Categorically distinct subsets allow flexible memory-selection in hybrid search

Nurit Gronau[1], Makaela Nartker[2], Sharon Yakim[1,3], Igor Utochkin[4] & Jeremy Wolfe[56,]*

[1]The Open University of Israel, Israel

[2]Johns Hopkins University, USA

[3]The Hebrew University of Jerusalem, Israel

[4]University of Chicago, USA

[5]Brigham and Women's Hospital and [6]Harvard Medical School, USA

*Corresponding author:
Prof. Nurit Gronau
Psychology & Education Department
Cognitive Science Studies
The Open University of Israel
1 University Rd., Raanana, 4353701, PO box 808, Israel.
nuritgro@openu.ac.il

Acknowledgements

We would like to thank Mimi Shwartz, Moshe Shinfeld and Shunit Cohen for assisting in data collection. This study was supported by the Israeli Science Foundation (ISF) grant 1622/15 (to NG) and by the National Institutes of Health (NIH) grant EY017001 (to JMW).

Abstract

In many every-day situations, we search our visual surroundings for any one of many memorized items held in memory, a process termed *hybrid search*. In some cases, only a portion of the memorized mental list is relevant within a specific visual context, thus, restricting memory search to the relevant subset would be desirable. Previous research had shown that participants largely fail to ‘partition’ memory into several distinct subsets, on a trial-by-trial basis. However, given the known role of semantic content in long-term memory organization, we hypothesized that semantically-defined subset categories might serve as a more powerful means for flexible memory-selection in dynamic hybrid search situations. Experiment 1 revealed that, indeed, semantic characteristics (i.e., object category), but not perceptual features (e.g., arbitrary color), can provide a firm basis for flexible memory partitioning. Experiments 2 and 3 further showed that such memory partitioning is costless and is independent of the nature of the surrounding visual distractors (i.e., a categorically-homogeneous or heterogeneous display). These findings demonstrate that confining one’s memory search to a currently relevant subset of items is highly effective when the different memory subsets are defined by clear semantic categories. The results underscore the importance of conceptual information in the organization of *activated long-term memory* (aLTM), and in forming the basis for a flexible, trial-by-trial, memory-selection. Our findings further highlight the relationship between visual search and memory search, and they may shed light on the processes contributing to a successful construction of bounded episodes in long-term memory.

Keywords: Hybrid search, memory search, visual search, memory selection, categorical processing, activated long-term-memory (aLTM), event memory

Introduction

Many every-day tasks involve an ongoing interaction between visual search and memory search. When visually searching the closet for items needed for an upcoming trek, we may also be searching memory through a preplanned ‘mental list’ of our outdoor clothes. This type of combined visual search (search the closet) and memory search (scanning a list in memory) is known as *hybrid search* (see, e.g., Schneider & Shiffrin, 1977). Visual and memory search may seem at first sight similar, yet different mechanisms likely underlie these two processes. While it has long been known that latencies in visual search increase roughly linearly with the number of items in the display, Wolfe (2012) has shown that reaction times in memory search increase logarithmically with the number of items in memory: adding a second item to a memory-set costs more than adding the tenth or even the 100th item to that same memory list.

Often, at least in the lab, all or most of the memorized items in one’s mental list are relevant to the current visual search, but in some cases only a subset of them will be relevant in a specific search context. In Figure 1 (left panel), imagine that you are searching for a memorized list of necessary camping gear, kept in the basement (A) as well as your set of hiking clothes, in the closet upstairs (B). If you are holding in mind two or more subsets of items or lists, is it possible to restrict memory search only to the currently relevant subset? That is, if you are in the basement (right panel), can you restrict memory search to the basement subset, or does the learned global "trekking" list automatically trigger a search of the entire memory set, even if there are no clothes in the basement? Transitioning back to the laboratory setting, what happens when different location contexts, or task-relevant cues, shift continuously between trials, and how efficiently can memory search adapt to current task demands? This research endeavors to investigate whether and how we can *flexibly* shift between various memory subsets when task-relevant contexts frequently change.

Figure 1: *Left - Memory search*: If your mental list contains both camping gear (A) and hiking clothes (B), is it possible to restrict your memory search to the gear if you are in the basement, where no clothes are stored? *Right - Visual search*: When searching the basement, what are the properties that may guide visual search, and how do they interact with memory search? From a theoretical standpoint, how flexible is memory search through different subsets (lists), when contexts (e.g., locations, or other task-relevant cues) constantly shift? See elaboration in text below.

Previous hybrid search studies have shown that we can search for dozens of specific targets types at the same time ('search for this tent, this shirt, these sunglasses', etc., Wolfe, 2012), and we can also hold in mind several simultaneous item *sets*, or categories (e.g., 'search for any camping equipment, hiking clothes, or trekking accessories', e.g., Cunningham & Wolfe, 2014). Unsurprisingly, search performance is slower in the latter case, due to the use of more broadly defined search criteria (e.g., Cunningham & Wolfe, 2014; Wolfe et al., 2017). Less known, however, is whether guidance towards multiple lists can be performed in an adaptive manner, allowing us to *confine* memory search to a particular list or a portion of it, while discarding other portions, as a function of trial context or memory-set content. Namely, if one memorizes a list containing several 'sublists' or categories, is it possible to select only a specific sublist on a certain trial, and to flexibly shift between the different sublists on a trial-by-trial basis, thus allowing a more efficient memory search?

Moreover, the interplay between visual search and memory search under such conditions warrants deeper investigation, as, by large, the two types of processes are guided and influenced by different factors. When searching the visual field for particular items belonging to any of several sets or categories, perceptual attributes characterizing these categories, if available, would typically guide visual attention (e.g., Hout et al., 2017; Nako et al., 2014; Maxfield & Zelinsky, 2012; Wolfe, 2021; Yang & Zelinsky, 2009; Yu et al., 2023). Thus, if you are in the basement, attention will be guided to objects that have basic features that might be consistent with what you are looking for (e.g., colorful equipment such as #1, #3 or #5 in

Figure 1, right). Once identified as skis, #5 could be dismissed as categorically incorrect, since it is not a camping or a trekking equipment. Other items like the wrench (#2) or the rake (#4) might not attract attention at all because they have the 'wrong' basic features. Finally, some items (e.g. #1) have features consistent with one's search, and once attended, they prove to belong to the correct category, but are they part of the required list? It is at this point that it becomes necessary to probe memory and ask if the camping cooler was, in fact, part of the gear you were looking for. The central question of this paper is whether the memory list that is probed includes only the relevant subset of camping gear items (A), or whether you are obligated to search the entire memory set (A and B) before concluding that the cooler is in fact irrelevant to your current goals. If it is possible to select a subset, how easy is it to *switch* memory search through one subset with another if the item in question shifts from being a piece of gear to being a piece of clothing? As previously noted, this question is compelling not only for its exploration of the efficacy of memory search processes but also for its broader implications concerning our cognitive flexibility in adapting to a dynamic visual environment.

Importantly, Wolfe (2012) and subsequent work have shown that the memory sets in hybrid search can contain very large numbers of targets, reaching hundreds (Cunningham & Wolfe, 2014). Clearly, these cannot be stored in limited-capacity working-memory storage. Rather, targets are likely stored in a temporary, accessible portion of long-term memory (LTM) often termed *activated LTM* (aLTM, e.g., Cowan, 1988, 1999; Cunningham & Wolfe, 2014; though our use of aLTM differs somewhat from Cowan's original usage. See also, Oberauer, 2001, 2002; and Baddeley's 'episodic buffer', Baddeley, 2000). Due to aLTM's tight associations with long-term memory, memory-search through this temporary store is most probably influenced by conceptual factors (e.g., Cowan, 1988; Woltz & Was, 2006, 2007), and affected only marginally by perceptual factors. Thus, while visual guidance prioritizes features such as color, orientation or shape, memory guidance is typically dominated by the semantic content of the items, such as their category. Given these apparent differences, any attempt to *confine* search to a particular subset of items, within a specific trial, might yield effective results in one type of search, but not the other. For instance, at the visual search level, despite the requirement to attend to one category and discard another, the concurrent activation of both sets (across trials) may cause potential carry-over effects. As an example, if a member of a target set becomes an irrelevant distractor on a subsequent trial, the fact that both sets are prioritized across the entire block may elicit an erroneous response and/or slowed target latencies by the irrelevant distractor (e.g., *varied-mapping*, Schneider and Shiffrin, 1977).

At the memory search level, in contrast, separate sets of categories may enjoy relative autonomy and isolation, particularly if they cohere around a unifying semantic theme. Indeed, as mentioned above, semantic content has long been known to be a prominent factor of LTM organization (e.g., Anderson & Pichert, 1978; Bartlett, 1932; Bower, 1970; Brewer & Treyens, 1981; Collins & Loftus, 1975; Mandler, 1967; Mervis & Rosch, 1981; see also Konkle et al., 2010), and as such, search through an activated portion of LTM may be constrained by conceptual boundaries. Consequently, memory search through a particular memorized subset might be immune to influences from other memorized subsets. Thus, despite multiple set activation across the entire block or task, memory search may benefit from efficient selection, operating at least partially independently from visual search.

Interestingly, however, recent research has suggested that people can actually fail to successfully limit their memory search to specific, relevant sub-lists during hybrid search settings. Rather, they search through the *entire* memory set, regardless of the specific task requirements. For instance, Boettcher, Drew and Wolfe (2018) asked participants to memorize a set of 16 target objects, divided into two 8-item lists. Subsequently, subjects performed a visual search task in which only one of the two lists was relevant on each trial. Prior to the search task, participants were trained and tested on their memory for the items, to assure that they had explicit knowledge of each of the memory subsets (lists). Results showed that under conditions that required switching between the memory subsets on a trial-by-trial basis, the ability to 'partition' memory into two subset lists and to restrict memory search to the relevant set only was very poor. That is, during the visual search task, the mean latency for target detection in the partitioned subsets resembled latencies when searching memory through an *unpartitioned,* 16-item list. Participants in the partitioned condition, thus, appeared to be searching through the entire memory set rather than through the specific task-relevant (8-item) subset only, suggesting that they were unable, or unwilling, to flexibly shift memory subsets.

Importantly, in most of these experiments, the items in the different lists/subsets were chosen arbitrarily, such that objects within a list came from a wide range of semantic contexts and were not conceptually linked to each other. As mentioned above, perhaps subsets need to be more semantically coherent in order to streamline memory search and enable one to temporarily discard an irrelevant list. Indeed, in one of their studies, Boettcher et al. (2018) tested partitioning with subsets derived from different everyday real-world categories. Surprisingly, however, they still observed no flexible trial-by-trial memory partitioning. That is, participants continued to behave as if they were searching through the entire memory set rather than through the relevant subset/category only. One limitation of that study, however, was the use of relatively wide object categories that may have suffered from semantic overlap with each other. For example, three out of four categories (*grocery items, candy,* and *kitchen items*) were associated with food (the fourth category being *jewelry*). In the absence of clearer boundaries between the different subsets in memory, participants may have failed to ignore or inhibit an irrelevant memory subset. Alternatively, observers may have simply found the partitioning process as too difficult under these conditions (Boettcher et al., 2018).

Using a slightly different hybrid search paradigm, Nordfang & Wolfe (2018) examined whether bottom-up signals in a visual search display could affect the selection of a memory subset with well-defined semantic categories. Specifically, the authors asked whether knowledge extracted from the visual display about potentially relevant memory targets could restrict memory search to only these task-relevant items. For instance, if a memory set were composed of both coins and animals, could a *visual* display containing only round items facilitate *memory* search restricted to just the coins, while discarding the non-relevant animals? In contrast to Boettcher et al. (2018), Nordfang & Wolfe (2018) emphasized the use of categories that clearly differed both visually and conceptually (i.e., keyboard characters, simple drawings, photos of animals, and photos of scenes), thus, memory sets were highly distinct from each other. The authors found some support for flexible selective memory search under these conditions, yet only when using very small memory set sizes. That is, when holding several subsets of categories in memory, each containing a very small number of items (e.g., two objects), participants were able to restrict their memory search to a single category relevant to the visual search display (e.g., *coins*, in a display containing only *round* items). When using large memory set sizes, however, no memory-partitioning benefit in search time was observed, and performance resembled memory search through a single 16-item set, rather than through the relevant 8-item subset. Once again, then, evidence for memory partitioning behavior on a trial-by-trial basis was, at best, ambiguous. Note that in the Nordfang & Wolfe (2018) study, it could be important that items from different subsets/categories were learnt and memorized as one large mixed set, rather than as several distinct subsets memorized separately. That is, although categories clearly differed in nature, they may not have been explicitly *encoded* as separate subsets, undermining any benefit that partitioning could have had during the recognition (i.e., visual search) phase.

In spite of these negative findings, it still seems likely that some sort of partitioning of aLTM does occur in the real world. Conceptual knowledge forms part of the substrate for long-term memory structure (e.g., Anderson & Pichert, 1978; Collins & Loftus, 1975; Mervis & Rosch, 1981). The purpose of the present study was, therefore, to examine the effects of this fundamental, categorical organizing principle on hybrid-search. We addressed two questions: 1) Can clear categorical structure provide support for a more streamlined memory-selection process, and 2) can utilizing separate encoding 'sessions' for each sublist ensure that observers memorized categories as distinct subsets, rather than as some unified (mixed) memory set (e.g., Nordfang & Wolfe, 2018). We hypothesized that despite the simultaneous maintenance of both sublists in an active state, participants would be able to flexibly shift between them in memory on a trial-by-trial basis. This would imply that they could streamline memory search even in dynamic situations, speeding up overall visual search and facilitating target detection. Finding a partition-by-category effect would highlight the relations between visual search and memory search processes, potentially demonstrating their partial independence. Moreover, on a broader scale, delineating the conditions under which memory search is immune to non-relevant (albeit currently active) information may shed light on the processes contributing to the formation of bounded 'event-like' episodes in long-term memory (see *General Discussion*).

To test memory partitioning-by-category, we used photographs of real-world objects, depicting well-defined semantic categories, such as animals, vehicles, musical instruments, etc. In addition, to create very clear visual distinctiveness among categories, images of objects in different memory sets (categories) were tinted with different shades of color (e.g., all *animals* might be blue, while all *vehicles* might be orange; see Figure 2 for an illustration). As noted above, while we hypothesized that perceptual factors would exert a rather minor influence on memory-search processes, we assumed that these color masks might assist participants in delineating the different memory subsets during their encoding. That is, the colors might strengthen categorical boundaries among memory subsets and enhance their distinction during the memorization phase. The use of arbitrary color masks also allowed for assessing the independent contribution of visual-perceptual versus semantic effects on memory partitioning behavior (see below). We hypothesized that when objects from minimally-overlapping perceptual and conceptual categories were memorized separately, participants would be able to narrow memory search on a trial-by-trial basis to only the subset of items that is relevant on that trial. Put differently, visual search RTs would be faster when a 16-item memory set is partitioned into two 8-item subsets containing distinct categories and colors, compared to subsets that share a category/color. Additionally, we expected performance in the former type of subsets to mirror that of an 8-item, non-partitioned set, while performance in the latter should resemble that of a unified, 16-item non-partitioned set (regardless of target presence).

To foreshadow our results, in a series of three studies, semantic content (i.e., category), but not arbitrary perceptual content (color), was indeed shown to function as a powerful facilitator of memory partitioning, leading to more efficient search for memorized objects in visual displays.

Experiment 1: Assessing the independent contribution of semantic and visual-perceptual effects to memory partitioning

Method:

The study design, hypothesis, analysis plan and exclusion criteria were preregistered at OSF osf.io/jue2y (Wolfe et al., 2020) where the data can also be found.

Participants. Twenty-five participants (16 females, mean age 26, range 18-41) took part in the experiment for payment. Sample size was determined based on effect sizes from prior research of the mean differences between the two unpartitioned hybrid searches with 8 and 16 item visual set sizes (the '8Alone' vs. '16Alone' conditions of Boettcher et al., 2018). Power calculations indicated that a sample of 22 observers should be adequate to detect an effect size of d=0.9 with a power of 0.9 and 0.01 chance of a Type-1 error. Due to a larger number of conditions in the present experiment, however, the number of trials per condition was reduced (relative to Boettcher et al., 2018), potentially reducing the overall experimental power. We therefore decided to run a slightly larger number of subjects. Twenty-eight participants were recruited to the experiment from the Harvard Decision Science Laboratory study pool as well as the general population in Cambridge, Massachusetts. Two of these did not complete the full experiment and one had very high error rates (>2 standard deviations above participants' average performance), and thus were removed from the data analysis. All participants had normal or corrected-to-normal vision, gave informed consent, and were compensated $10/h for their participation. The Partners Healthcare Corporation Institutional Review Board approved all experimental procedures.

Stimuli & Design. We used the Hybrid Visual and Memory Search task (Wolfe, 2012), in which participants searched on each trial for one of several memorized targets among a varying number of object distractors. As mentioned earlier, we were mainly interested in the effects of semantic distinctiveness on memory partitioning, but we wished to examine whether color might additionally assist in dissociating two episodic memory sets during encoding and memorization. We therefore orthogonally manipulated semantic (i.e., categorical) and visual-perceptual (i.e., color) factors, by choosing objects from distinctive semantic categories and tinting achromatic stimuli of different sets in different colors. In the semantic conditions, objects in the two memory subsets either belonged to the same, or to different categories; in the perceptual/color conditions, the two subsets shared the same color, or differed in color. The combined factors of category and color created four Partition conditions, each containing 16 target objects divided into two memory subsets of 8 items. The four partition conditions were (see Figure 2): Same Category – Same Color, Same Category – Different Color, Different Category – Same Color, and Different Category — Different Color. In addition, there were two regular, unpartitioned conditions containing 8 (“Regular8”) or 16 (“Regular16”) items, each characterized by a unique category and color. Altogether, eight distinct object categories and eight distinct colors were used to create the six different conditions. Two categories and two colors were used for the two "Regular" conditions. Two more colors were needed for the two "same color" partition conditions and four more color were needed for the two "different color" partition conditions. Similarly, the partition conditions used six different categories. The eight object categories were: Animals, vehicles, furniture, clothes, fruit & vegetables, manipulable tools and kitchenware, musical instruments, and sweets. The eight colors used to tint the objects were red, blue, green, yellow, orange, cyan, pink and purple. For each participant, categories and colors were randomly matched to create the six conditions.

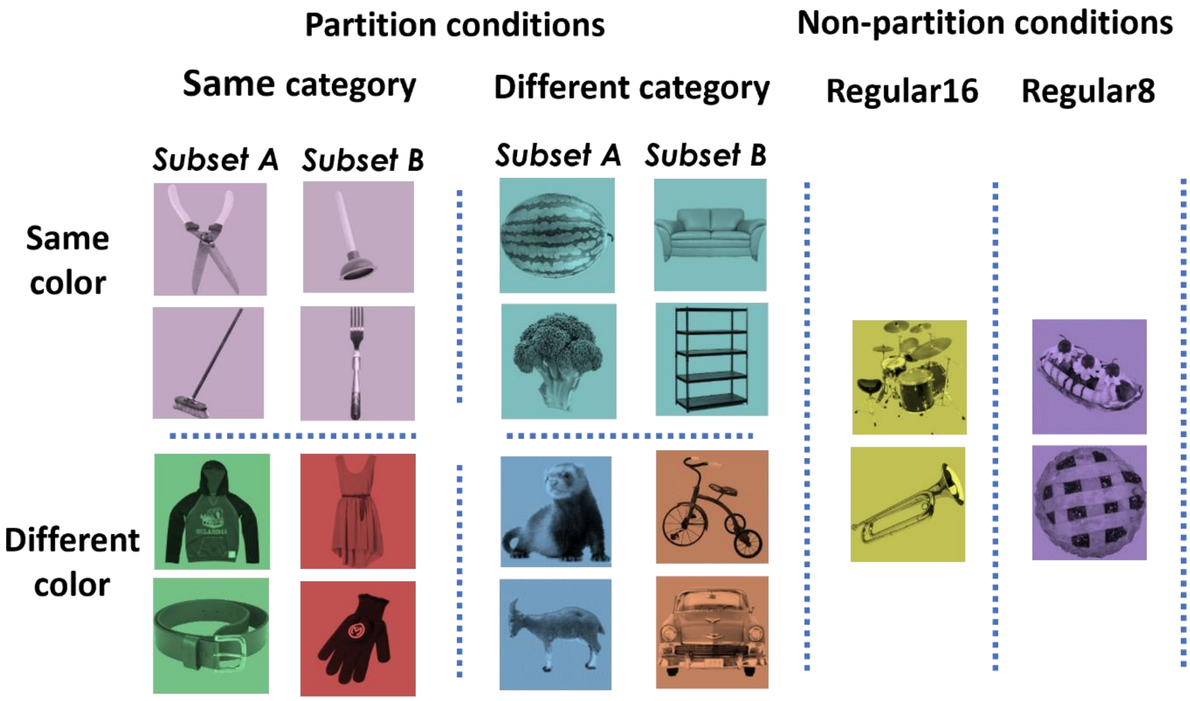

Figure 2: Examples of stimuli in the four Partition conditions and the two 'Regular' conditions of Experiment 1. Within the *Partition conditions*, participants were required to memorize 16 items, divided into two subsets of 8 items each (only 2 shown here), with Category (same/different) and Color (same/different) factors orthogonally manipulated. Thus, for instance, a subject could memorize two tool subsets tinted in pink in the 'Same Category – Same Color' condition, a clothing subset tinted in green and another clothing subset tinted in red in the 'Same Category – Different Color' condition, a fruit/vegetables subset and a furniture subset, both tinted in cyan, in the 'Different Category – Same Color' condition, and an animal and a vehicle subset tinted in blue/orange, respectively, in the 'Different Category – Different Color' condition. In addition, there were two *Non-partitioned conditions* containing 16 yellow ("Regular16") or 8 purple ("Regular8") items. The eight object categories and eight colors were randomly matched to create these six category-color combinations.

Each category contained 50 stimulus exemplar images, randomly selected from a larger pool of 100 images. Out of these 50 images, the to-be-memorized target stimuli, as well as the distractors in the visual search task were randomly extracted for each participant, in each memory set/subset. Thus, for instance, the Regular8 condition and each of the subsets within the Partition conditions contained eight sampled target stimuli, and an additional forty-two images used as distractors. In the Regular16 condition, the entire stimulus pool (containing 100 images) was randomly divided into two smaller 'subsets' of 50 stimuli (unbeknownst to the participants), each containing eight targets and 42 distractors. Although participants memorized the 16 targets as one integrated list, and saw all 84 distractors throughout the block, a specific target was always presented on a given trial among distractors belonging to its 'subset' *only* (with trials of the two hidden 'subsets' intermixed within the block). This procedure enabled, within each trial, an identical target-to-distractor presentation ratio across all conditions, for all stimulus sets and subsets.

Stimuli were drawn from 2,400 photographs of objects used in the picture memory experiments of Brady et al. (2008), and subtended approximately 3.2° by 3.2°, at an approximately 60 cm viewing distance. Importantly, colors and categories were assigned randomly to the different conditions, such that for each participant, each category/color was presented in only one condition. There were six blocks in the experiment, each containing trials taken from a single condition. The order of blocks was determined randomly.

Procedure and Apparatus. Each block consisted of a memorization phase and a visual search phase. At the beginning of each memorization phase, participants were presented with the targets to be memorized. Each target was presented individually for 3 seconds in the center of the display. Following the memorization phase, participants performed a memory-recognition test, in order to assure that items were learnt to a satisfying level: Target items in each memory set were presented in random order along with new distractor items taken from the same category and color, and participants indicated with a button press whether each object was "old" (i.e., part of the memory set, 50% of test items) or "new" (a distractor, 50% of test items). Participants had to pass two memory-recognition tests with scores above 95% correct to move on to the visual search phase for that block. If this criterion was not reached, the memory set items were presented again and the memory-recognition test was repeated (for a maximal number of five memory tests). In the four Partition conditions, participants first memorized one subset of 8 objects, successfully passed the memory-recognition test, and then memorized and were tested on the second 8-item subset (subset order was determined randomly). As in Boettcher et al., 2018, each memory subset was randomly assigned to one of two types of background displays – either *light gray* or *dark gray*. The background color appeared during item memorization and was meant to later operate as a cue indicating which subset of memory items should be searched on each trial, during the visual search phase (see below). For the two unpartitioned (Regular) memory conditions, the background color (light/dark gray) was fixed, and was randomly determined. Once participants passed two successive recognition tests, they moved on to the visual search phase. One exception to this was the Same Category – Same Color Partition condition, in which participants performed an additional memory classification test, due to the fact that the two subsets could not be easily dissociated based on mere category or color. Participants were thus required to indicate in this additional test if a specific object was from the first memory subset, the second subset, or not a target (i.e., a new item) by a button press ("1","2","3", respectively; all items appeared on a white, neutral background). The participants had to pass this test once with a 95%

correct level performance, otherwise, the test repeated. As mentioned above, this classification test was not necessary in the remaining Partition conditions, in which the two memory subsets could be easily distinguished based on category and/or color. As a further manipulation check for participants' ability to correctly dissociate the two memory subsets within the Same Category – Same Color Partition condition, this test repeated also at the end of the visual search phase.

In the visual search phase, participants searched through visual displays evenly divided between set sizes of six or 12 items. Of these displays, 50% contained a previously memorized target with the remaining distractors taken from the same category and color. Thus, in a 12-item set size display, there might be a red giraffe target and 11 other red animal distractors that had not been memorized as potential targets. The remaining 50% of trials were target-absent. Observers were instructed to click on the target as quickly as possible. If they believed no target was present, they were instructed to click on a “no target” box positioned on the left side of the screen. As mentioned earlier, for the partition conditions, the background color in each trial (light/dark gray) indicated which subset of targets was relevant for the current trial. This was only critical in the Same Category – Same Color Partition block, where it served as the cue for the relevant memory subset. In the other partitioned conditions, the background could aid partition but it was not a necessary cue, since identification of the color/category status of any item would signify the relevant subset. Background color in the unpartitioned memory conditions was identical for the search and the memorization phases.

There were 16 practice trials and 120 experimental trials in each block (2 visual set size conditions * 2 target-presence conditions * 30 items per block type). All experiments were programmed in Matlab Version 9.0.0 using the Psychtoolbox, Version 3.0.11 (Brainard, 1997). The stimuli were presented on a 24-in. screen with a refresh rate of 60 Hz, on an iMac, model A1225 (EMC 2211), with the resolution set to 1,280 X 960 pixels and an 85-Hz refresh rate.

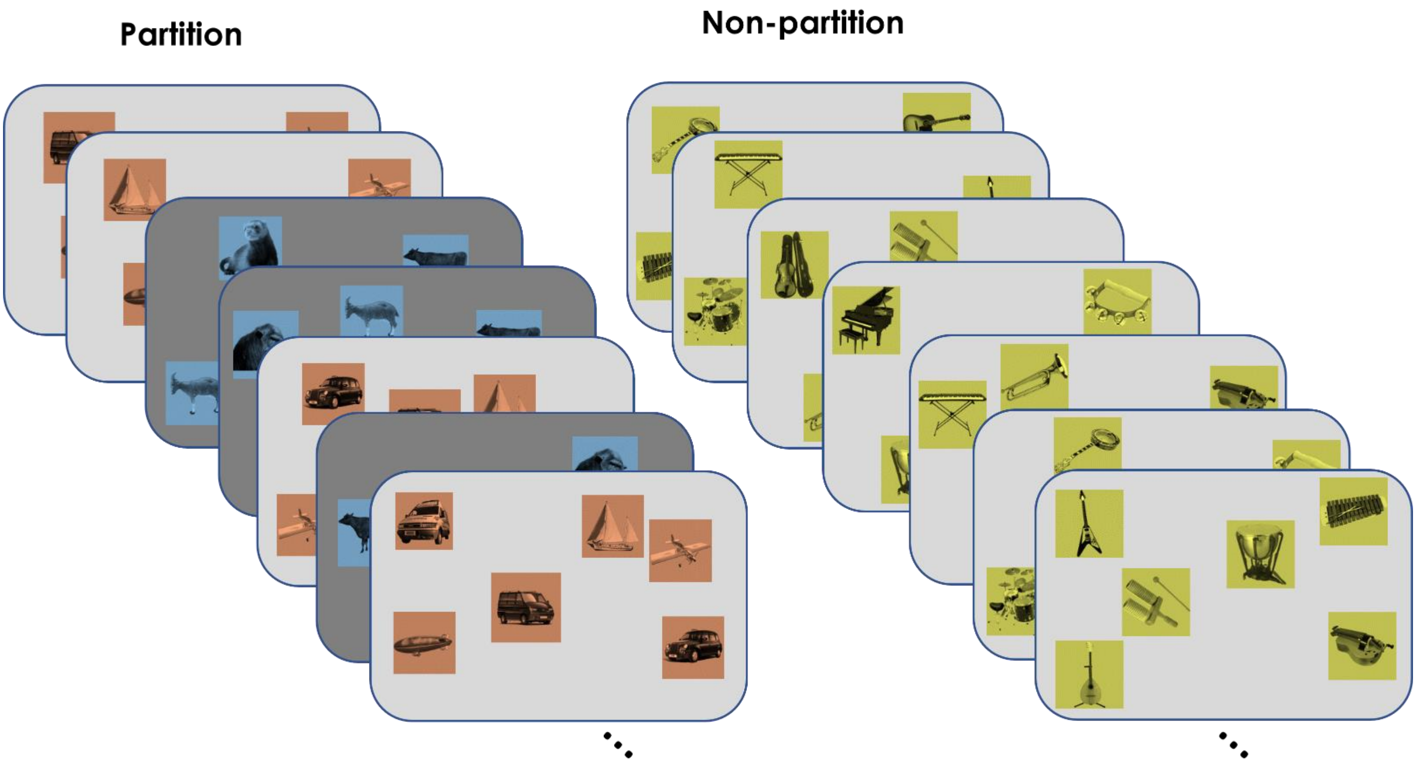

Figure 3: Trials in the visual search task. *Left*: an example of trails in the 'Different Category – Different Color' Partition block, containing an 8-item target subset of vehicles tinted in orange and another 8-item subset of animals tinted in blue. *Right*: an example of a Regular16 block, comprising a memory set of 16 musical instruments, tinted in yellow. In the partition conditions, the background color switched according to the memory-subset, while the background was constant (either light or dark grey, determined randomly) in the two non-partition ('Regular') conditions. Image examples in the figure are presented only for illustrative purposes, and are not scaled to their real world hue, size or quantity (6/12 items per trial).

Results & Discussion:

Memorization findings: As mentioned above, prior to each visual search block, participants had to pass two recognition tests with scores above 95% correct. Effectively, reaching a score of 95% on the recognition test allowed zero or, at most one erroneous response (depending on the specific condition). On average, participants performed 2.7 recognition tests for each memory set/subset, across the different conditions (0.028 errors). For the Same Category – Same Color Partition condition, participants additionally performed 1.8 memory-classification tests on average before the visual search phase (0.035 errors), and

1.5 tests after the search phase (0.031 errors). This improvement in memory test performance was expected following the extensive training of the two subsets in the visual search task. Overall, the results of the memory tests indicate that participant were proficient with the items in each memory set/subset.

Visual search findings: Figure 4 presents mean RTs for correct responses in Experiment 1 (90% of overall trials), as a function of the different memory-search and visual set size conditions. Note that "perfect" memory-partitioning results should have resembled the Regular8 results, while a complete failure of partitioning should have looked like the Regular16 results. An inspection of Figure 4 reveals that conditions involving partition by different categories (orange lines) in fact fell somewhere in between the two Regular conditions (dark/light blue lines). Nevertheless, the Different-Category Partition conditions were more efficient than the Same-Category Partition ones (yellow lines), implying that the use of distinctive categories indeed assisted memory-partitioning. The Color partitioning factor was ineffective altogether (dashed lines). In addition, subjects may have paid an overall partitioning cost, as indicated by the difference between the Partition and ‘Regular’ (unpartitioned) conditions.

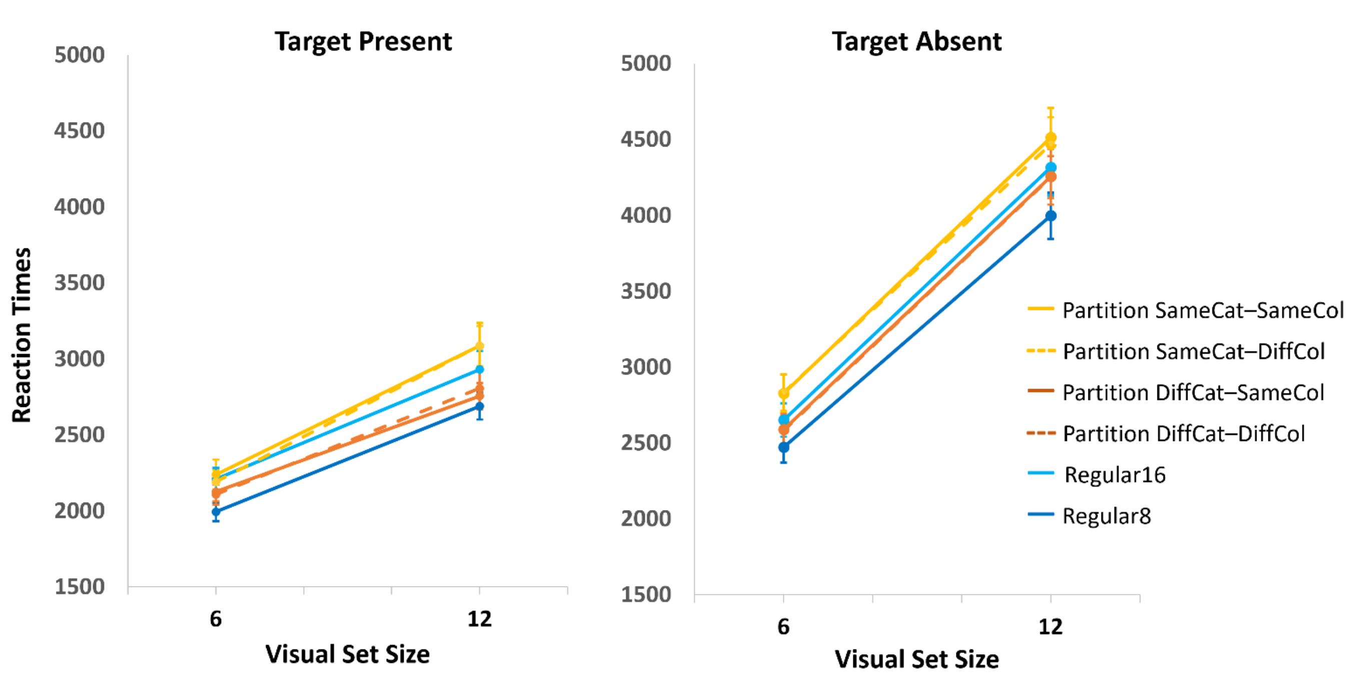

Figure 4: Mean reaction times in Experiment 1 for the different conditions, as a function of visual set size. Same/Diff Cat = Same/Different Category; Same/Diff Col = Same/Different Color. Error bars indicate ± 1 standard error of the mean.

To gain a better understanding of these findings, we conducted a two-stage analysis of the visual search reaction times (RTs). It is important to note that in the Partition conditions, we manipulated category and color independently, allowing us to assess the distinct contributions of each factor (as well as their potential interaction) to memory-partitioning. In the first stage of analysis, we examined the effects of these two factors on memory-partitioning. Subsequently, based on these results and on the finding that color was in fact ineffective - we excluded this factor from further consideration and proceeded to compare the two category-partition conditions with the two non-partitioned ('Regular') conditions (detailed below).

To determine the role of category and color in partitioning, we conducted a two-way repeated measures ANOVA on the four partition conditions (across visual set sizes). The ANOVA included Category (Same/Different) and Color (Same/Different) as factors. Within the Target-present trials, this analysis yielded a statistically significant category effect ($F(1,24)=10.79$, $p=0.003$, $\eta^2_p=0.31$, BF incl = 9.56[1]), but no color ($F(1,24)=0.04$, $p=0.84$, $\eta^2_p=0.002$, BF incl = 0.22) or color by category interaction ($F(1,24)=0.04$, $p=0.84$, $\eta^2p=0.002$, BF incl = 0.28) effects. Similar results were seen in the Target-absent trials: a statistically significant category main effect ($F(1,24)=11.6$, $p=0.002$, $\eta^2_p=0.33$, BF incl = 7.06), but no color effect ($F(1,24)=0.02$, $p=0.9$, $\eta^2_p=0$, BF incl = 0.22) or color by category interaction ($F(1,24)=0.03$, $p=0.86$, $\eta^2p=0.001$, BF incl = 0.28). The analysis of the four partition conditions, thus, clearly showed that categorical identity, but not color overlay, serves as an important factor in memory partitioning. Note that when using stimuli that appear in their natural colors, visual appearance and categorical identity tend to be positively correlated (i.e., certain categories tend to be characterized by certain colors, shapes, textures, etc.). To overcome this dependency between visual appearance and category, and to allow an independent measure of the two factors, we used achromatic images that were tinted by arbitrary hue overlays, as explained earlier. However, this linkage of category with an arbitrary color seems to have rendered the color inefficient as a memory-partitioning factor. Since it was ineffective, color was excluded from the following analysis, thus, each partition condition was averaged across the two color conditions.

Next, we assessed the direct relations between the partitioned and the unpartitioned memory conditions. As mentioned above, according to the partitioning-by-category hypothesis, the Different-Category Partition condition should ideally resemble, or match the unpartitioned, Regular8 condition; while the Same-Category partition condition (in which no partitioning behavior is expected due to ill-defined subset boundaries) should match the Regular16 condition. We could refer, thus, to partitioning by different categories as, theoretically, containing a set size of 8 items, and to the Same-Category partition condition as equivalent to an 'effective' set size of 16 items.

To test our hypothesis, a three-way repeated-measures ANOVA was conducted, including Visual set size (6/12), 'Effective' memory set size (8/16), and Memory-partitioning (Partition/Regular) as factors. Within the Target-present trials, this analysis yielded a robust visual set size main effect ($F(1,24)= 368$, $p<0.001$, $\eta^2p =0.94$, BF incl > 1000), a statistically significant memory set size effect ($F(1,24)= 13.05$, $p=0.001$, $\eta^2p =0.35$, BF incl > 1000), and a marginally significant partition effect ($F(1,24)= 4.23$, $p=0.051$, $\eta^2p =0.15$, BF incl = 1.74). None of the interactions were statistically significant (all Fs<2.4, all ps>0.13), aside from a visual set size by memory set size interaction ($F(1,24)= 4.6$, $p=0.042$, $\eta^2p =0.16$, BF incl = 0.59). Follow-up contrasts that examined more closely the memory set size effect (across visual set-sizes), showed a significant difference between the two Partition conditions ($t(24)= 3.24$, $p=0.004$, Cohen's d = 0.65, BF10 = 2.93), as well as the two Regular conditions ($t(24)= 2.54$, $p=0.018$, d = 0.512, BF10 = 11.6).

Similar results were observed with the Target-absent trials: a strong visual set size main effect ($F(1,24)= 405$, $p<0.001$, $\eta^2p =0.95$, BF incl > 1000), a statistically significant memory set size effect ($F(1,24)= 9.16$, $p=0.006$, $\eta^2p =0.28$, BF incl = 399), and a significant partition effect ($F(1,24)= 8.23$, $p=0.008$, $\eta^2p =0.25$, BF incl = 9.75). The interaction effects were non-significant, aside from the small, marginally significant interaction of visual set size by partition ($F(1,24)= 3.93$, $p=0.06$, $\eta^2p =0.14$, BF incl = 0.24) (all Fs<3.4, all ps>0.08). The follow-up contrasts concerning the memory set size effect, yielded once again a significant difference between the two Partition conditions ($t(24)= 3.34$, $p=0.003$, Cohen's d = 0.67, BF10 = 14.54). Slower RTs in the Regular16 than the Regular 8 were obtained, yet these failed to reach significance level ($t(24)= 1.73$, $p=0.097$, d= 0.35, BF10 = 0.78).

Taken together, the results above showed a medium-to-large effect of partitioning-by-category, and a small-to-medium effect of memory set-size, yet the findings also stressed an important Partitioning main effect, i.e., slower RTs for partitioned than unpartitioned memory conditions. This partitioning 'shift' in RTs implies that there may be some general cost component when dividing a memory set into several smaller subsets. What might cause this apparent cost? One probable account is the switching of memory subsets on a trial-by-trial basis, which likely consumes cognitive control capacity (see, e.g., Monsell, 2003; Monsell & Driver, 2000; Rubinstein, Meyer & Evans, 2001). That is, the *memory-selection* process may be costly, as it involves the simultaneous maintenance of two memory subsets, while only one subset is effectively relevant on each trial. In this case, a partitioning cost should be inherent to any memory-partitioning paradigm requiring flexible subset shifting. There might be an alternative (not mutually exclusive) account, however, which is associated with the specific *visual-perceptual* factors characterizing partitioned, but not unpartitioned memory conditions in the current paradigm. Recall that the background color in the partition conditions (i.e., light/dark gray) served as a cue for the relevant memory subset, alternating between trial types. In the 'Regular' (unpartitioned) conditions, in contrast, the background color remained constant across trials within the block. It is possible that the alternation of the background caused some distraction in the partition conditions, resulting in overall slower RTs than the unpartitioned conditions. Regardless of the underlying reason, shifting visual and/or memory settings is likely costly. Indeed, an analysis of the difference between the 'shift' trials (i.e., trials in which memory subsets changed) and the 'non-shift' trials (i.e., trials 2-5 after change) within the partition conditions provided some support for this hypothesis, though only within the target-absent trials: statistically significant slower responses were obtained for switch (mean= 3587 ms) than for non-switch (mean= 3468 ms) trials ($t(24)= 3.72$, $p=0.001$, $d= 0.74$, $BF10 = 32.59$, computed across partition and visual set size conditions). No such difference was seen among the target-present trials (mean switch trials = 2542, mean non-switch trials = 2527, $t(24)= 0.41$, $p=0.68$, $d= 0.08$, $BF10 = 0.23$). While a shifting account, thus, received only partial support by this analysis, it yet mirrored the larger partitioning effect seen in the target-absent than the target-present trials.

To gain a deeper insight into the partitioning cost observed in Experiment 1 and to enhance memory partitioning efficiency, several adjustments were implemented in the subsequent experiment. First, visual presentation conditions were matched across partitioned and unpartitioned ('Regular') memory conditions. That is, rather than alternating the cue for the relevant memory subset in the partition conditions only, a similar cue now alternated also in the unpartitioned conditions. This ensured consistent visual presentation conditions across trials in both condition types. Notably, while the cue indicated the relevant subset in partitioned conditions, no subsets were designated in unpartitioned conditions as there was only one, unified memory set; thus, the cue in the 'Regular' conditions was only inserted as a means for matching visual presentation conditions, and it was alternated randomly across trials (see details in Experiment 2). If the partitioning cost observed in Experiment 1 stemmed from *visual* distraction of the alternating cue, this distraction should now impact both conditions equally, effectively nullifying the cost associated solely with the partitioned conditions. Conversely, if the partitioning cost is inherent to memory-selection processes and arises from shifting *memory* subsets across trials, it should remain unaffected by controlling for cue visual presentation across experimental conditions.

Second, to enhance memory-partitioning performance, we altered the nature of the visual cue representing the relevant memory subset. Instead of background color (light/dark gray), we employed *spatial location* (left/right screen side) to denote which subset targets were searched on each trial. Specifically, in partition conditions, participants learned each subset within one of two screen sides and subsequently searched for the relevant targets within the same side. For example, if a search trial included items from a subset initially memorized on the left, the target and distractors would all appear on the left during visual search. This way, each memory subset was linked to a specific spatial location, both during memorization and visual search phases. Since spatial location is typically considered an important visual feature that might be uniquely bound to object identity (e.g., Golomb et al., 2014; Hollingworth, 2006; Kahneman et al., 1992; Pertzov and Husain, 2013; Treisman & Gelade, 1980), we hypothesized it might offer a stronger association with items in the different memory subsets than a background color. Furthermore, location could represent an important aspect of real-world partitioning opportunities not captured by color. For instance, one memory subset might be useful in one location (e.g., basement) while another is relevant elsewhere (e.g., bedroom). Consequently, using a location cue for the partition conditions could improve subset segregation during target encoding/memorization as well as during the visual search phase (see details below, for the equivalent manipulation within the non-partitioned conditions).

Finally, we removed the superficial tint layer of stimuli due to its ineffectiveness in this experiment. There were two partition conditions (Same-category, Different-category), in which images were presented in their natural colors. We anticipated observing a partitioning-by-category effect, i.e., faster search times for the Different-category than the Same-category Partition condition, irrespective of target presence/absence in the visual display.

Experiment 2: Using spatial location as a partitioning cue and controlling for visual factors across conditions

Method

As in the previous study, all materials for this study were preregistered at OSF osf.io/3vja5 (Wolfe et al., 2023).

Participants. The actual power obtained for the difference between the two partition conditions (across color and visual set size) in Experiment 1 was 0.87 and 0.89 in the target-present and target-absent trials, respectively. We thus decided to run a similar number of participants in Experiment 2. Subject recruitment continued until reaching a number of participants similar to our previous experiment. Thirty-two participants were recruited. One did not complete the experiment due to technical problems, four participants had excessive erroneous responses in the visual search task (>2 standard deviations above participants' average performance), and one participant had very high error rates (>2 standard deviations above group average) in the memory-classification test. These participants were discarded from the data analysis. Thus, 26 observers (15 females, mean age 26, range 18 - 36) were included in the experiment, taking place in the Visual Cognition lab of the Open University of Israel. All participants had normal or corrected-to-normal vision, gave informed consent, and were compensated a sum equivalent to $15/h for their participation. The Ethical committee of the Open University of Israel approved all experimental procedures.

Stimuli & Design. Without the color (hue overlay) factor there were only four conditions – two Partition (Same/Different-Category subsets) and two Regular (set sizes 8 and 16) conditions. Five image categories were allotted randomly to the four conditions (with two categories needed in the Different-Category Partition condition) – animals, vehicles, furniture, clothes & produce. As mentioned above, all images appeared in their natural colors on a light gray background. Due to the decrease in the number of conditions, the number of trials per condition was increased to 200. There were four blocks in the experiment, one for each condition. The order of blocks was determined randomly. All other aspects of stimuli and design were similar to Experiment 1.

Procedure and Apparatus. The experimental procedure was also similar to the previous experiment. As mentioned earlier, one of the critical changes was that the cue indicating the relevant memory subset was changed from a background color to screen location. That is, in the partition conditions, rather than appearing on different background displays, the two memory subsets appeared on different sides of the display. The assignment of subsets to the left or the right side of the screen was randomly determined for each condition and subject. During the memorization phase, items from the first subset were presented on the left half of the screen. They were memorized and tested while appearing on that side at its center. After successfully completing the memory-recognition test, participants memorized a second memory subset, with stimuli appearing on the right side of the screen. When participants moved on to the visual search task, target and distractors' location matched the location of that subset in the memorization phase. Thus, for instance, if the trial included an animal subset that was originally memorized on the left, all animals appeared on the left during the visual search task, and vice versa for the transportation subset (see Figure 5, left). Note that although the spatial cue indicated which subset of targets was relevant for the current trial, it was only critical in the Same-Category partition block. In the Different-category block, subset relevance could also be derived from the categorical identity of the currently displayed items.

For the unpartitioned memory conditions (Regular8, Regular16), a single memory set was memorized and tested using images that were presented at the screen center (Figure 5, right). However, to equate location-switches across trials within the Regular and Partition conditions during the visual search phase, Regular8 and Regular16 search stimuli were presented in a random fashion, either to the left or the right of the screen. Spatial location was thus meaningless in these conditions. As mentioned above, to the extent that visual presentation differences may have contributed to the partitioning cost seen in Experiment 1, such a cost should be reduced (or perhaps eliminated) in the present experiment.

Stimuli were presented on IBM computers running Windows 10, with a 15-in. CRT monitor (resolution 1280 x 960 pixels) and an 85-Hz refresh rate, at the Visual Cognition lab of the Open University of Israel. All other aspects of the experiment were identical to Experiment 1.

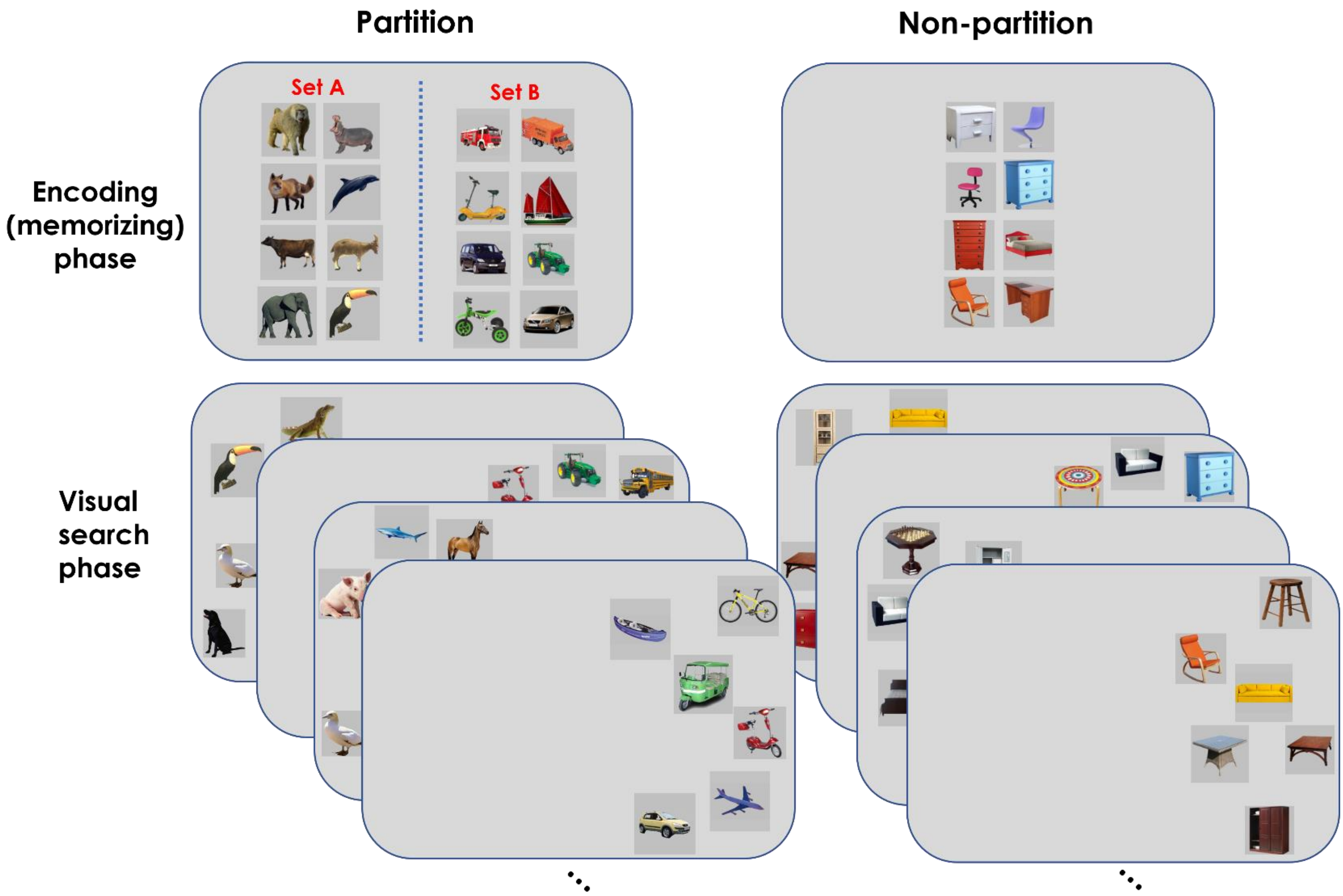

Figure 5: The encoding (memorization) and visual search displays in Experiment 2. *Left*: an example of a Different-Category Partition block. During memorization, items from one subset appeared on the left half of the screen and were tested there. After completing the test, participants memorized a second subset on the right side (subsets are presented simultaneously for illustrative purposes only). In the visual search task, targets and distractors matched the subset's location from the memorization phase. Trials from both subsets were randomly intermixed. *Right*: an example of a Regular8 block. To ensure equal presentation conditions in the Regular and Partition conditions during the visual search phase, Regular8 and Regular16 search stimuli were randomly presented on the left or right side of the screen in each trial. Spatial location was irrelevant and meaningless in these conditions. Note that image examples are presented only for illustrative purposes, and are not scaled to their real world size or quantity (6/12 items per trial).

Results & Discussion:

Memorization findings: On average, participants performed 2.9 recognition tests for each memory set/subset, across the different conditions (0.038 errors). For the Same-Category Partition condition, participants additionally performed 1.5 memory-classification tests on average before the visual search phase (0.07 errors), and 1.2 tests after the search phase (0.06 errors).

Visual search findings: Figure 6 presents mean RTs for correct responses in Experiment 2 (89% of overall trials), as a function of the different memory-search and visual set size conditions. Clearly, to a first approximation, partition-by-category was effective, as the difference between the Same- and Different-Category Partition conditions remained, replicating the findings from Experiment 1. Furthermore, the partitioning cost now seemed to disappear, as partition with different categories produced results similar to those produced by a single memory set of 8 items, while two subsets from the same category produced results similar to a memory set size of 16.

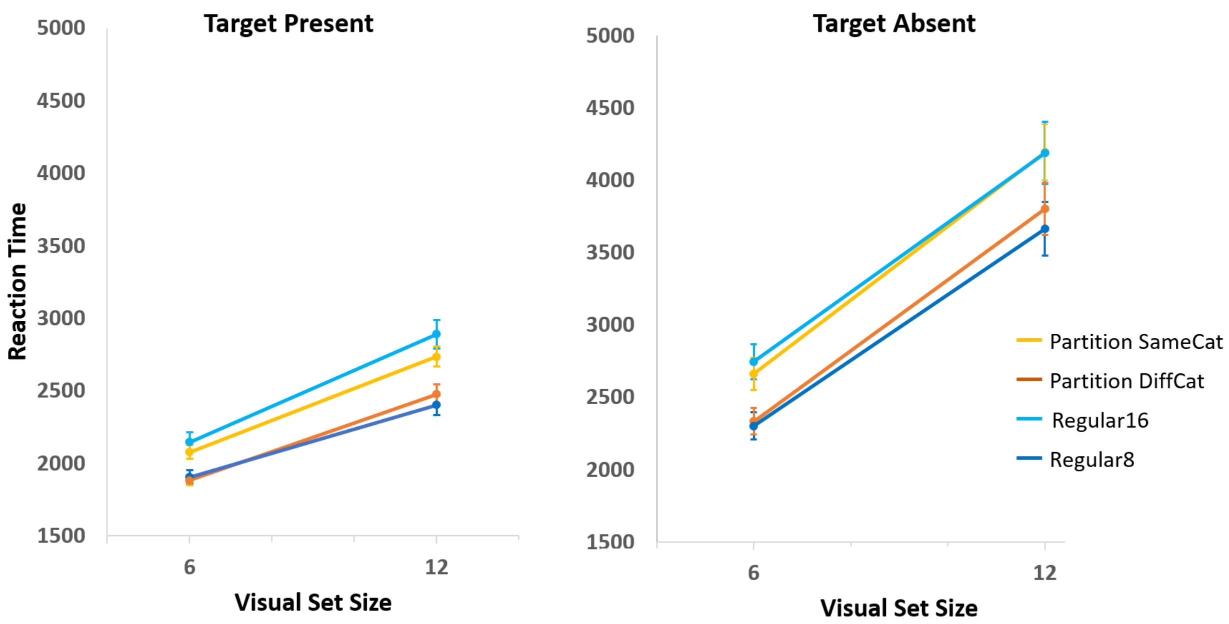

Figure 6: Mean reaction times in Experiment 2 for the different conditions, as a function of visual set size. SameCat = Same Category, DiffCat = Different Category; Error bars indicate ± 1 standard error of the mean.

To test the significance of these effects, a three-way repeated-measures ANOVA was once again conducted, including Visual set size (6/12), 'Effective' memory set size (8/16), and Memory-partition (Partition/Regular) as factors. Within the Target-present trials, this analysis yielded robust visual set size ($F(1,25)= 340$, $p<0.001$, $\eta^2p =0.93$, BF incl > 1000) and memory set size main effects ($F(1,25)= 36.6$, $p<0.001$, $\eta^2p =0.59$, BF incl > 1000), but no partition main effect ($F(1,25)= 0.67$, $p=0.42$, $\eta^2p =0.03$, BF incl = 0.27). There was an interaction of visual set size by memory set size ($F(1,25)= 9.84$, $p=0.004$, $\eta^2p =0.28$, BF incl = 1.23) and a triple interaction ($F(1,25)= 4.53$, $p=0.043$, $\eta^2p =0.15$, BF incl = 0.43), while the rest of the interactions were non-significant (all Fs<2, p>0.1). A closer look at the specific comparisons of interest unsurprisingly yielded significant effects of partitioning by category (Same-Category vs. Different-Category Partition: $t(25) = 3.42$, $p = 0.002$, $d = 0.67$, $BF_{10} = 17.45$) and of memory set size among the two unpartitioned memory conditions (Regular16 vs. Regular8: $t(25) = 5.15$, $p < 0.001$, $d = 1.01$, $BF_{10} = 937$).

A similar pattern of results was obtained with the Target-absent trials: a strong visual set size ($F(1,25) = 280$, $p < 0.001$, $\eta^2p = 0.92$, BF incl > 1000) and memory set size main effect ($F(1,25) = 16.26$, $p < 0.001$, $\eta^2p = 0.39$, BF incl > 1000), but no partitioning main effect ($F(1,25) = 0.71$, $p = 0.79$, $\eta^2p = 0.003$, BF incl = 0.18). None of the interactions were statistically significant (all $Fs < 2.25$, all $ps > 0.1$). As for the follow-up comparisons of interest, a significant effect was obtained when directly contrasting the two partition conditions ($t(25) = 2.25$, $p = 0.034$, $d = 0.44$, $BF_{10} = 1.74$), as well as when comparing the two unpartitioned memory conditions ($t(25) = 3.59$, $p = 0.001$, $d = 0.71$, $BF_{10} = 25.7$).

Overall, the results of Experiment 2 replicated the partitioning-by-category effect seen in Experiment 1. Namely, hybrid search was significantly more efficient when a memorized set of items was partitioned into two smaller subsets containing different categories than when it contained same-category items. However, the insertion of a spatial cue offered no advantage relative to the color background cue. Directly comparing the partitioning-by-category effect between the two experiments yielded null results ($F(1,49) = 0.21$, $p = 0.65$, $\eta^2p = 0.0008$, BF incl = 0.29). Furthermore, the failure to partition memory in the Same-Category condition in Experiment 2 suggests that spatial location *alone* is insufficient for effective partitioning. Nonetheless, no partitioning cost was observed in the present study, suggesting that flexible shifting of memory sets (on a trial-by-trial basis) may actually be costless, at least when using highly distinct object categories. Presumably, the 'shift' in RTs seen in the previous study resulted from unbalanced visual presentation settings across the Partition and the Regular memory conditions, rather than from a switch cost inherent to the memory-selection process.

One important question is nevertheless yet to be resolved. Note that in Experiments 1 and 2, all the distractors came from the category of the targets for the relevant memory set or subset. If the targets were animals, so were all the distractors. As a result, within the Different-Category Partition condition, each distractor was also a cue to the relevant subset identity ('Is this *animal* a target or not?'). In the Same-Category Partition condition, in contrast, the distractors in both subsets overlapped in category, thus the only indication for the relevant subset on a given trial was an arbitrary cue (i.e., background color or screen side). This asymmetry among conditions likely rendered subset recognition more difficult in the Same- than the Different-Category Partition condition. In addition, subjects may have adopted a general target-familiarity strategy in the Same-Category condition (e.g., 'detect *any* animal target, regardless of sublist'), since there was no true incentive to constrain one's search to the specific subset-relevant targets.

To deal with these potential limitations and to control for the asymmetry between the two Partition conditions, in Experiment 3 we introduced several important changes. First, distractors were selected from non-target categories, forming a heterogeneous, rather than an homogeneous categorical visual display. Returning to the real world analogy, the camping gear (tent, sleeping bag, backpack) might all be in the basement but that basement would also include irrelevant distractors from other categories (gardening equipment, tools, etc.). To the extent that partitioning-by-category is dependent on the existence of same-category distractors in the display, any partitioning effect within the Different-category condition should be dramatically weakened (regardless of target presence). Second, partition was challenged by inserting two types of *lure* items to the visual search task. The first type, Same-Category Lures, consisted of items from the target category that did not belong to the memorized subset. For example, in an animal-target trial, a lure item could be an animal not included in the memorized list. Given a categorically heterogeneous display, these lures assured that participants were not simply performing a categorization task ('is there *any animal* in the display'), rather, they were searching for the specific list of memorized targets (a fox, a cow, a dolphin, etc.). The second type, Other-Set Lures, comprised targets from the 'wrong' subset. If animals are targets when presented on the left, a target animal appearing amongst right-sided items is a lure that should be avoided. Importantly, within the context of an heterogeneous display containing both types of lure items, there is no obvious indication of the relevant subset targets on either partition condition, aside from the pre-learnt cue (e.g., screen side). Thus, the asymmetry between the two partition conditions was largely minimized. In addition, the Other-Set lures specifically required that participants be aware of the to-be-searched subset on each trial. A high proportion of false-positive responses to these lures would suggest that observers may not be fully aware of the relevant subset on a given trial, or they might be disregarding their knowledge of the task-relevant subset (potentially responding based solely on general target familiarity). However, if accuracy rates remain high despite some distraction caused by the lures (e.g., slower responses to target items), it would indicate that participants are indeed aware of subset relevance. Yet, they may be unable to perfectly confine their visual search to the relevant targets alone. This implies that both memory subsets/categories are somewhat active during task performance. Despite this concurrent activation of multiple sets or categories at the *attentional* (visual search) level, selective partitioning by category is expected at the *memory* search level. Therefore, observing an interference effect with the lures, alongside a constrained memory search within different categorical subsets, would suggest that visual search and memory search operate relatively independently from each other.

Experiment 3: Using heterogeneous distractors and inserting lure items

Method:

As in previous studies, all materials for this study were preregistered at OSF osf.io/7mwkb (Gronau et al., 2024).

Participants. In Experiment 2, the actual power obtained for the difference between the two partition conditions was 0.91 in the target-present trials, but only 0.58 in the target-absent trials. Based on a conservative estimated effect size of dz=0.65 and an estimated power of 0.9 (given an alpha of .05), we estimated a required sample size as 28 participants for Experiment 3. Thirty participants (21 females, mean age 27, range 20 - 33) were recruited for payment in the Visual Cognition lab of the Open University of Israel. Two participants had excessive erroneous responses in the visual search task (>2 standard deviations above participants' average performance), and thus were removed from the data analysis. Subject recruitment continued until reaching twenty-eight qualifying participants.

Stimuli, Design & Procedure. As in Experiment 2, spatial location was used to cue which subset was relevant in the Partition conditions. Five categories of objects (animals, vehicles, furniture, clothes & produce) were randomly assigned to the four memory (two Partition and two Regular) conditions. Sampling of memory-target and memory-distractor stimuli was identical to previous studies. Importantly, and in contrast to previous experiments, object distractors in the visual search task were taken from a pool of 400 item images. This pool included a wide variety of object categories that did not overlap with any of the five memory-target categories (e.g., tools, appliances, stationary, toys, sport equipment, etc., see examples in Figure 7).

Furthermore, as mentioned above, we introduced two types of lure items to the visual search task. The *Same-Category Lures*, consisting of items that belonged to the target category but were not part of the memorized subset, were present in all conditions and were meant to prevent a simple categorization strategy (e.g., 'search for *an animal*'). If some basic feature(s) could guide attention to the target category, observers might find a same-category lure quite quickly (Cunningham & Wolfe, 2014; Levin et al., 2001; Nako et al., 2014; Yang & Zelinsky, 2009). However, upon detection of the target category, they would need to search the relevant part of the memory set in order to respond only to a specifically memorized target. The *Other-Set Lures*, in contrast, were inserted only during the two partition conditions and consisted of targets from the other (irrelevant) memory subset (e.g., a vehicle image, in an animal-target subset trial). As specified above, these lures were inserted to ensure that participants were explicitly aware of the relevant memory subset on each trial, and were not responding on the basis of mere target familiarity (‘is there *any target* in the display’). Lure prevalence was 50% and target- and lure-presence were orthogonal to each other (i.e., 50% of the trials on each of the target-present/absent conditions included a lure item). On 'Regular', control trials, all lures were non-target members of the target category (since there could be no lures from another memory subset). On Partition trials, the two types of lure items were divided randomly. Thus, each type of lure appeared on approximately 25% of target-present/absent trials.

Due to the use of two different types of lure items in the partition conditions, a larger number of trials was required in these conditions, relative to the Regular, unpartitioned conditions. Thus, the Partition conditions contained 288 trials per block, whereas the Regular conditions contained 144 trials per block, yielding a total of 864 trials in the experiment. All other aspects of the design and the procedure were identical to the Experiment 2.

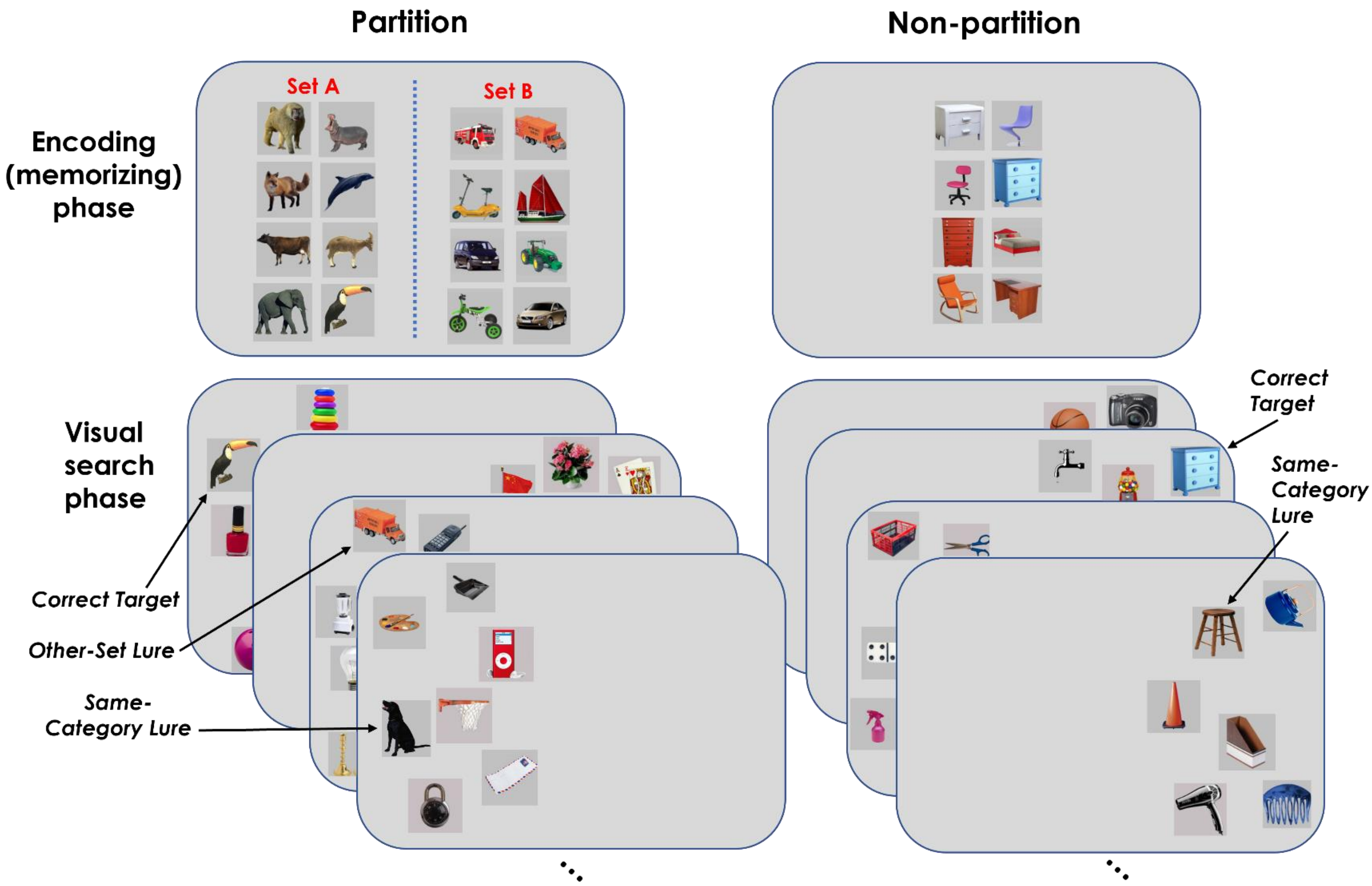

Figure 7: The encoding (memorization) and visual search displays in Experiment 3. *Left*: an example of a Different-Category Partition block. During the visual search task, stimuli's side indicated the relevant to-be-searched subset. In the present example, a toucan is a memorized animal target, appearing in the correct (left) side of the screen. An orange truck is an Other-Set lure since it is a vehicle target that appears in the incorrect (left) side. A dog is a Same-Category lure, as it belongs to the relevant category, but is not part of the memorized list. *Right*: an example of a Regular8 block. As in Experiment 2, to equate visual presentation settings across conditions, stimuli were randomly presented on the left or the right side of the screen. Only Same-Category lures were inserted in the Regular conditions (as there were no memory subsets). Image examples are presented only for illustrative purposes, and are not scaled to their real world size or quantity (6/12 items per trial).

Results & Discussion:

Memorization findings: On average, participants performed 2.3 recognition tests for each memory set/subset, across the different conditions (0.022 errors). For the Same-Category Partition condition, participants additionally performed 1.5 memory-classification tests on average before the visual search phase (0.064 errors), and one test after the search phase (0.016 errors).

Lure findings: Prior to examination of the main partition results, we wish to ensure that participants obeyed task-instructions by responding to the specific memorized targets, and discarding (as much as possible) lure items. Figure 8 presents the effects of the lure items on performance in Experiment 3. An inspection of the figure indicates that observers paid an overall cost in RT and accuracy for both types of lure items. Despite this cost, accuracy performance was remarkably high, suggesting that participants largely managed to refrain from falsely responding to these foil objects (see analysis below).

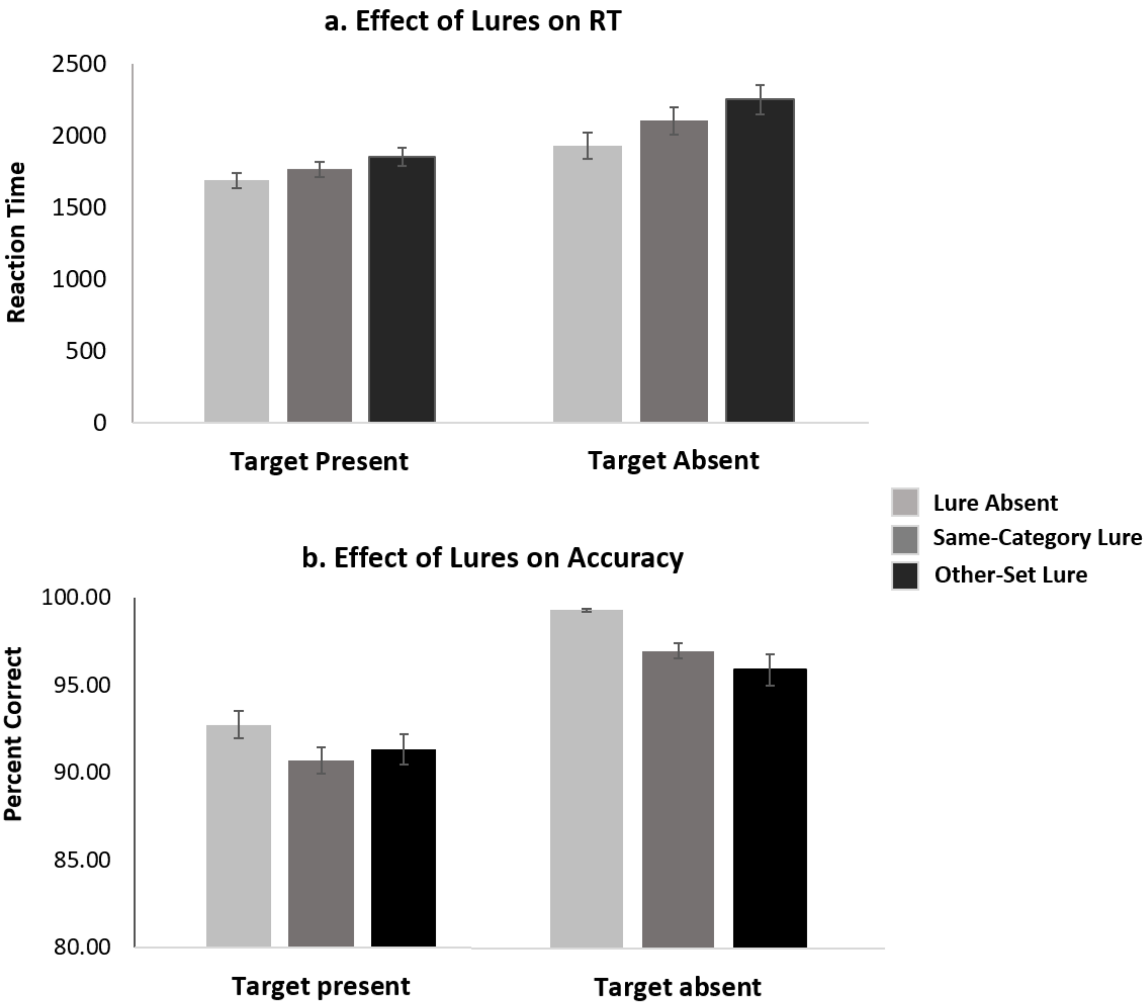

Figure 8. Effects of lure items on performance in Experiment 3. *Upper panel*: mean RT in the Lure-absent and the two types of Lure-present conditions (computed across all memory and visual set-size conditions, separately for target-absent and target-present trials). *Lower panel*: Proportion of correct responses within these conditions. Error bars indicate ± 1 standard error of the mean.

To assess the effect of the lure items on search RTs, we performed two independent contrasts, the first comparing Lure-absent and Lure-present trials (across both types of lures), and the second directly comparing responses in the two types of lure-item trials. A within-subject contrast examining the overall effect of lures on search RTs (i.e., Lure-absent vs. Lure-present trials) yielded a statistically significant result within both Target-present ($t(27)=11.99$, $p<0.001$, $d=2.27$, $BF10>1000$) and Target-absent ($t(27)=12.08$, $p<0.001$, $d=2.28$, $BF10>1000$) trials. When directly comparing RTs for the two types of lure trials, the RT distraction was larger for Other-set lures than for Same-category lures (Target-present trials: $t(27)=3.87$, $p<0.001$, $d=0.73$, $BF10=51$; Target-absent trials: $t(27)=3.93$, $p<0.001$, $d=0.74$, $BF10=60$). These results indicate that the lures' presence was overall distracting. That is, detecting a *non-target* animal (when the relevant memory subset contained animals), or a *target vehicle* (belonging to the other memory subset, within partition-trials) significantly slowed participants' responses. Both types of distractions were expected. Any target category lures that were detected would require a search of the relevant memory set (see, e.g., Drew et al., 2017; Hout et al., 2017; Nako et al., 2014; Maxfield & Zelinsky, 2012; Wolfe, 2021; Yang & Zelinsky, 2009; Yu et al., 2023). More speculatively, the cost of the other-set lures may have arisen from a momentary doubt that the correct subset was prioritized for the current trial. Notably, this distraction was observed regardless of the partition condition, indicating that the two subsets were maintained in a highly active state, whether containing items from same or from different categories (Other-set lure vs. Lure-absent trials, in Same-category partition: $t(27)= 10.63$, $p<0.001$, $d= 2.01$, $BF10 >1000$; in Different-category partition: $t(27)= 5.78$, $p<0.001$, $d= 1.09$, $BF10 >1000$). A similar overall distraction effect was observed when comparing the arcsine-transformed accuracy rates of lure-absent vs. lure-present trials (Target-present: $t(27)=3.53$, $p=0.002$, $d=0.67$, $BF10=24$; Target-absent: $t(27)=8.57$, $p<0.001$, $d=1.62$, $BF10>1000$). There was no difference, however, between the distraction levels of the two lure types (Target-present: $t(27)=0.78$, $p=0.44$, $d=0.15$, $BF10=0.27$; Target-absent: $t(27)= 0.65$, $p=0.52$, $d=0.12$, $BF10=0.24$).

And yet, in spite of these distracting effects, participants were generally successful in withholding their responses to the lure items (at the cost of longer RTs), as evident from their overall high accuracy performance in lure-present trials (94% on average, compared to 96% in lure-absent trials, when averaged across target-presence conditions). In fact, the minimal level of correct responses was 89% for trials containing Other-Set lures (within the Regular16 condition) and 90% for trials containing Same-Category lures (within the Same-Category partition condition). These high accuracy levels clearly indicate that participants were aware of the relevant memory subset on each trial, and were not responding on the basis of simple categorization or target familiarity. Thus, we can now proceed to examine our main question of interest regarding the ability to flexibly partition memory into two categorically-defined memory subsets in the context of a heterogeneous visual search display.

Visual search findings: Figure 9 presents mean RTs for correct responses in Experiment 3, as a function of the different memory-search and visual set size conditions. Note that the Other-Set lure items were present only in the two partition conditions, as there was no division into different subsets in the Regular conditions. To allow a fair comparison of all *four* conditions (i.e., the two partition and two 'Regular' conditions), we excluded at a first stage trials containing the Other-Set lures (approximately 25% of partition trials) from the main analysis. This analysis is presented in Figure 9a. We then analyzed the 'Other-Set' lure trials (appearing only in the two partition conditions) separately, see Figure 9b.

An examination of Figure 9a reveals that RTs in Experiment 3 were markedly faster than those of Experiment 2. A between-subjects t-test conducted on the overall RTs of the two experiments yielded a significant difference for both target-present ($t(52)= 8.19$, $p<0.0001$, $d= 2.23$, $BF10 >10000$) and target-absent ($t(52)= 8.08$, $p<0.0001$, $d= 2.2$, $BF10 >10000$) trials. In addition, the partition results of Experiment 2 appear to be replicated, at least for the target-present trials, in Experiment 3 (see detailed analysis below).

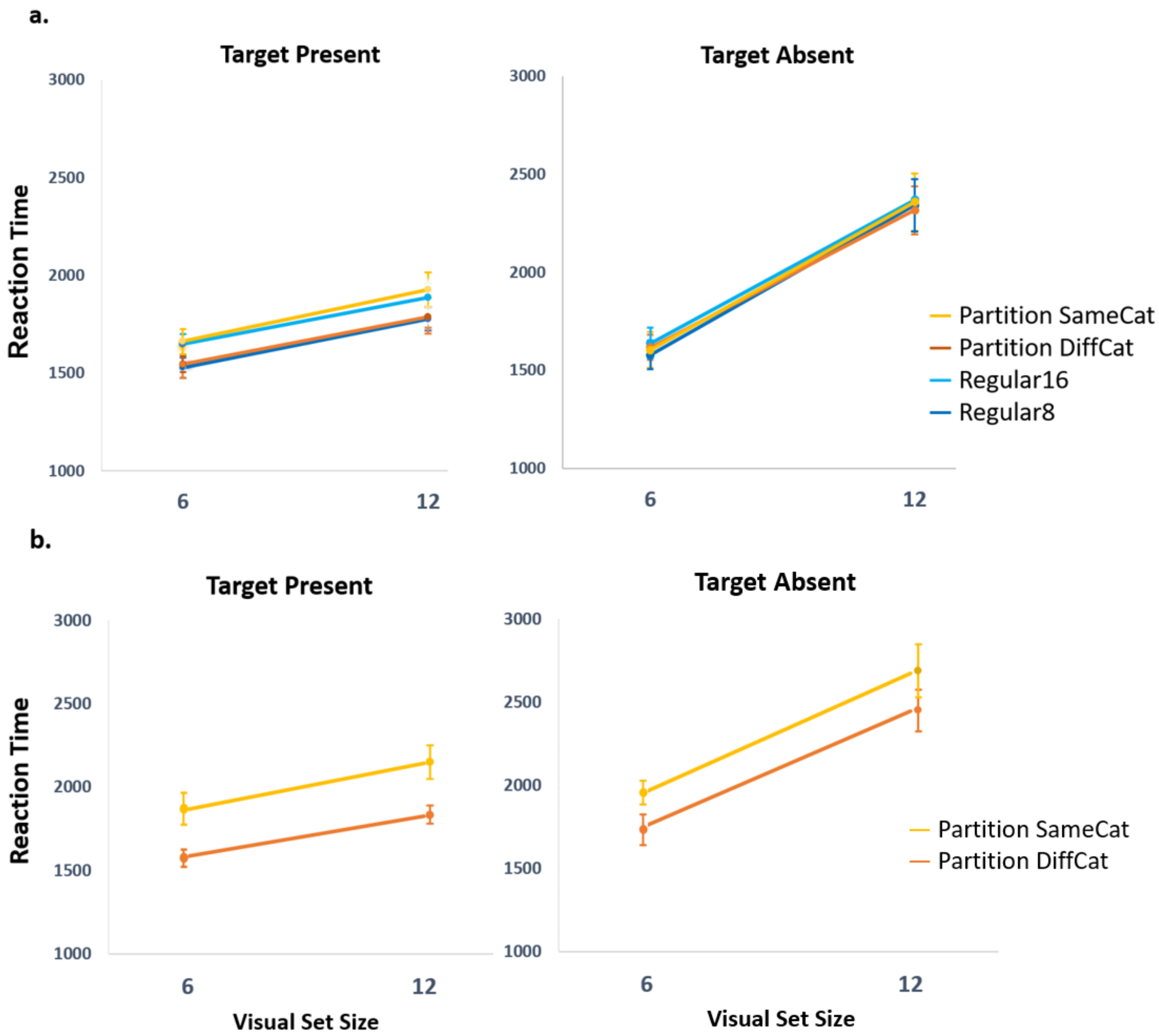

Figure 9. Mean reaction times in Experiment 3 for the different conditions, as a function of visual set size. *Upper panel* (a): In the main analysis there were Lure-Absent and Same-Category Lure trials, which were present in all four conditions. The Other-Set Lures were present only in the two Partition conditions, thus, trials containing these lures (approximately 25% of partition trials) were excluded from this analysis. *Lower panel* (b): RTs in the remaining Other-Set Lure trials, within the two partition conditions. SameCat = Same Category, DiffCat = Different Category; Error bars indicate ± 1 standard error of the mean.

First, it is worth thinking about why this version of the experiment is faster. In Experiment 2, if the observer attended to an item, they would find a member of the target category (e.g., an animal). They would then need to search memory to determine if it was a target. Each deployment of attention in the visual display would provide an opportunity for a partitioned or unpartitioned memory-search, such that RTs in each trial reflected the overall memory-search time summed across distractors (and, potentially the target). In Experiment 3, in contrast, most selected items would not provoke a memory search because distractor items would be identified as members of an irrelevant category (‘This *sofa* could never be a target animal’). The lures might generate a memory search, and the target would do so, but the quick rejection of most items would produce short RTs. On target-absent trials, only the occasional lure would require memory search, making it even harder to see evidence for memory-partitioning.

And yet, despite the rapid rejection of distractors in Experiment 3, the target-present trials still showed evidence for memory partitioning (Figure 9a, left). A three-way repeated-measures ANOVA conducted on the correct RTs of the Target-present trials, including Visual set size (6/12), ‘Effective’ memory set size (8/16), and Memory-partition (Partition/Regular) factors, yielded a robust visual set size ($F(1,27)= 190$, $p<0.001$, $\eta^2p =0.32$, BF incl > 1000) and a memory set size main effect ($F(1,27)= 11.5$, $p=0.002$, $\eta^2p =0.08$, BF incl > 1000), with no partition main effect ($F(1,27)= 0.7$, $p=0.41$, $\eta^2p =0.002$, BF incl = 0.20). There were no interaction effects (all Fs<0.5, all ps>0.5). Follow-up planned comparisons (across visual set-size conditions) yielded a significant effect of partitioning by category (Same-Category vs. Different-Category partition: $t(27)= 2.35$, $p=0.03$, $d= 0.44$, BF10 = 2.05) and of memory set size in the two unpartitioned memory conditions (Regular16 vs. Regular8: $t(27)= 2.19$, $p=0.04$, $d= 0.41$, BF10 = 1.54).

The analysis of the Target-Absent trials showed a strong visual set size effect ($F(1,27)= 159$, $p<0.001$, $\eta^2p =0.86$, BF incl > 1000) but no other main effects or interactions (All Fs<1.35, p>0.25) (Figure 9a, right). As previously discussed, the absence of a memory set-size effect in both the Regular and Partition conditions can likely be attributed to the rapid rejection of irrelevant distractors, due to their categorization as members of an irrelevant category. In contrast, trials containing a same-category lure (e.g., a non-target animal) likely provoked a memory search, but such a process, induced by an individual lure, was presumably not sufficiently powerful to produce a memory set-size effect. It's worth noting that the target-absent trials in Experiments 1 and 2 exhibited a pronounced memory set-size effect, which emerged from the pooled memory search processes across all same-category distractors in the display.

With regards to the Other-Set lure items, these were present only in the two partition conditions, as mentioned earlier. We therefore analyzed trials containing these lures separately (see Figure 9b, lower panel). An inspection of the figure reveals that a categorical-partition effect was obtained in both Target-present ($t(27)=4.57$, $p=0.0001$, $d=0.86$, BF10=271) and Target-absent ($t(27)=2.78$, $p=0.01$, $d=0.53$, BF10=4.66) trials. The former findings corroborate on our main analysis, indicating that partitioning-by-category is possible when targets appear within a heterogeneous visual search display. The latter findings with the target-absent trials was smaller than the one observed in the target-present trials, and was necessarily caused by the Other-Set lure items. This partitioning effect could be explained by two possible accounts. One explanation is that the memory-partition effect with the lures is akin to the one observed with the 'correct' targets (in the Target-present trials). In this scenario, once an item from the irrelevant subset category was detected (e.g., a vehicle in an animal subset trial), it evoked a memory-search process that resulted in target (lure) identification. It was only at a subsequent stage that task requirements were considered, typically leading to classification of the target as a ‘lure’ item and withdrawing one’s response. The difference in reaction times between the Same- and Different-category partition conditions, therefore, reflects variations in their effective memory-set size, since both elicited memory search processes. Alternatively, a memory search process might have been triggered solely by the Same-category condition, as the initial categorization of the lure item in the Different-category condition rendered it irrelevant. In other words, the lure in the latter case (belonging to the other-set category) could be rapidly rejected at an early categorization stage, preempting the memory-search stage. While both suggestions seem sensible, we cannot clearly determine at this point which of the two better account for the observed findings. Further research is required to resolve this point.

General Discussion

The present study investigated whether a flexible, efficient partitioning of memory into two categorically-defined subsets can take place on a trial-by-trial basis. In three experiments using a hybrid search paradigm, flexible partitioning-by-category was evident when using semantically restricted, clearly defined object categories (such as animals, vehicles, clothes, etc.). Experiment 1 revealed that semantic characteristics, but not perceptual features (such as an arbitrary color), provide a firm basis for efficient memory partitioning. Experiment 2 further showed that such memory partitioning can be costless, as long as partitioned and unpartitioned conditions enjoy similar visual presentation conditions. Finally, Experiment 3 demonstrated that when subsets differ in category, a selective memory search process takes place even when the surrounding visual distractors are categorically-heterogeneous. Furthermore, features common to a target category may guide attention to likely targets, including lure items. These lures, belonging to the either of the prioritized categories (i.e., non-target objects from the relevant category or target objects from the irrelevant category), can lead to distraction. Despite this distraction, participants generally exhibit relatively high accuracy rates by successfully refraining from responding to the lure items.

As mentioned in the Introduction, previous findings had failed to demonstrate an efficient partitioning-by-category effect when using memory set sizes comparable to the ones used here. In one study (Boettcher et al., 2018, Experiment 2) no flexible partitioning was seen, perhaps because of a partial semantic overlap between the categories used. Nordfang & Wolfe (2018) did find evidence for memory-partitioning, but only when using very small memory subsets. In that case, the failure to flexibly partition memory into larger subsets may have resulted from an insufficient distinction between the subsets during the memorization phase (i.e., categories were learned as one, intermixed memory set). In the present experiment, flexible, trial-by-trial memory partitioning was obtained with the use of well-defined conceptual borders, promoted by separation of the subsets in time and, perhaps, spatial location, during encoding.

It is well-established that semantic and categorical knowledge serve as structuring principles in LTM (e.g., Anderson & Pichert, 1978; Bartlett, 1932; Bower, 1970; Brewer & Treyens, 1981; Collins & Loftus, 1975; Mandler, 1967; Mervis & Rosch, 1981). As noted earlier, in the present experiments we examined the effect of semantic/categorical memory on ‘activated LTM’, a portion of LTM in which the memorized items are temporarily stored for the subsequent Hybrid search task (aLTM, e.g., Cowan, 1988, 1999; see also Oberauer, 2001, 2002; Woltz & Was, 2007). Cowan defined aLTM as “a subset of features that are in a temporarily activated state, making these items more rapidly and reliably accessible than other items in LTM” (Cowan, 2019). Boettcher et al. (2018) and Cunningham & Wolfe (2014) adapted and adopted the construct of aLTM as part of their account of Hybrid search. Here, aLTM refers to the set(s) of items that are relevant to the current task. This does not involve a strong commitment to any particular model of activated (or available) LTM. Our results demonstrate that this temporary, aLTM store of pre-learned items, can be organized by prior semantic knowledge, and can be used in a flexible way as a function of the relevant retrieval context. Indeed, it is not surprising that such a store is sensitive to, and affected by semantic information (e.g., Cowan, 1988; Woltz & Was, 2006, 2007), since the activated representations are tightly linked to LTM content (see also Baddeley’s ‘episodic buffer’, Baddeley, 2000). The novel aspect of the present study, however, relates to the finding that memory search within aLTM can be efficiently *confined* to a certain semantic category, even when multiple sets of categories are simultaneously activated. While previous research has documented the influence of several simultaneous categories on Hybrid search (e.g., Cunningham & Wolfe, 2014; Wolfe et al., 2017), our current findings offer further insights into the nature of these categorical representations. Specifically, they imply that searched targets that are grouped via strong conceptual and/or categorical links, may be stored in aLTM in a rather isolated manner that allows flexible selection of relevant-only target items. Visual guidance, then, may operate on various categorical templates that may interfere with each other at a category decision, or a response selection stage, but not at the target verification (i.e., memory search) stage. It appears that the processes of visual search and memory search are thus largely independent, with *selective* memory search emerging as an inherent and fundamental characteristic of well-bounded semantic sets.

To offer a more comprehensive understanding of memory partitioning, consider the schematic model that illustrates memory selection within a dynamic environment, as depicted in Figure 10. In this representation, each polygon within a given condition represents one of two memory subsets. The connections between nodes within the polygon signify the strength of item associations, while the lines connecting different polygons denote between-subset connections. When subsets are characterized by highly distinctive categories (A), they exhibit strong intra-subset semantic associations (depicted by thick lines) and weak inter-subset connections (illustrated by dashed, thin lines). The robust associative links among items within each category arise from the structural organization of long-life conceptual knowledge. As the environment is dynamic and changes on a trial-to-trial basis, only one memory subset is prioritized on each trial (shown in black), while the other is somewhat deactivated or attenuated (depicted in gray), contingent upon cue-relevance. However, even the attenuated subset retains sufficient activation to evoke a distracting effect when acting as an irrelevant lure (as observed in Experiment 3). Despite this concurrent activation at the visual-search level, memory search through one relevant subset remains impermeable to the other, owing to the clear semantic boundaries between the subsets. That is, structural constraints in LTM (and, accordingly, aLTM), prevent the leakage of information from one subset to another.

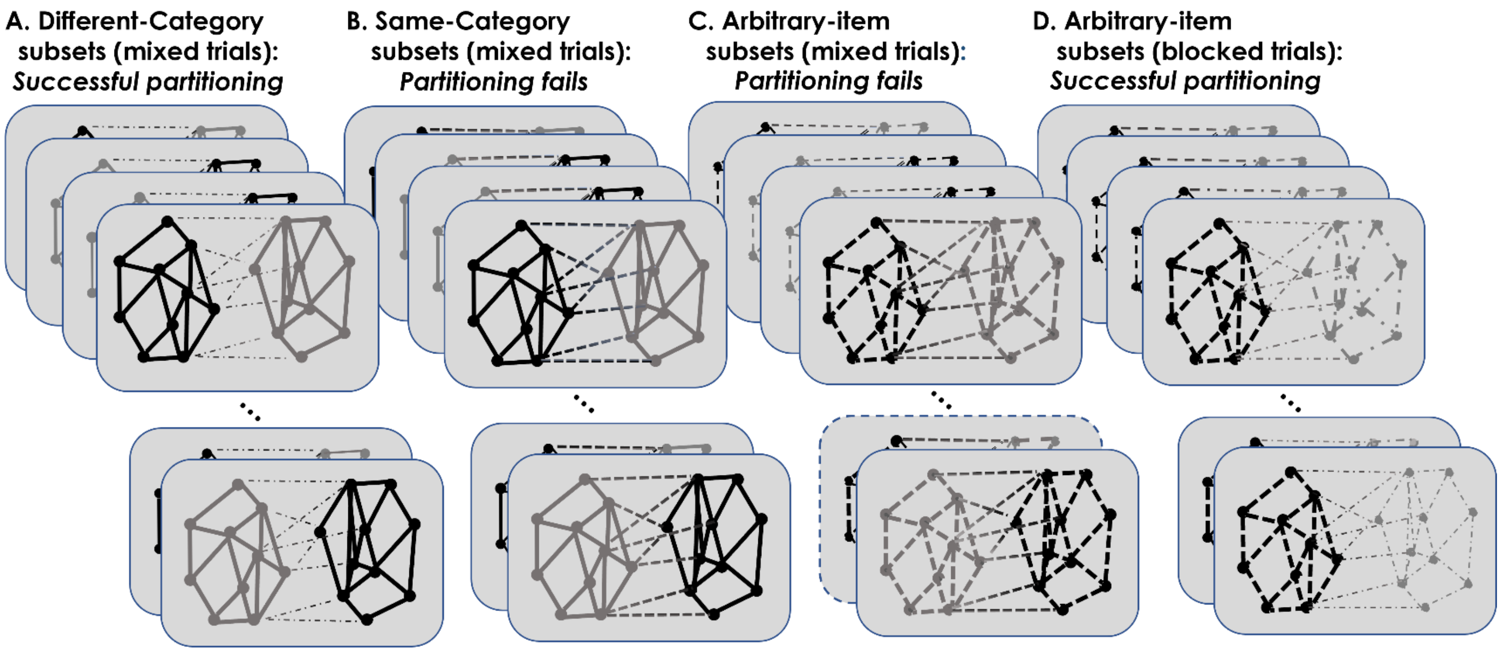

Figure 10. Suggestive model of memory selection within changing (mixed trials, A-C) and fixed, repeated (blocked trials, D) environments. Polygons represents the two memorized subsets, intra-polygon connections signify subset-item association strength, and inter-polygon connections denote between-subset item connection strength. Successful memory partitioning is allowed, on a trial-to-trial basis, when subsets are characterized by highly distinctive semantic categories (A), but not when subsets contain identical categories (B) or arbitrary, unrelated (C) items. Memory search can be confined to the latter type of items only when these are blocked (D), yet under such conditions (e.g., Boettcher et al., 2018) one may doubt whether aLTM is genuinely partitioned. See main text for explanations.

Importantly, the relative immunity of category-dependent targets within memory-search does not apply to ill-defined subsets comprising same-category objects (B). Their strong categorical overlap and inter-subset connections (depicted by thick dashed lines), cause a "spill over" effect during memory-search when simultaneously activated. Based on previous research (Boettcher et al., 2018), a similar situation presumably arises with subsets comprising arbitrary items that lack conceptual distinctiveness. In this scenario (C), intra-subset associations are typically based on temporary, experiment-based training rather than on a long-term, stable aLTM structure, thus rendering them short-lived (denoted by thick, dashed intra-node connections). Importantly, the lack of clear boundaries between subsets makes them penetrable during memory-search processes (indicated by the dashed lines connecting the two subsets). In fact, Boettcher et al. (2018) demonstrated that efficient memory selection can only be achieved with this type of subsets when they are presented in separate blocks. In such cases (D), memory partitioning likely arises from the repeated activation of relevant-set items in aLTM, strengthening their intra-associative connections, alongside the temporary decay of irrelevant-set connections over time and/or task performance (indicated by thin dashed lines within the irrelevant subset, as well as between the two subsets). However, it's worth noting that a block-design lacks the important requirement for the continual activation of *multiple* memory subsets and flexible shifting between them, casting doubt on whether aLTM is genuinely partitioned under such conditions.

Episodes, Events and Memory-Partitioning

From a broader perspective, in addition to contributing to understanding Hybrid search, our findings may shed light on the fundamental principles underlying organization and segmentation processes in long-term memory. Although stimuli in the world are typically perceived within a continuous stream of information, they tend to be encoded, organized and remembered as distinct episodes or events. Extensive research has explored the factors influencing the representation of such bounded events and their segmentation (see, e.g., Newtson, 1973; Zacks & Tversky, 2001; Zacks et al., 2009). Among these factors are changes in temporal and spatial context, as well as conceptual and perceptual elements, which aid mnemonic binding and episodic memory organization (e.g., Baldwin et al., 2001; Ezzyat & Davachi, 2011; Lohnas et al., 2023; Schapiro et al., 2013; Zacks et al., 2009). While speculative, one could consider the subset encoding phase preceding the visual search task in the current research as the creation of a brief "event-like" episode. In this phase, participants explicitly learned groups of images as cohesive memory 'subsets', connected via experimental instructions, and distinguished from other subsets by differences in time and spatial location (or background color in Experiment 1). This spatial and particularly temporal variation among memorized lists likely enhanced associations within each subset, while sharpening boundaries between the different subsets. Interestingly, we found that although participants could clearly distinguish the two subsets on a conscious level, as evident from explicit memory test performance, these subsets were effectively represented as separate 'episodes' only when comprising different semantic categories. Hence, the mere existence of within-set conceptual links (characterizing *both* Different- and Same-category partition conditions), along with between-set spatial and temporal differences, were insufficient for the formation of truly bounded, distinctive episodic 'events'. Furthermore, the findings of Nordfang & Wolfe (2018) suggest that temporal separation between the memorization phases of the different subsets, combined with well-defined categorical membership, might be a critical factor for establishing discrete episodes. Notably, in their study, different categorical subsets were learned as one intermixed memory set, resulting in relative poor partitioning behavior. In the present study, in contrast, categorical subsets were memorized and tested serially, creating a clear episodic distinction between sublists. Although speculative, it appears that effective partitioning-by-category might thus rely on both categorical and temporal/episodic separation during encoding.

Moreover, although event-segmentation studies demonstrate that perceptual changes, such as switches in visual appearance, sound, and particularly movement, contribute to the formation of event boundaries (e.g., Hard et al., 2006; Newtson et al., 1977; Zacks, 2004), the use of color as a distinguishing feature was proven to be inefficient for memory-partitioning in the present study (Experiment 1). It is important to note, however, that in our research, we deliberately chose the color manipulation to be arbitrary and orthogonal to categorical identity. In contrast, previous event-segmentation studies, examining the formation of separated memory episodes, often associated alleged 'perceptual' factors with meaningful actions or goals (Baird & Baldwin, 2001; Baldwin et al., 2001; Zacks, 2004). As such, the potential contribution of purely perceptual features to episodic event formation and segmentation remains to be explored in future studies.

In sum, across a series of three studies, we demonstrated that flexible partitioning of memory into distinct memory subsets, or event-like episodes, is highly efficient when subsets are defined by semantic category. Memory selection in such cases is confined within categorical boundaries, preventing 'slippage' of search to other currently irrelevant subsets. We conclude that despite multiple set activation across an entire block or task, memory search may benefit from relative immunity, operating at least partially independently from visual search.

Footnotes:

1. We computed Bayes factors (referred to as BF, see Dienes, 2011; Jeffreys, 1961), representing the likelihood ratio of the data under the assumption of the presence of an effect. For pairwise comparisons, BF_{10} was used, indicating the Bayes factor in favor of H1 over H0. For ANOVAs, we used the measure of BF_{incl}, that compares models which contain the effect of interest to equivalent models stripped of the effect (an analysis suggested by Sebastiaan Mathôt, see JASP Team, 2022). According to widely accepted benchmarks, BF values of 10 or greater indicate a strong effect, $3<BF<10$ suggests a moderate effect, $0.33 <BF< 3$ implies an inconclusive result, and $BF< 0.33$ indicates the absence of an effect (i.e., a clear support for H0). All Bayesian computations were conducted using the JASP software (2022, version 0.16.3).
2. Cohen's *d* effect size, computed for all pairwise contrasts, is defined as the standardized mean difference between the two conditions. According to common rules of thumb, *d* equaling 0.2, 0.5, and 0.8 stand for small, medium, and large effects, respectively (see J. E. Cohen, 1988).

References:

Anderson, R.C., & Pichert, J.W. (1978). Recall of previously unrecallable information following a shift in perspective. *Journal of Verbal Learning and Verbal Behavior*, *17,* 1–12.

Baddeley, A. (2000). The episodic buffer: a new component of working memory?. *Trends in cognitive sciences*, *4*(11), 417-423.

Baird, J. A., & Baldwin, D. A. (2001). Making sense of human behavior: Action parsing and intentional inference. In B. F. Malle, L. J. Moses, & D. A. Baldwin (Eds.), *Intentions and intentionality: Foundations of social cognition* (pp. 193–206). Cambridge, MA: MIT Press.

Baldwin, D. A., Baird, J. A., Saylor, M. M., & Clark, M. A. (2001). Infants parse dynamic action. *Child Development, 72,* 708–717.

Bartlett, F.C. (1932). *Remembering: A Study in Experimental and Social Psychology*. Cambridge University Press.

Boettcher, S. E., Drew, T., & Wolfe, J. M. (2018). Lost in the supermarket: Quantifying the cost of partitioning memory sets in hybrid search. *Memory & Cognition*, *46*(1), 43-57.

Bower, G. H. (1970). Organizational factors in memory. *Cognitive psychology, 1*(1), 18-46.

Brady, T. F., Konkle, T., Alvarez, G. A. and Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. Proceedings of the National Academy of Sciences, USA, *105*(38), 14325-14329.

Brainard, D. H., & Vision, S. (1997). The psychophysics toolbox. *Spatial vision*, *10*(4), 433-436.

Brewer, W. F., & Treyens, J. C. (1981). Role of schemata in memory for places. *Cognitive Psychology, 13*(2), 207-230. doi:10.1016/0010-0285(81)90008-6

Cohen, J. (1988). Set correlation and contingency tables. *Applied Psychological Measurement, 12*(4), 425-434. doi:10.1177/014662168801200410

Collins, A. M., & Loftus, E. F. (1975). A spreading activation theory of semantic processing. *Psychological Review*, *82*, 407–428.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin, 104*, 163–191.

Cowan, N. (1999). An embedded-processes model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control*. New York: Cambridge University Press.

Cowan, N. (2019). Short-term memory based on activated long-term memory: A review in response to Norris (2017). *Psychological Bulletin, 145*(8), 822-847.
https://doi.org/10.1037/bul0000199

Cunningham, C. A., & Wolfe, J. M. (2014). The role of object categories in hybrid visual and memory search. *Journal of Experimental Psychology: General*, *143*(4), 1585.

Dienes, Z. (2011). Bayesian versus orthodox statistics: Which side are you on?. *Perspectives on Psychological Science*, *6*(3), 274-290.

Drew, T., Boettcher, S. E. P., & Wolfe, J. M. (2017). One visual search, many memory searches: An eye-tracking investigation of hybrid search. Journal of Vision, 17(11), 5-5. Retrieved from https://doi.org/10.1167/17.11.5

Ezzyat, Y., & Davachi, L. (2011). What constitutes an episode in episodic memory? *Psychological Science, 22*(2), 243–252. https://doi.org/10.1177/0956797610393742

Golomb, J. D., Kupitz, C. N., & Thiemann, C. T. (2014). The influence of object location on identity: A "spatial congruency bias." *Journal of Experimental Psychology: General*, *143*(6), Article 6. https://doi.org/10.1037/xge0000017

Gronau, N., Wolfe, J., Nartker, M., & Utochkin, I. (2024, May 6). Hybrid Partition Semantic Unrelated distractors (Exp3). Retrieved from osf.io/zbgtv

Hard, B. M., Tversky, B., & Lang, D. (2006). Making sense of abstract events: Building event schemas. *Memory & Cognition, 34*, 1221–1235.

Hollingworth, A. (2006). Scene and position specificity in visual memory for objects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*(1), 58-69. https://doi.org/10.1037/0278-7393.32.1.58

Hout, M. C., Robbins, A., Godwin, H. J., Fitzsimmons, G., & Scarince, C. (2017). Categorical templates are more useful when features are consistent: Evidence from eye movements during search for societally important vehicles. *Attention,Perception, & Psychophysics*, *79*(6), 1578–1592. DOI 10.3758/s13414-017-1354-1

JASP Team (2022). JASP (Version 0.16.3)[Computer software].

Jeffreys, H. (1961) *Theory of Probability,* 3rd Ed. London: Oxford University Press.

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology, 24*, 175–219. doi:10.1016/0010-0285(92)90007-O

Konkle, T., Brady, T. F., Alvarez, G. A., & Oliva, A. (2010). Conceptual distinctiveness supports detailed visual long-term memory for real-world objects. *Journal of experimental Psychology: general*, *139*(3), 558.

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 animal in preattentive vision. *Perception and Psychophysics, 63*(4), 676–697.

Lohnas, L. J., Healey, M. K., & Davachi, L. (2023). Neural temporal context reinstatement of event structure during memory recall. *Journal of Experimental Psychology: General*.

Mandler, G. (1967). Organization and memory. In *Psychology of learning and motivation* (Vol. 1, pp. 327-372). Academic Press.

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. https://doi.org/10.1080/13506285.2012.735718

Mervis, C. B., & Rosch, E. (1981). Categorization of natural objects. *Annual Review of Psychology, 32*, 89–115.

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences, 7*, 134–140. doi:10.1016/S1364-6613(03)00028-7

Monsell, S., & Driver, J. (Eds.). (2000). *Control of cognitive processes: Attention and performance XVIII*. Cambridge, MA: MIT Press.

Nako, R., Wu, R., Smith, T. J., & Eimer, M. (2014). Item and category-based attentional control during search for real-world objects: Can you find the pants among the pans?. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(4), 1283.

Newtson, D., Engquist, G., & Bois, J. (1977). The objective basis of behavior units. *Journal of Personality and Social Psychology, 35*, 847–862.

Nordfang, M., & Wolfe, J. M. (2018). Guided search through memory. *Visual Cognition*, *26*(4), 285-298.

Oberauer, K. (2001). Removing irrelevant information from working memory: A cognitive aging study with the modified Sternberg task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 948–957.

Oberauer, K. (2002). Access to information in working memory: Exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 411–421.

Pertzov, Y., & Husain, M. (2014). The privileged role of location in visual working memory. *Attention, Perception, & Psychophysics*, *76,* 1914-1924. doi:10.3758/s13414-013-0541-y

Schapiro, A. C., Roger, T. T., Cordova, N. I., Turk-Browne, N. B., & Botvinick, M. M. (2013). Neural representations of events arise from temporal community structure. Nature Neuroscience, 16(4), 486–492. https://doi.org/10.1038/nn.3331

Rubinstein, J. S., Meyer, D. E., & Evans, J. E. (2001). Executive control of cognitive processes in task switching. *Journal of Experimental Psychology: Human Perception and Performance, 27*(4), 763–797. https://doi.org/10.1037/0096-1523.27.4.763

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review, 84*, 1–66.

Swallow, K. M., Barch, D. M., Head, D., Maley, C. J., Holder, D., & Zacks, J. M. (2011). Changes in events alter how people remember recent information. *Journal of Cognitive Neuroscience, 23*(5), 1052–1064. https://doi.org/10.1162/jocn.2010.21524

Woltz, D. J., & Was, C. A. (2006). Availability of long-term memory during and after attention focus in working memory. Memory & Cognition, 34, 668–684.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97–136. https://doi.org/10.1016/0010-0285(80)90005-5

Woltz, D. J., & Was, C. A. (2007). Available but unattended conceptual information in working memory: Temporarily active semantic content or persistent memory for prior operations? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 155–168.

Wolfe, J. M. (1994). Guided Search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review, 1,* 202–238. doi:10.3758/BF03200774

Wolfe, J. M. (2012). Saved by a log: How do humans perform hybrid visual and memory search?. *Psychological Science*, *23*(7), 698-703.

Wolfe, J. M. (2021). Guided Search 6.0: An updated model of visual search. *Psychonomic Bulletin & Review*, *28*(4), 1060-1092.

Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search. *Nature Human Behaviour, 1* (3), 1–8.

Wolfe, J. M., Alaoui Soce, A., & Schill, H. M. (2017). How did I miss that? Developing mixed hybrid visual search as a ‘model system’for incidental finding errors in radiology. *Cognitive Research: Principles and Implications*, *2*, 1-10.

Wolfe, J., Gronau, N., & Nartker, M. (2020, June 22). Hybrid Partition Semantic Visual Factors (Exp. 1). Retrieved from osf.io/jue2y

Wolfe, J., Nartker, M., & Gronau, N. (2023, January 10). Contribution of semantic factors to partitioning memory sets in hybrid search (Exp. 2). Retrieved from osf.io/3vja5

Yang, H., & Zelinsky, G. J. (2009). Visual search is guided to categorically-defined targets. Vision Research, 49(16), 2095-2103. Retrieved from https://doi.org/10.1016/j.visres.2009.05.017

Yu, X., Zhou, Z., Becker, S. I., Boettcher, S. E. P., & Geng, J. J. (2023). Good-enough attentional guidance. Trends in Cognitive Sciences, 27(4), 391-403. Retrieved from https://doi.org/10.1016/j.tics.2023.01.007

Zacks, J. M. (2004). Using movement and intentions to understand simple events. *Cognitive Science*, *28*, 979–1008.

Zacks, J.M., Speer, N.K., & Reynolds, J.R. (2009). Segmentation in reading and film comprehension. *Journal of Experimental Psychology:General*, *138*, 307–327. https://doi.org/10.1037/a0015305

Zacks, J. M., & Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin*, *127*, 3–21. https://doi.org/10.1037/0033-2909.127.1.3