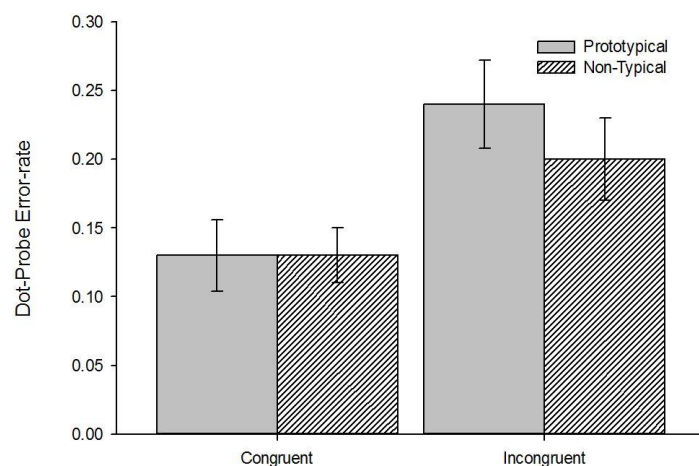

Figure 8. Dot-probe accuracy for studied objects in Exp. 1.

For novel objects, trials were collapsed across Dot Location as congruent and incongruent locations were relative to the prototypical object (i.e., a congruent-prototypical trial was the same as an incongruent-nontypical trial). A paired-samples t-test was conducted to investigate whether attention was captured by a novel but prototypical object relative to a novel-nontypical object. The results revealed no significant differences the two type of targets ($p > .05$; Figure 9).
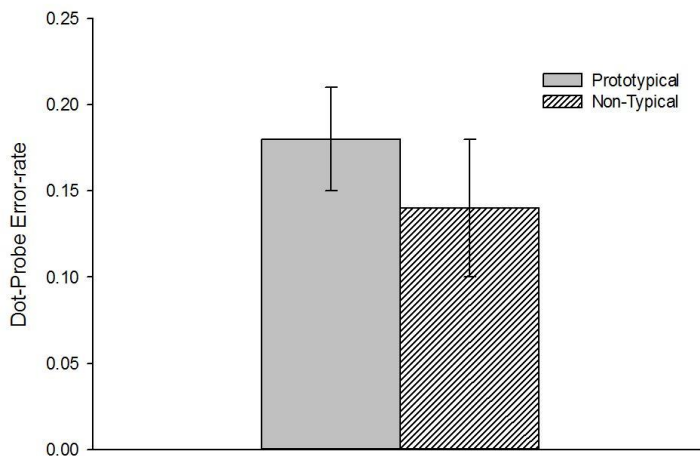

Figure 9. Dot-probe accuracy for novel objects in Exp. 1.

Dot-Probe RT. Only correct dot-probe trials were analyzed. For studied objects, the mean RTs for each participant were entered into a 2 (Typicality: Prototypical, Non-typical) x 2 (Dot Location: Congruent, Incongruent) repeated measures ANOVA. The results showed a significant main effect of Dot Location, $F(1, 59) = 4.59$, $p < .05$, $\eta^2_p = .07$, revealing that participants were slower at responding to the dot-probe when it was presented at the opposite side of the studied object ($M = 346$ ms, $SE = 10$ ms) than when it was presented at the same side as the study object ($M = 330$ ms, $SE = 11$ ms). The main effect of Typicality and the Typicality x Dot Location

interaction were non-significant, reflecting an automatic capture of attention by previously

encoded objects (see Figure 10), regardless of whether the encoded object was a typical or an

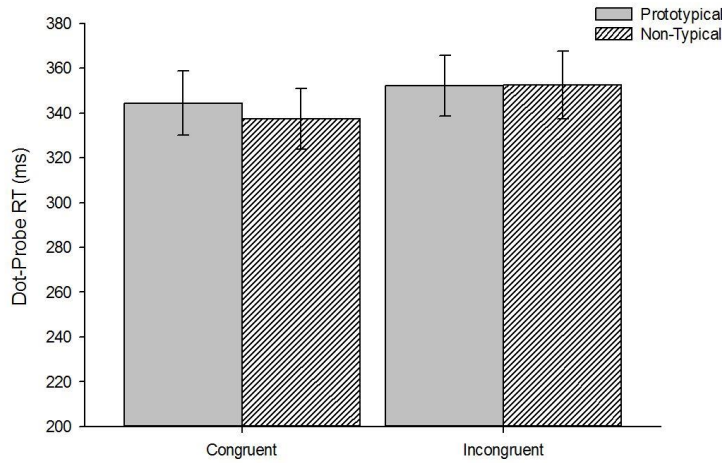atypical exemplar of the target category.

Figure 10. Dot-probe RT for studied objects in Exp. 1.

For novel objects, a paired-samples t-test was computed to examine the possibility of slower

responses for prototypical objects relative to nontypical objects, however, no such difference was
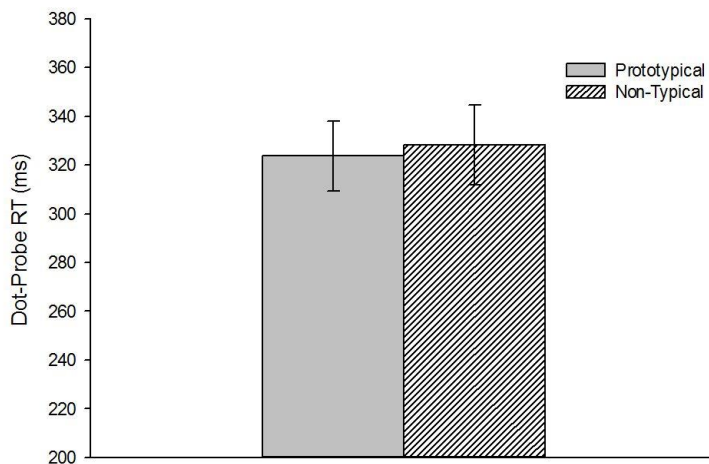
observed (p > .05; Figure 11).

Figure 11. Dot-probe RT for novel objects in Exp. 1.

The goal of Experiment 1 was to examine if observers are able to rely on specific episodic representations, rather than on object typicality, to construct categorical search templates. Observers encoded a set of typical and non-typical objects from various categories and later completed a categorical search task in conjunction with a dot-probe response task. Although the encoding phase and the search phase were independent from each other, observers indeed relied on the representations encoded earlier when presented with a categorical search cue, as they were significantly faster to create search templates, reflected in shorter trial initiation times for studied objects relative to novel objects. More importantly, they were also faster at finding studied objects than novel objects (regardless of the typicality of the encoded object). These interesting effects suggest that episodic memory can be used to identify categorically-diagnostic target features, as observers were not explicitly trained on the different target categories and their exemplars. Moreover, as predicted by the contingent involuntary orienting hypothesis, spatial attention was automatically captured by objects matching the search template during dot-probe trials. Participants made more dot-probe response errors, and were much slower, when the location of a studied object (prototypical or non-typical) was incongruent with the location of the dot-probe. This was not the case for novel objects. These results suggest, that participants successfully encoded the specific exemplars of the target categories during the memory phase, and later when cued with that target category, the features of encoded exemplar were activated relative to any other set of features. Thus, observers relied on episodic memory to form search templates almost exclusively, regardless of the typicality of the encoded object. However, when an episodic representation is not available, observers seem to default to object typicality to construct their search template, given the classic typicality effect observed during search trials. Such gist-based representations are accurate enough to efficiently guide the search,

yet, they are not detailed enough to automatically capture attention during the dot-probe trials. This suggest that observers will rely on single episodic experiences to from categorical search templates when possible, as this type of search templates are more detailed and accurate than those created from semantic memory, resulting in the automatic attention capture and shorter search times.

EXPERIMENT 2

While a category name may activate the most typical category-related features, the same name may also activate a very specific set of features in semantic memory depending on semantic context where the search will be conducted. For example, searching for boots in a jungle likely activates a different set of features than searching for boots in a department store. Moreover, if a categorical target is presented again in a later trial, then the previously found target may serve as the search template for the target category, regardless of the semantic context (Awh, Belopolsky, & Theeuwes, 2012; Huang, Holcombe, & Pashler, 2004). In order to assess the role of semantic context and target repetition in the formation of search templates, observers in Experiment 2 searched for categorically-defined targets using a variation of the scene-preview method (see Hollingworth, 2012; Võ & Wolfe, 2012). The experimental paradigm and design was similar to that of Experiment 1, however, there was no encoding phase nor was MDS data used to quantify the similarity between objects.

Participants. A power analysis (1-β = .95; α = .05; one-tail) conducted on pairwise effect size of dot-probe response times observed in experiment 1 of Reeder, Zoest, & Peelen (2015; N = 12, d = -0.917) suggested a sample size of 15 participants. However, with the goal of increasing the number of observations given the Dot-Probe trial breakdown in the present study (e.g., Studied Prototypical-Inconsistent, Novel NonTypical-Consistent), 67 undergraduate Psychology students participated in exchange for partial course credit. All participants (M_{age} = 18.5 years, SD = 0.94; 42 female) were Native-English speakers with normal or normal-to-corrected vision, and normal self-report color vision.

Stimuli and Apparatus. The experiment was conducted using 21.5-inch Dell monitors with screen resolution of 1920 x 1080 and sampling rates at 60Hz, with all experimental procedures handled by E-Prime 2.0 software (Psychology Software Tools, 2006). A pool of 48 distinct scenes was collected from the internet. For each scene category, a matching target category was identified (i.e., an object category that could be expected to be found in that scene category). For instance, if the scene was football field, then the object category was helmet. Once the match between an object category and each scene was made, four different exemplars for the target category were collected: two exemplars were semantically-consistent with the scene and while the other two were not (they were consistent with a different scene). For football field and helmet, for example, there were two exemplars of a football helmet (consistent), one exemplar of a roman helmet (consistent with the scene coliseum), and one of a military helmet (consistent with the scene military base; see Appendix D). Additionally, 46 "filler" scenes with a single semantically-consistent object category were collected to be used as filler search trials. The same distractor categories, and their respective exemplars, from Experiment 1 were used as distractors for all search trials (these were all semantically unrelated to the target scenes). None of the scenes used depicted objects from either the target or the distractor categories. All object images were resized to 100 x 100 pixels, on average.

The stimuli used in Experiment 2 were normed in a pilot study, in which participants (N = 9) were presented with each scene (e.g., gym) for three seconds. Following the offset of the scene, four exemplars of a target category (e.g., bike) appeared on the screen and participants were asked to pick the two exemplars that match the previous scene the best.

Procedure and Design. Similar to Experiment 1, participants completed 10 practice search trials and 10 practice dot-probe trials prior to the experimental blocks. However, there were only two blocks of trials, each consisting of 96 trials (72 search trials and 24 dot-probe trials). At the beginning of each trial in Block 1, participants were presented with a randomly selected scene prime, along with its corresponding categorical target cue above it, followed 2500 ms later by a prompt asking them to press SPACE when they were ready to initiate the trial (see Figure 12). Similar to Experiment 1, the offset of the scene and target cues was followed by central fixation cross presented for 500ms. On 75% of trials (72 trials) participants searched through an array of 16 objects for the target and pressed SPACE to terminate the search, erasing the search display and replacing the objects with numbers. Afterward, they used the mouse to click the location that contained the target. Targets repeated three times (each time in a different location) and in random order throughout Block 1 (i.e., 24 different search targets were used in total). All search trials were target-present, and all target objects semantically matched the scene that accompanied the target cue (e.g., football helmet when searching for helmet in a football field scene).
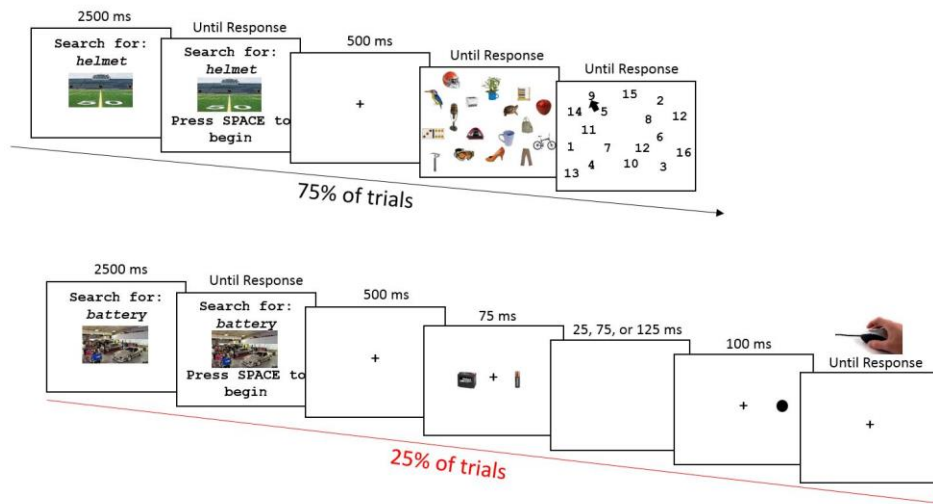

Figure 12. Example of search and dot-probe trials for block 1.

On 25% of trials (24 trials), participants were presented with two object primes from the target category, each at 8.5° of visual angle from either side of the fixation cross, for 75 ms. One of the primes semantically matched the scene while the other did not. For example, if the scene prime was a chemistry lab and the target cue was coat, then one of the primes was a lab coat and the other one was a rain coat. It is important to note that target categories used in dot-probe trials were different than those used in search trials for Block 1. The same dot-probe procedure as in Experiment 1 was adopted, in which the two primes disappeared after 75 ms and a black dot-probe appeared 25, 75, or 125 ms afterward for a duration of 100 ms on either side of the fixation cross. Participants clicked the left or right mouse button as quickly as possible to indicate the location of the dot. In half of the probe trials (12 trials), the dot-probe appeared on the opposite side as the semantically-matching prime (i.e., semantic-matching trials), while in the other half, it appeared in the opposite location of the semantically-mismatching prime (i.e., semantic-mismatching trials; see Appendix B for a trial breakdown). If observers use the scene prime to identify critical target-features, then they should be slower, and make more mistakes, when responding to the dot-probe in semantic-matching than in semantic-mismatching trials. A four second time-penalty was imposed when participants made incorrect responses for both search and dot-probe trials. For dot-probe trials, participants were also prompted to respond more quickly if their response latency was greater than 1500 ms.

Block 2 was similar to Block 1, and it examined whether observers adopted previously found targets as a search template when a categorical search cue is repeated. Search trials followed the same procedure as Block 1. Participants searched for 24 novel targets paired with novel scenes (each one repeated three times) randomized throughout the block (72 search trials total). As for the dot-probe trials, the procedure was same as in Block 1, with the sole difference

42

that the target categories were those used during search trials in the previous block. However, the

categorical target cues were now presented in conjunction with novel scenes from that used in

Block 1. One of the object primes in Block 2 dot-probe trials was semantically related to the

novel scene, while the other was a similar, but not identical, object as the search target from

Block 1 for that target category. For example, if the repeated target cue was helmet (football field

scene in Block 1) and the novel scene was a military base, then one of the primes was an infantry

helmet and the other was a different football helmet from the one used as a search target in Block
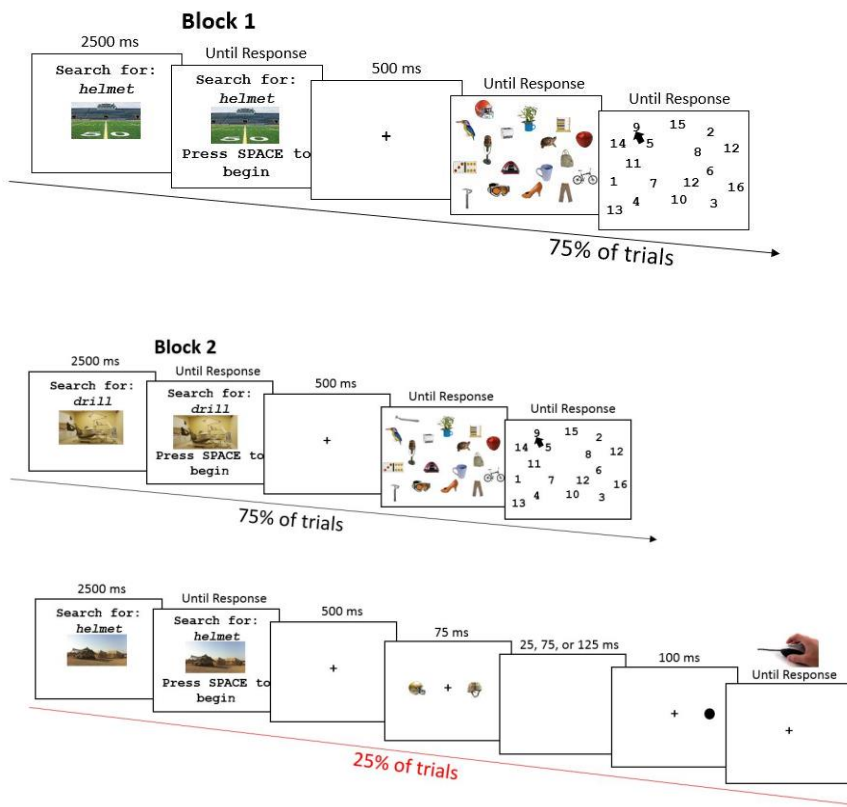
1 (see Figure 13).

Figure 13. Example of search and dot-probe trials for block 2 (bottom) in reference to search
target in block 1 (top).

Scenes were counterbalanced across participants in such way that some participants viewed any given scene (e.g., military base) in Block 1 while others viewed it in Block 2. In half of the dot-probe trials in Block 2 (12 trials) the dot-probe appeared at the opposite location as the "old" target prime from Block 1 (i.e., episodic-matching trials) while in the other half it appeared at the opposite location as the scene-matching prime (i.e., semantic-matching trials). Thus, if observers included the repeated target from Block 1 in their search template during Block 2 dot-probe trials then they should be slower, and make more mistakes, when responding to the dot-probe in episodic-matching than in semantic-matching trials.

Results and Discussion

As in Experiment 1, the significance level for all analyses was set at .05, all multiple comparisons were subjected to Bonferroni corrections, and the same outlier filtering techniques were used prior to data analysis. This resulted in 3.5% of trials Search trials (339 trials) and 1.1% of Dot-Probe trials (36 trials) being dropped across participants, with the remaining 9,309 Search trials and 3,180 Dot-Probe trials included in the analyses.

Trial Initiation. Similar to Experiment 1, trial initiation times for all trials were examined. A paired-samples t-test between trial initiation times in Block 1 and Block 2 revealed no significant difference (p > .05). However, it was observed that trial initiation times were seemingly longer in Experiment 2 than in Experiment 1, which could suggest that, generally speaking, it takes longer for observers to convert scene semantics into search templates than it takes them to retrieve encoded episodic representations. To examine this, trial initiation times for both blocks were collapsed and compared to those observed in Experiment 1. Independent-samples t-tests confirmed this, as trial initiation times for studied objects in Experiment 1 (M = 517 ms, SE = 20 ms) were shorter than those observed in Experiment 2 (M = 661 ms, SE = 41

ms), t(129) = 3.11, p < .01, Cohen's d = .54. However, trial initiation times in Experiment 2 were also longer than those for novel objects in Experiment 1 (M = 538 ms, SE = 22 ms), t(129) = 2.57, p < .01, Cohen's d = .45, which could also suggest that observers simply took longer to disengage from the scene before initiating the trial.

Search Trials. Only correct search trials were analyzed. Overall, search times decreased from Block 1 to Block 2, however, a paired-samples t-test revealed that this difference was not reliable (p > .05). In contrast to trial initiation times, search times for Experiment 2 were remarkably short in comparison to those observed in Experiment 1. Therefore, search times across the two blocks were collapsed and compared to search times for the studied and novel targets in Experiment 1. Independent-samples t-test showed that observers were faster at finding the target when they were semantically primed by a scene (M = 973 ms, SE = 29 ms) relative to novel (M = 1633 ms, SE = 43 ms), t(129) = 13.03, p < .01, Cohen's d =2.29, and to studied targets from Experiment 1 (M = 1324 ms, SE = 38 ms), t(129) = 7.54, p < .01, Cohen's d = 1.32 (see Figure 14).
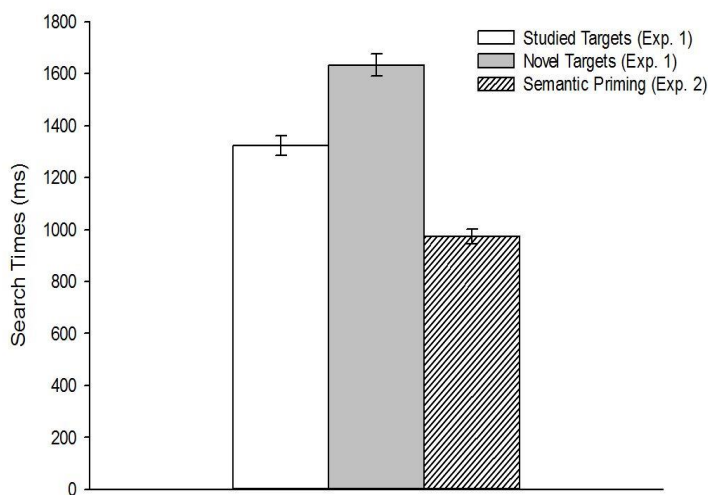

Figure 14. Search times for the studied and novel targets in Experiment 1 and for semantic-priming used in Experiment 2.

45

These results suggest that observers were indeed considering the scene prime as a valid source of information to construct their search templates, and that semantic priming may produce much more detailed templates than those produced from episodic memory, as the set size for the search array remained equal across the two experiments (see General Discussion for an alternative explanation).

Dot-Probe Accuracy. Dot-probe trials were collapsed across the 25, 75, and 125 ms ISIs prior the onset of the dot-probe. The mean error rate for each participant was analyzed separately across the two blocks. For Block 1, a paired-samples t-test revealed that observers made significantly more mistakes during dot-probe trials when the dot appeared at the opposite side as the semantically-matching prime (M = .22, SE = .03) relative to when it appeared at the opposite side as the semantically-mismatching prime (M = .12, SE = .02), t(66) = 5.26, p < .00, Cohen's d = 1.29. This suggests that observers used the scene prime to construct categorical search templates, as targets that semantically-match the scene prime automatically captured attention, relative to objects from the same target category that did not match the scene.

Dot-probe trials in Block 2, on the other hand, assessed the contributions of episodic memory (i.e., target repetition) on the formation of categorical search templates. However, the difference between the error rates for semantic-matching (M = .13, SE = .03) and episodic-matching trials (M = .09, SE = .02) was only marginally significant, t(66) = 2.07, p = .042, Cohen's d = .51 (two-tailed), and in the opposite direction what was expected if repeated targets were included in search templates (see Figure 15). This suggests that, just as in Block 1, attention was strongly captured by objects that matched the semantic content of the scene.
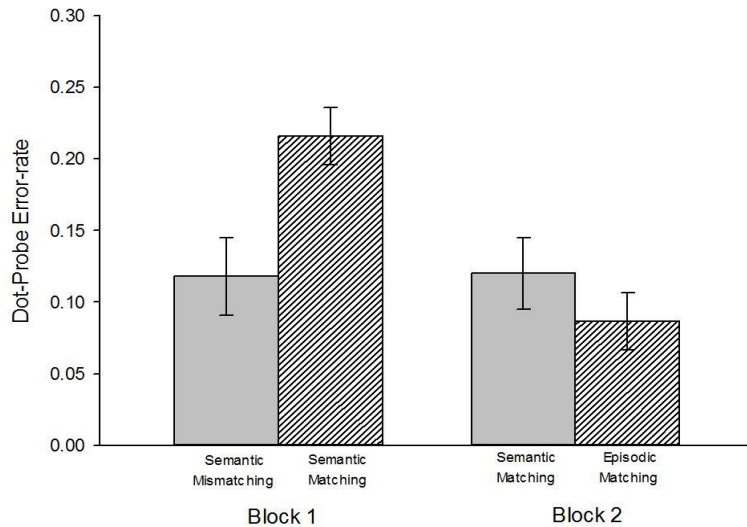
Figure 15. Dot-probe accuracy for Exp. 2.

Dot-Probe RTs. Only correct dot-probe trials were analyzed. The mean RTs for each participant were analyzed separately across the two blocks. For Block 1, a paired-samples t-test revealed that observers were slower to respond to the dot-probe when this appeared at the incongruent side of the semantically-matching prime (M = 394 ms, SE = 22 ms) than when it appeared at the congruent side as the semantically-matching prime (M = 369 ms, SE = 16 ms), however, this difference was only marginally significant, $t(66) = 1.88$, $p = .065$, Cohen's $d = .46$. This reflects, in general, that observers tend to reliably integrate the scene context their search template, as spatial attention was immediately captured by objects matching the scene prime and had to be shifted to the opposite location in order to respond to the dot-probe.

As with error rates, dot-probe trials in Block 2 examined if observers included previously encountered targets as categorical search templates. There was no statically significant difference, however, between episodic-matching and semantic-matching trials in Block 2, $p > .05$ (see Figure 16).
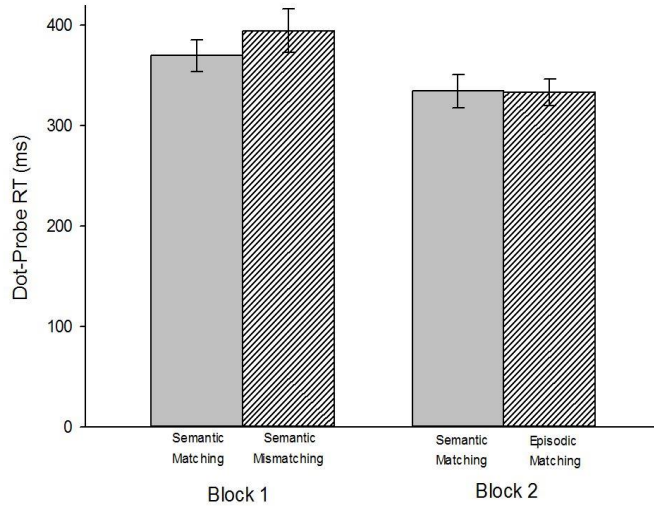
Figure 16. Dot-probe RT for Exp. 2.

The goals of Experiment 2 were to examine if scene semantics can be used to identify categorically-diagnostic target features, and if so, whether or not observers will rely on scene semantics when target cues from previous, but different scenes, are repeated. Observers were provided with target cues accompanied by a scene primes before completing a categorical search task in conjunction with a dot-probe response task. In Block 1, observers integrated the scene prime into their categorical search templates, and although this process seem to take longer, it resulted in a much more detailed search template, as visual search was significantly faster relative to search times observed in Experiment1. Moreover, spatial attention was captured by objects that matched the scene prime during dot-probe trials, as observers made more response errors for incongruent relative to congruent trials. In Block 2, observers were presented with the same target cues as in search trials in Block 1 during dot-probe trials, although cues were accompanied by different scene primes than the ones in the previous block. No statistically significant effect was observed in terms of response errors and response times during dot-probe trials, however, the pattern of result seems to suggest observers did not rely on previously

48

encountered targets to form search templates but continued to use the scene prime instead. These

results from Experiment 2 suggest that when the semantic context is useful, observers

consistently utilize it to inform categorically-diagnostic target features.

GENERAL DISCUSSION

Summarizing, the goals of my master's thesis were 1) to examine if observers can rely on a different mechanism other than semantic memory to form search templates (e.g., episodic memory), and 2) to explain how these two information sources compete with each other to construct an efficient categorical search template. To accomplish these goals, I conducted two experiments adopting an attentional-capture paradigm in conjunction with a categorical search task, which evaluated whether observers rely on internal (Experiment 1) and external (Experiment 2) episodic and semantic cues to identify categorically-diagnostic target features.

In Experiment 1, observers were able to use internal cues drawing from either a recent experience with a category exemplar (e.g., intentional encoding of objects) or semantic memory (e.g., typicality). The typicality effect observed in previous studies was replicated, in which prototypical objects were found much faster than non-typical objects, suggesting that long-term learned statistical regularities of object categories and their features are remarkably influential in categorical visual search. Furthermore, the nature of the categorical cue, whether it came from a studied or novel object category, not only influenced trial initiation times, but also had a large effect on search times (regardless of typicality). Visual search was conducted significantly faster for studied targets relative to novel targets, suggesting that observers can successfully rely on episodic details of objects to form categorical search templates. An alternative explanation, however, is that, given that the studied objects were previously attended during the memory phase, observers' attentional control-settings are attuned to this selection-history (see Awh et al., 2012), and thus studied objects automatically capture attention and "pop-out" among distractors. This might not be the case in the present study, as search times were longer than what pop-out

effects typically reflect. For example, Wang, Cavanagh, and Green (1994) observed pop-out effects for inverted Ns and Zs among normal Ns and Zs to be between 500 ms and 550 ms across different array set size. Usually, pop-out effects are characterized by a parallel-search of the entire visual array, in which very shallow search slopes are observed, as search times remain constant regardless of the set size. The present study did not manipulate set size, however, the search times observed for studied objects (1280-1360 ms) were too long for the target object to stand out among distractors. Moreover, trial initiation times were significantly shorter when target cues related to studied relative to novel objects, which suggest that observers were quickly retrieving objects from the study phase to use as categorical search templates.

Additionally, evidence from the dot-probe trials suggests that the contents of VWM reflect an episodic-based search template. Observers made significantly more response errors when the location of an encoded object was incongruent with the location of the dot-probe, given the task pressure of making a response as quickly as possible. This was not the case for novel target categories and their exemplars, which suggests that when presented with studied category cues, episodic representations for the target objects are strongly activated. Moreover, observers were also much slower at responding to the dot-probe when it was opposite to an encoded object. However, it is important to mention that the size of this effect was relatively small. On average, Reeder and colleagues (2013, 2015) observed RT differences for congruent and incongruent dot-probe trials ranging from 20 to 45 ms. The RT difference in the present study was much shorter (16 ms), although it is possible that the same task-demands of making quick responses, which evoked the response error effects, might have reduced the difference in RTs. Future research using eye-tracking may support the evidence found in the present investigation.

In Experiment 2, observers were able to use contextual information from the search display to form highly detailed categorical search templates. More importantly, in Block 2 of Experiment 2, observers had the opportunity to use previously encountered targets as search templates or to continue to rely on the semantic context of the display to do so. The results show that observers, when given the opportunity, incorporate the search context into their search template. In Block 1, observers made more response errors, and were relatively slower, when the location of objects that were semantically related to the scene primes were incongruent with the location of the dot-probe. Moreover, when target cues later repeated in Block 2, observers continued rely on semantic primes presented at the beginning of each trial to form their search template, rather than incorporate features of targets found in Block 1, as they made relatively more response errors for the semantically related prime. However, an alternative explanation is that observers were not able to incidentally encode target objects in Block 1, even after three repetitions, due to the lack of meaningful spatial associations between the objects in the search array. For instance, Draschkow, Wolfe, and Võ (2014) reported higher recall rates for target objects found when search through real-world scenes than when participants were asked to intentionally memorize the objects in those scenes. This counterintuitive effect, however, was not observed when the object array was embedded in neutral background (e.g., brick wall), which suggests that the meaningful associations between objects in the array facilitates incidental encoding. Since the present study used search arrays with essentially no meaningful association between objects (white background with randomly located objects), the findings of Draschkow et al. (recently replicated by Josephs et al., 2016) suggests that it is possible that observers in the present study were not successful at incidentally encoding the targets in LTM. It is also important to note that the "repeated" exemplars used in dot-probe trials of Block 2 were not

identical to those used in the search trials of Block 1. This was purposefully done to encourage the inclusion of LTM representations for those exemplars in the search template rather than attention being capture due to repetition-priming. Still, given the surprisingly fast search produced by semantic priming, and the relatively higher error-rate for semantically-matching objects during the dot-probe trials, it is more plausible that observers integrated the semantic content of the scene into their search template, without perhaps even probing their memory for previously encountered targets, as the scene prime provided enough information to quickly identify the categorically-diagnostic target features.

Another important aspect that possibly influenced the continued reliance on the semantic context over LTM representations in Experiment 2 was that semantic primes were always valid during search trials, as the search targets was always related to the scene. If, however, semantic primes stopped being a reliable source of information it is possible that observers would have change strategies and turned to their LTM representations for the incidentally encoded targets to form search templates. In fact, recent research by Bravo and Farid (2016) suggests that when observers undergo extensive training to remember target exemplars of an object category (e.g., watches), they are able to switch between each representation depending on the search context (type of distractor make-up of the search array) when the context is cued before the search trial. However, it is possible that observers cannot switch strategies when only incidentally encoding target objects, as the LTM representations may not be retrievable due to retroactive interference. Future research can investigate this possibility by directly manipulating the validity of the semantic prime during later search trials by simply varying the proportion of valid and invalid semantic primes in Block 2 of Experiment 2, relative to a group of observers who receive only valid semantic primes throughout the experiment.

As noted before, the results of Experiment 2 suggest that observers can and will incorporate scene semantics into categorical search templates. What was unexpected from the present investigation was the efficiency of this process. Compare to observers in Experiment 1, observers in Experiment 2 took longer to initiate each trial, but were much faster at searching for categorical targets when relying on scene primes to identify critical target-features, even though observers in Experiment 1 had previously seen (and intentionally encoded) the target objects. These results could reflect a trade-off between the latency to create search templates and the precision of such representations, with more detailed categorical templates taking longer to be constructed. However, it is also plausible that observers in Experiment 2 took longer to initiate each trial because they had to disengage from a visually-vivid image, regardless of whether or not they had already created their search template. Of more interest are the search time differences across the two experiments. There are two potential explanations for such results. The first, of more practical importance, is that the target stimuli used in Experiment 1 came from the same database as the distractor stimuli, while in Experiment 2, the target stimuli were downloaded from stock images on the internet and the distractor stimuli was not (same stimuli as in Experiment 1). The target stimuli in Experiment 2 was resized to match the distractor stimuli, however. Given that the same exact distractor categories and exemplars were used in both experiments, it is possible that the target-distractor similarity (Duncan & Humphreys, 1989) in stimuli characteristics in Experiment 1 might have made the search harder relative to Experiment 2. However, such explanation would warrant much shorter search times for Experiment 2, where they seem long enough (973 ms) that the "pop-out" account proposed for Experiment 1 may not by compatible here either. Nonetheless, it is possible to test this explanation empirically by using the same target and distractor stimuli in a between-subjects design, in which one group of

observers encode objects and later search for them while another group of observers are semantically primed to search for those same exact objects. This between-groups design would allow for the direct comparison across conditions to ensure that the present effect is not due to the stimuli differences.

The second explanation, of more theoretical importance, is that semantic priming, relative to episodic retrieval, is an inherently more efficient mechanism due to the large network of semantic associations learned through the lifespan. This may allow the observer quickly produce and upload a highly detailed search template onto VWM. On the other hand, it is also possible that semantic priming does not activates "complete" search templates but instead it only activates the specific features of an object category associated with the semantic prime. In the case of football field and helmet, only the cage-like mask characteristic of all football helmets is activated, rather than a single and concrete representation of a football helmet (as would be expected if one intentionally retrieves a specific target exemplar from episodic memory). This means that possibly, scene semantics allows the observer to distinguish an object from other subordinate-level objects of the same category. Such distinction is not possible when only categorical (verbal) information about the target object is available. Future research should investigate such account in relation to a non-categorical template search, as it would provide a baseline for search performance, as well as for the latency for "setting-up" the search template, too ensure that these data are not a result of the current experimental design.

It has been previously noted that, although current models of visual search are able to account for the efficiency of categorical searches, they should also explain the mechanisms by which observers identify categorically-diagnostic target features (Levin et al., 2001). Recent research on categorical visual search has focused on demonstrating the precision and

effectiveness of word cues in allowing observers to form highly detailed search templates necessary to find targets (e.g., Maxfield & Zelinsky, 2012; Schmidt & Zelinsky, 2009). Additionally, search times, time-to-target and verification, are shorter when target objects are typical exemplars of the target category (Maxfield, Stalder, & Zelnisky, 2014).The present study expanded upon this previous research and investigated the mechanisms involved in search template formation when episodic information about the categorical target is available to observers. It has been demonstrated that through training observers are able to identify categorically-diagnostic target features to build flexible, but specific, search templates in preparation for search tasks (Bravo & Farid, 2012). However, the present study supported the idea that observers are able to construct highly detailed search templates in the absence of training if specific experiences with target categories are available. More importantly, these specific experiences with the target categories and their exemplars came from a different, supposedly unrelated memory test, and not from the search task itself. This implies that when a recent experience with the target category is readily available to be incorporated into the search template, observers will prefer to do so, possibly because it produces a more accurate search template.

Similar to the impact of episodic memory in categorical search, the effects of the semantic context of the search display on search template formation have not been directly investigated in the literature. The present study aimed to address this and examined how observers incorporate information from the search environment into their search template. Furthermore, the present study also tested whether observers use episodic information from previous searches regardless if such information is inconsistent with the semantic context of the search display. However, the present study suggests that observers continuously rely on scene-

semantics to identify categorically-diagnostic target features, regardless of information retrieved from previous searches. This perhaps occurs because semantic priming is much faster, and potentially less effortful, than memory retrieval, especially when the semantic context of the search is always a reliable source of information.

Lastly, it has also been recently proposed that differences in task performance caused by internally generated visual representations (i.e., representations formed in the absence of physical visual stimulation), such as in categorical visual search, may be influenced by individual differences in visual imagery, expertise with different object categories, and occipital lobe anatomy (Reeder, in press). The present study cannot attest to this account, however, it is possible that the results observed here might have been influenced by individual differences in memory retrieval abilities (see Brewer & Unsworth, 2012). Future research on the topic of individual differences in visual processing is necessary to elucidate such account.

In conclusion, this investigation provided some insights into the mechanisms observers use to identify categorically-diagnostic target features, which have not been directly addressed by models of visual search (e.g., Duncan & Humphreys, 1989; Wolfe et al., 1989, 2007). The findings from this thesis reflect a more comprehensive understanding of how categorical search templates are formed as well as when different sources of information (e.g., episodic and semantic memory) are preferred by observers to construct an efficient categorical search template.

REFERENCES

Alexander, R. G., & Zelinsky, G. J. (2011). Visual similarity effects in categorical search. *Journal of Vision*, *11*, 1–15.

Awh, E., Belopolsky, & Theewes, J. (2012). Top-down versus bottom-up attentional control: A failed theoretical dichotomy. *Trends in Cognitive Sciences*, *16*(8), 437-443.

Bacon, W. F., & Egeth, H. E. (1994). Overriding stimulus-driven attentional capture. *Perception & Psychophysics*, *55*, 485-496.

Beck, V. M., Hollingworth, A., & Luck, S. J. (2012). Simultaneous control of attention by multiple working memory representations. *Psychological Science*, *23*(8), 887-898.

Bravo, M. J., & Farid, H. (2009). The specificity of the search template. *Journal of Vision*, *9*, 1–9.

Bravo, M. J., & Farid, H. (2012). Task demands determine the specificity of the search template. *Attention, Perception & Psychophysics*, *74*, 124–131.

Bravo, M. J. & Farid, H. (2016). Observers change their target template based on expected context. *Attention, Perception, & Psychophysics*, *78*, 829-837.

Brewer, G. A., & Unsworth, N. (2012). Individual differences in effect of retrieval from long-term memory. *Journal of Memory and Language*, *66*, 407-415.

Castelhano, M. S., Pollatsek, A., & Cave, K. (2008). Typicality aids search for an unspecified target, but only in identification, and not in attentional guidance. *Psychonomic Bulletin & Review*, *15*, 795–801.

Castelhano, M. S., & Heaven, C. (2010). The relative contribution of scene context and target features to visual search in scenes. *Attention Perception & Psychophysics*, *72*(5), 1283–1297.

Chun M.M., & Jiang, Y. (1998). Contextual cuing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology*, *36*, 28–71.

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, *96*(3), 433–458.

Evans, K. K., & Treisman, A. (2005). Perception of objects in natural scenes: is it really attention free? *Journal of Experimental Psychology: Human Perception and Performance*, *31*(6), 1476–1492.

Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception & Performance*, *18*, 1030-1044.

Folk, C. L., Leber, A. B., & Egeth, H. E. (2002). Made you blink! Contingent attentional capture produces a spatial blink. *Perception & Psychophysics*, *64*(5), 741-753.

Hollingworth, A., & Henderson, J. M. (2002). Accurate visual memory for previously attended objects in natural scenes. *Journal of Experimental Psychology: Human Perception & Performance*, *28*(1), 113-136.

Hollingworth, A. (2012). Task specificity and the influence of memory on visual search: Comments on Võ and Wolfe (2012). *Journal of Experimental Psychology: Human Perception and Performance*, *38* (6), 1596-1603.

Horowitz, T. S., & Wolfe, J. M. (1998). Visual search has no memory. *Nature*, *395*(6693), 575-577.

Hout, M. C., & Goldinger, S. D. (2010). Learning in repeated visual search. *Attention, Perception, & Psychophysics, 72,* 1267–1282.

Hout, M.C., Papesh M.H., Goldinger S.D. (2013). Multidimensional scaling. *Wiley Interdisciplinary Review Cognitive Science*, *4*, 93–103.

Hout, M. C. & Goldinger, S. D. (2014). Target templates: The precision of mental representations affects attentional guidance and decision-making in visual search. *Attention, Perception, & Psychophysics*, *77* (1), 128-149.

Hout, M. C., Goldinger, S.D., & Brady, K.J. (2014) MM-MDS: A multidimensional scaling database with similarity ratings for 240 object categories from the massive memory picture database. *PLoS ONE*, *9*(11): e112644. doi:10.1371/journal.pone.0112644

Huang, L., Holcombe, A. O., & Pashler, H. (2004). Repetition priming in visual search: Episodic retrieval, not feature priming. *Memory & Cognition*, *32*(1), 12-20.

Josephs, E. L., Draschkow, D., Wolfe, J. M., & Võ, M. L. H. (2016). Gist in time: Scene semantics and structure enhance recall of searched objects. *Acta Psychologica*, *169*, 100-108.

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.

Levin, D. T., Takarae, Y., Miner, A. G., & Keil, F. (2001). Efficient visual search by category: Specifying the features that mark the difference between artifacts and animals in preattentive vision. *Perception & Psychophysics*, *63*(4), 676-697.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*, 279-281.

Malcolm, G. L., & Henderson, J. M. (2009). The effects of target template specificity on visual search in real-world scenes: Evidence from eye movements. *Journal of Vision*, *9*, 1–13.

Maljkovic, V, & Nakayama, K. (1994). Priming pop-out: I. Role of features. *Memory and Cognition*, *22*, 657-672.

Maxfield, J. T., & Zelinsky, G. J. (2012). Searching through the hierarchy: How level of target categorization affects visual search. *Visual Cognition*, *20* (10), 1153–1163.

Maxfield, J. T., Stalder, W.D., & Zelinsky, G. J. (2014). Effects of target typicality on categorical search. *Journal of Vision*, *14* (12), 1-11.

van Moorselaar, D., Theeuwes, J., & Olivers, C. N. L. (2014). In competition for the attentional template: Can multiple items within visual working memory guide attention? *Journal of Experimental Psychology: Human Perception and Performance*, *40*(4), 1450-1464.

Murphy, G. L. & Brownell, H. H. (1985). Category differentiation in object recognition: Typicality constraints on the basic category advantage. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*, 70-84.

Olivers, N. C., Peters, J., Houtkamp, R., & Roelfsma, P. R. (2011). Different states in visual working memory: When it guides attention and when it does not. *Trends in Cognitive Sciences*, *15*, 327-334.

Pansky, A. & Koriat, A. (2004). The basic-level convergence effect in memory distortions. *Psychological Science*, *15*(1), 52-59.

Peterson, M. S., Kramer, A. F., Wang, R. F., Irwin, D. E., & McCarley, J. S. (2001). Visual search has memory. *Psychological Science*, *12*, 287-292.

Psychology Software Tools, E-Prime Version 2.0 (2006).

Reeder, R. R., & Peelen, M. V. (2013). The contents of the search template for category-level search in natural scenes. *Journal of Vision*, *13* (3), 1-13.

Reeder, R. R., Zoest, W. van, & Peelen, M. V. (2015) Involuntary attentional capture by task-irrelevant objects that match the search template for category detection in natural scenes. *Attention, Perception, & Psychophysics*, *77* (4), 1-11.

Reeder, R. R. (in press). Individual differences shape the content of visual representations. In-press at: *Vision Research*.

Robbins, A., & Hout, M. C. (2015). Categorical templates: Typical category members are found and identified quickly during word-cued search. *Summary published in* Object Perception, Attention, and Memory (OPAM) 2015 Conference Report, *Visual Cognition*. doi: 10.1080/13506285.2015.1093247.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic Objects in Natural Categories. *Cognitive Psychology, 8*, 382-439.

Schmidt, J., & Zelinsky, G. J. (2009). Search guidance is proportional to the categorical specificity of a target cue. *The Quarterly Journal of Experimental Psychology*, *62*, 1904–1914.

Spotorno, S., Malcolm, G. L., & Tatler, B. W. (2014). How context information and target information guide the eyes from the first epoch of search in real world scenes. *Journal of Vision*, *14*(2), 1–21.

Spotorno, S., Malcolm, G. L., & Tatler, B. W. (2015). Disentangling the effects of spatial inconsistency of targets and distractors when searching in realistic scenes. *Journal of Vision*, *15*(2), 1-21.

Sun, S. L., Shen, J., Shaw, M., Cant, J. S., & Ferber, S. (2015). Automatic capture of attention by conceptually generated working memory templates. *Attention, Perception, & Psychophysics*, *77*, 1841-1847.

Torralba, A., Oliva, A., Castelhano, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. *Psychological Review*, *113*(4), 766-786.

Ullman, S., Vidal-Naquet, M., & Sali, E. (2002). Visual features of intermediate complexity and their use in classification. *Nature Neuroscience*, *5*(7), 682–687.

Vickery, T. J., King, L. W., & Jiang, Y. (2005). Setting up the target template in visual search. *Journal of Vision*, *5*(1):8, 81–92.

Võ, L. H. & Wolfe, J. M (2012). When does repeated search in scene involve memory? Looking at versus looking for objects in scene. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(1), 23-41.

Võ, M. L. H. & Wolfe, J.M. (2013). The interplay of episodic and semantic memory in guiding repeated search in scenes. *Cognition*, *126*, 198–212.

Võ, M. L. H. & Wolfe, J. M (2015). The role of memory for visual search in scene. *Annals of the New York Academy of Sciences*, *1339*, 72-81.

Wang, Q., Cavanagh, P., & Green, M. (1994). Familiarity and pop-out in visual search. *Perception & Psychophysics*, *56*(5), 495-500.

Williams, C. C., Henderson, J. M., & Zacks, R. T. (2005). Incidental visual memory for targets and distractors in visual search. *Perception & Psychophysics*, *67*(5), 816-827.

Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology. Human Perception and Performance*, *15*(3), 419–433.

Wolfe, J. M., Klempen N., & Dahlen K. (2000). Post-attentive vision. *Journal of Experimental Psychology: Human Perception & Performance*, *26*(2), 693–716.

Wolfe, J. M., Butcher, S. J., Lee, C., & Hyle, M. (2003). Changing your mind: On the contribution of top-down and bottom-up guidance in visual search for feature singletons. *Journal of Experimental Psychology: Human Perception and Performance*, *29*, 483–502.

Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews: Neuroscience*, *5*, 1–7.

Wolfe, J. M. (2007). Guided Search 4.0: Current progress with a model of visual search. In W. Gray (Ed.), *Integrated models of cognitive systems* (pp. 99–119). New York: Oxford.

Wyble, B., Folk, C., & Potter, M. C., (2013). Contingent attentional capture by conceptually relevant images. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(3), 861-871.

Yang, H., & Zelinsky, G. J. (2009). Visual search is guided to categorically-defined targets. *Vision Research*, *49*(16), 2095–2103.

Appendix A. Table summarizing the trial breakdown collapsed across the 6 blocks of trials in Experiment 1.

Total trials in Experiment 1: 192	Search Trials (75%, 144 trials)	Dot-Probe Trials (25%, 48 trials)
Search Target (Studied-Nontypcial)	36	n/a
Search Target (Studied-Prototypical)	36	n/a
Search Target (Novel-Nontypical)	36	n/a
Search Target (Novel-Prototypical)	36	n/a
Dot-Probe (Studied Proto-Incongruent)	n/a	6
Dot-Probe (Studied NonTyp-Incongruent)	n/a	6
Dot-Probe (Studied Proto-Congruent)	n/a	6
Dot-Probe (Studied NonTyp-Congruent)	n/a	6
Dot-Probe (Novel Proto-Incongruent)	n/a	6
Dot-Probe (Novel NonTyp-Incongruent)	n/a	6
Dot-Probe (Novel Proto-Congruent)	n/a	6
Dot-Probe (Novel NonTyp-Congruent)	n/a	6

Appendix B. Table summarizing the trial breakdown across the two blocks in Experiment 2.

Total trials in Experiment 2: 192	Search Trials (75%, 144 trials)	Dot-Probe Trials (25%, 48 trials)
Block 1 total trials: 96		
Repeated Search Target (24 x 3)	72	n/a
Dot-Probe (Semantic-Matching)	n/a	12
Dot-Probe (Semantic-Mismatching)	n/a	12
Block 2 total trials: 96		
Repeated Search Target (24 x 3)	72	n/a
Dot-Probe (Episodic-Matching)	n/a	12
Dot-Probe (Semantic-Matching)	n/a	12
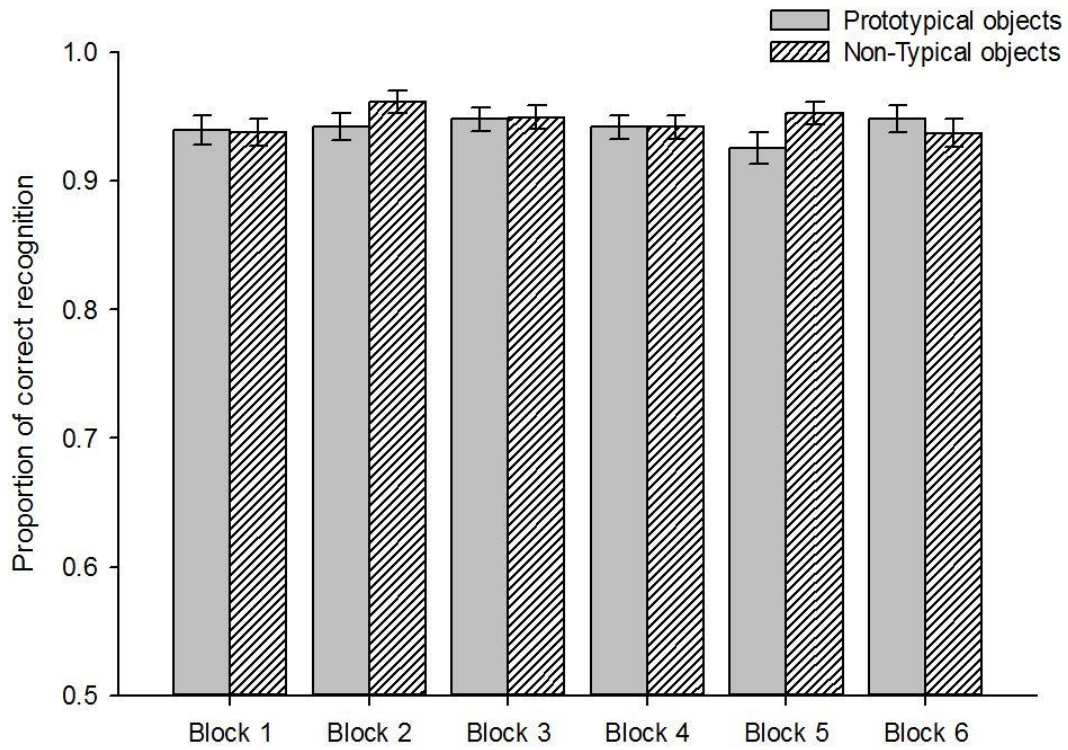
Appendix C. Proportion of correct 2-AFC recognition across trial blocks in Experiment 1.

APPENDIX D

Appendix D. Object categories, scene categories, and object exemplars used as target stimuli in Experiment 2.

Target Category	Scene	Matching Object	Mismacthing Object
cake	wedding	wedding cake	birthday cake
glass	fancy dinner	wine glass	pint glass
lamp	living room	floor lamp	gas lamp
knife	kitchen	kitchen knife	combat knife
fish	room	goldfish	swordfish
bear	iceberg	polar bear	teddy bear
mouse	mouse-trap	mouse	computer mouse
battery	car shop	car battery	aa battery
plant	desert	cactus	tulip
razor	barber shop	barbershop razor	gillette
brush	bathroom sink	toothbrush	paint brush
drill	dentist office	dentist drill	construction drill
snack	movie theater	popcorn	orange
boat	river	row boat	yacht
key	door entrance	key	computer key
jacket	ski slope	ski jacket	motorcycle jacket
truck	boys room	toy truck	real truck
butterfly	garden	real butterfly	butterfly clip
shot	bar	shot glass	medical shot
pin	sewing	sewing pin	push pin
skirt	st pattys	irish skirt	club skirt
pick	music store	guitar pick	pickax
weight	gym	barbell	paper weight
food	baseball game	hotdog	sushi
helmet	football field	football helmet	roman helment
boots	jungle	jungle boots	rain boots
mask	operating room	surgeon mask	wrestling mask
hat	baseball field	baseball hat	beach hat
ball	soccer field	soccer ball	basketball
pet	dog house	dog	cat
glasses	library	eyeglasses	sunglasses
shoes	office	shoes	bowling shoes
bell	classroom	school bell	church bell
gun	woods	hunting gun	water gun
car	driveway	sedan	golf car
goggles	skii mountain	skii googgles	swimming goggles
coat	chemistry lab	lab coat	rain coat
eggs	diner	eggs cooked	eggs
clock	living room	grandfather clock	classroom clock
flag	mount rushmore	american flag	football flag
paper	breakfast	newspaper	toilet paper
clip	pantry	chip clip	hair clip
glove	doctors office	latex glove	baseball glove
speaker	car interior	car speaker	concert speaker
stick	hockey rink	hockey stick	pool stick
bike	road	road bike	dirt bike
backpack	hiking path	hiking backpack	backpack
telephone	street sidewalk	payphone	cellphone

APPENDIX E

ACTION ON EXEMPTION APPROVAL REQUEST

LSU

Institutional Review Board
Dr. Dennis Landin, Chair
130 David Boyd Hall
Baton Rouge, LA 70803
P: 225.578.8692
F: 225.578.5983
irb@lsu.edu | lsu.edu/irb

TO: Megan Papesh
Psychology

FROM: Dennis Landin
Chair, Institutional Review Board

DATE: September 6, 2016

RE: IRB# E10027

TITLE: Visual Search Study

New Protocol/Modification/Continuation: New Protocol

Review Date: 9/6/2016

Approved_____X_____ Disapproved_____

Approval Date: 9/6/2016   Approval Expiration Date: 9/5/2019

Exemption Category/Paragraph: 2a

Signed Consent Waived?: No

Re-review frequency: (three years unless otherwise stated)

LSU Proposal Number (if applicable):

Protocol Matches Scope of Work in Grant proposal: (if applicable)

By: Dennis Landin, Chairman _____

PRINCIPAL INVESTIGATOR: PLEASE READ THE FOLLOWING –
Continuing approval is CONDITIONAL on:
1. Adherence to the approved protocol, familiarity with, and adherence to the ethical standards of the Belmont Report, and LSU's Assurance of Compliance with DHHS regulations for the protection of human subjects*
2. Prior approval of a change in protocol, including revision of the consent documents or an increase in the number of subjects over that approved.
3. Obtaining renewed approval (or submittal of a termination report), prior to the approval expiration date, upon request by the IRB office (irrespective of when the project actually begins); notification of project termination.
4. Retention of documentation of informed consent and study records for at least 3 years after the study ends.
5. Continuing attention to the physical and psychological well-being and informed consent of the individual participants, including notification of new information that might affect consent.
6. A prompt report to the IRB of any adverse event affecting a participant potentially arising from the study.
7. Notification of the IRB of a serious compliance failure.
8. SPECIAL NOTE: When emailing more than one recipient, make sure you use bcc. Approvals will automatically be closed by the IRB on the expiration date unless the PI requests a continuation.

*All investigators and support staff have access to copies of the Belmont Report, LSU's Assurance with DHHS, DHHS (45 CFR 46) and FDA regulations governing use of human subjects, and other relevant documents in print in this office or on our World Wide Web site at http://www.lsu.edu/irb

Appendix F. Copy of Institutional Review Board approval for Human Subject Research.

VITA

Juan D. Guevara Pinto is a third year student in the Cognitive and Brain Sciences doctoral

program at Louisiana State University, Baton Rouge, Louisiana. He received his bachelor's

degree in psychology from Georgia Southern University, Statesboro, Georgia, in 2014.