Addleman, D. A., Rajasingh, R., & Störmer, V. S. (2024). Attention to object categories: Selection history determines the breadth of attentional tuning during real-world object search. *Journal of Experimental Psychology: General.* https://doi.org/10.1037/xge0001575

Attention to object categories: Selection history determines the breadth of attentional tuning during real-world object search

Douglas A. Addleman[1,2], Reshma Rajasingh[1], & Viola S. Störmer[1]

[1]Dartmouth College, Department of Psychological and Brain Sciences

[2]Gonzaga University, Department of Psychology

Correspondence to: addleman@gonzaga.edu

Abstract

People excel at learning the statistics of their environments. For instance, people rapidly learn to pay attention to locations that frequently contain visual search targets. Here, we investigated how frequently finding specific objects as search targets influences attentional selection during real-world object search. We investigated how learning that a specific object (e.g., a coat) is task-relevant affects searching for that object and whether a previously frequent target would influence search more broadly for all items of that target's category (e.g., all coats). Across five experiments, one or more objects from a single category were likely targets during a training phase, after which objects from many categories became equally likely to be targets in a neutral testing phase. Participants learned to find a single frequent target object faster than other objects (Exp. 1, N=44). This learning was specific to that object, with no advantage in finding a novel category-matched object (Exp. 2, N=32). In contrast, learning to prioritize multiple exemplars from one category spread to untrained objects from the same category, and this spread was comparable whether participants learned to find two, four, or six exemplars (Exp. 3, N=72). These differences in the breadth of attention were due to variability in the learning environment and not differences in task (Exp. 4, N=24). Finally, a set-size manipulation showed that learning affects attentional guidance itself, not only post-selective processing (Exp. 5, N=96). These experiments demonstrate that the breadth of attentional tuning is flexibly adjusted based on recent experience to effectively address task demands.

Keywords: visual search; category search; selective attention; selection history

Significance: When searching for an object, people can look for its specific visual features – its color, for instance – or they can search more broadly for any object from a category (e.g., cars). But what determines the breadth of attentional tuning in naturalistic search? We hypothesized that recent experience affects how people focus their attention during real-world object search. Participants learned to attend to a particular object or multiple objects from the same category, and we assessed how these different experiences would affect future search. People focused their attention narrowly to a specific object if it was the only object they had previously looked for, but to an entire category if they had gained experience finding two or more objects from that category. This shows that people adaptively learn whether to search narrowly or broadly when looking for real-world objects.

Introduction

Occasionally, we search for items based on their simple visual features: we look for the red button to turn off an oven or for the last house on the left side of the street. But in the real world, attention is rarely guided based on one simple feature alone. More frequently, we have a rich and varied set of information available to help us find what we are looking for. In some cases, we rely on detailed visual information: when looking for our own keys, we may search for very specific shape features. In other cases, we rely on higher-level conceptual information: when looking for jeans in a clothing store, we might search for the category ‘pants’ more broadly. Prior research has shown that depending on task context and instructions, attention can be guided towards the visual details of individual objects or more broadly to sets of objects, potentially even to entire object categories. During visual search, for example, people find objects faster when pre-cued with an image of that specific object or when given only its category via a verbal label (Levin et al., 2001; Potter, 1975).

Visual search usually benefits from tuning attention as narrowly as our knowledge about a target allows. For instance, being cued with a target’s visual details during visual search is more beneficial than knowing only its semantic label (Schmidt & Zelinsky, 2009) or being cued by another object from the same category (Hout & Goldinger, 2015; Vickery et al., 2005; Wolfe et al., 2004). Indeed, knowing only an upcoming target’s category can sometimes even be costly, since attention can spread to non-target items that are perceptually or categorically similar to the target category (Alexander & Zelinsky, 2011; Moores et al., 2003), with even perceptually dissimilar but categorically matched distractors attracting attention (Yeh & Peelen, 2022; see also Reeder & Peelen, 2013). However, when potential search targets are highly uncertain – like when looking for clothing in the store – categorical templates can reduce the likelihood of missing potential targets. Broader templates like those for object categories can also be less effortful to maintain than highly specific ones, and there is evidence that people often use templates that are merely ‘good-enough’ for their current task (reviewed in Yu et al., 2023).

During real-world searches, people rarely have explicit instructions telling them how broadly to direct their attention. In the absence of instructions, the breadth of attention must be determined by other means; for instance, recent work has demonstrated that attention can be guided incidentally, in the absence of instructions, by a person’s experiences (Awh, Belopolsky, & Theeuwes, 2012). In particular, people direct attention to what they attended to in the past – that is, based on their *selection history*. Selection history can create long-term biases to search for locations or low-level perceptual features like color that are frequently associated with targets (Addleman & Störmer, 2023; Geng & Behrmann, 2005; Jiang et al., 2013; Sha et al., 2018). In one example, participants searched for a target letter ‘T’ among distractor ‘L’s, and in a training phase the target was more likely to be in one spatial region of the display than other regions (Jiang et al., 2013). Participants rapidly became faster at finding targets in that frequent region despite no explicit instructions to do so and no self-reported awareness of this spatial bias. They continued prioritizing this spatial region for hundreds of trials of neutral testing after the spatial bias was removed and even across a week-long delay between training and testing. In other studies, these effects of experience have been shown to affect the breadth of attentional templates for visual features. People can track the likelihood that an upcoming target will have a certain color or orientation and adjust their attentional templates to approximate the range of likely target features, or even shift attention towards or away from whole feature dimensions based on which feature dimension was previously more diagnostic in distinguishing targets from distractors (Lee & Geng, 2020; Witkowski & Geng, 2019; Witkowski & Geng, 2022; see also Müller et al., 1995).

Although this work demonstrates how attentional guidance and even the breadth of attentional tuning is affected by experience, it has so far been limited to simple visual features like location or color. It's not clear whether incidental learning can guide attention to the level of real-world objects or entire object categories. This is an important question for research on how attention works in the real world: in everyday life, we usually search for entire objects, such as our keys on a cluttered desk, a car in a parking lot, or milk containers at a supermarket. What role does our experience play when searching for these real-world objects? We hypothesized that attention to objects is strongly shaped by an individual's prior experience, and in particular the nature of recent search targets, similar that what has been shown for simple features (Addleman & Störmer, 2023; Geng & Behrmann, 2005; Jiang et al., 2013; Sha et al., 2018). More specifically, we predicted that attention would be narrowly focused on a specific object if only a particular object was relevant in the past but tuned more broadly – potentially to entire object categories – if a range of objects from a category was relevant previously. We were particularly interested in how much target variability is necessary to induce categorical tuning.

This study examined these topics by instructing participants to search among real-world objects for an orientation-defined target, while, unbeknownst to participants, increasing the likelihood that a specific object (Exp. 1 & 4) or multiple objects from one category (Exp. 3 & 5) would be the target. We examined whether learning facilitated faster search – and attentional guidance in particular – for the frequent target objects and whether this learning would spread to other objects of the same category (e.g., if a particular chair was a frequent target, would learning speed search for other, previously unseen, chairs). To anticipate our results, we found that people rapidly became faster at searching for the frequently experienced real-world object, an effect that continued into a neutral testing phase in which the probabilities of target objects were matched (Exp. 1a&b). When trained on only one frequent target object, learning was specific to exactly that target image (Exp. 2; replicated with a modified task in Exp. 4), but when at least two exemplars from one category were frequently search targets, learning spread to novel objects from the same category (Exp. 3). Interestingly, the extent of spreading to the entire category was essentially the same whether two, four, or six exemplars from the same category were frequently search targets during training. In a final experiment, using a set size manipulation, we showed that these response time advantages were due to attentional guidance and not only due to post-selection processes like the lowering of decision thresholds (Exp. 5). Collectively, these results demonstrate that selection history can adaptively broaden or sharpen attentional tuning to effectively guide visual search for real-world objects.

Experiment 1: People find frequent target objects faster than infrequent ones

This experiment tested whether participants would be biased towards a frequent real-world target object relative to an infrequent target object during visual search, and whether such biases would persist into a testing phase that lacked this frequency manipulation.

Method

Participants. We collected data in Experiment 1 until the end of the academic term, at which point 35 participants completed Experiment 1a and 25 participants completed Experiment 1b. We excluded participants based on three standard criteria used in visual search experiments in our lab (e.g., Addleman & Störmer, 2023), each applied to every epoch of three 56-trial blocks: participants must have had at least 80% accuracy; must have had fewer than 10% of trials with outlier RTs (faster than 200 ms or slower than 5000 ms); and must have taken no longer than 90 minutes to complete the experiment (intended to take 1 hour or less). We excluded 9 participants in Experiment 1a (1 for taking longer than 90 minutes and 8 for below 80% accuracy in at least one epoch) and 7 participants from Experiment 1b (3 for taking longer than 90 minutes, 3 for below 80% accuracy in at least one epoch; 1 for more than 10% outlier RTs in at least one epoch), resulting in data from 26 participants in Experiment 1a and 18 in Experiment 1b used in data analysis. Experiment 1a included 15 women, 9 men, one non-binary person, and one person who declined to report their gender, with a mean age of 21 (range: 18-27). Experiment 1b included 10 women and 8 men with a mean age of 20 (range: 18-25).

In all experiments, participants were volunteers from the University of California, San Diego subject pool compensated with extra course credit. Experiments were approved by the Institutional Review Board at UCSD and Dartmouth College and participants provided informed consent prior to participating. Participants performed the experiment online via their own devices and were asked to complete the experiment on a computer (not a mobile device) in full-screen mode. The experimental code prevented participation from mobile devices and that the display window was always at least 800 by 600 pixels.

Stimuli. A black fixation cross was present at the center of a 600 x 600 pixel display window throughout the experiment except during inter-trial feedback screens. Search items were images from an existing image set (Brady et al., 2013), converted to grayscale to prevent effects of attention to be guided to certain images based on low-level features such as color and presented at 90 by 90 pixels (approximately 1 x 1 degree visual angle on a 13-inch MacBook Pro viewed from 60 cm). Because we were interested in using realistic stimuli that have natural variation in low-level features between and within the categories, we did not control for any other low-level features in our stimulus set. The search array consisted of the same eight search objects per trial for each participant, each chosen to be categorically distinct from the other seven objects (Experiment 1a used one set of eight images and Experiment 1b another set). The same eight objects were present on each trial for each participant, equally spaced along a ring 150 pixels from fixation. The locations of images were randomly determined on each trial. The target image was either upright or upside-down, while each of the seven distractors was randomly rotated either 45 degrees clockwise or counterclockwise. Figure 1A shows an example display (objects are from Experiment 1a). All image files are available on OSF at https://osf.io/e3sfz.

Procedure. Following 28 practice trials, each participant completed 504 experimental trials separated into nine 56-trial blocks. Participants pressed the spacebar to start each trial, after which the search array was presented until they responded. Participants were instructed to indicate the orientation of an upright or upside-down target object among distractor objects tilted 45 degrees and press the ‘up’ arrow key for upright targets or the ‘down’ arrow key for upside-down targets. Instructions emphasized accuracy and speed, encouraging participants to respond as fast as possible and in under 3 seconds whenever possible. Accuracy and response time feedback followed each trial.

For each participant, one object was selected at random to be the frequent target. In the first six blocks (the training phase), that ‘rich’ object was the target on half of trials (versus a chance occurrence of ⅛; 28 trials per rich object per block), while the other half of trials presented each of the seven other ‘sparse’ objects equally often as the target (1/14, or 4 trials per sparse object per block). In the testing phase, all objects occurred equally often as the target in each block (⅛, or 7 trials per object per block). Testing immediately followed training, and participants were neither informed of these probability manipulations nor of any difference between training and testing. The target location and orientation (upright or upside-down) were random on each trial. Figure 1B is a schematic of Experiment 1's probability structure, with objects from Experiment 1b.

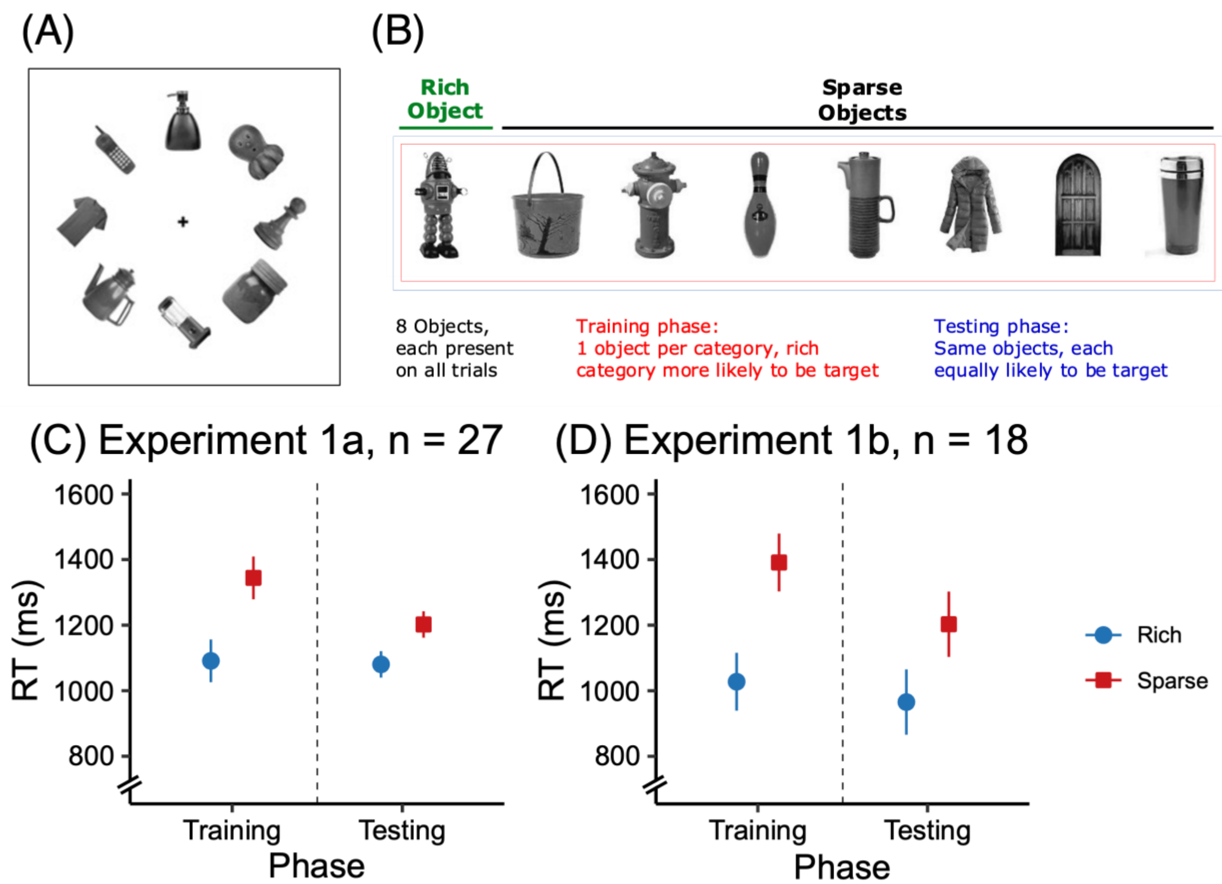

Figure 1. Experiment 1 design and results. A: An example search array from Experiment 1a. Participants searched for the vertical (upright or upside-down) target among tilted distractors and reported the target's orientation. Experiment 1b's task was identical except for the use of eight new images. B. A depiction of the probability manipulation in Experiment 1a and 1b (Experiment 1b's object images are shown). In a 336-trial training phase, one object (selected at random for each participant) was the target more often than the others (on ½ of trials), designed to induce experience-driven attention to the frequent target object. In the 168-trial testing phase, each object was equally likely to be the target, designed to determine if learning during the training phase persisted after the probability manipulation stopped. (C) Mean response times from 27 participants in Experiment 1a for each experimental phase separated by whether the target was the frequent target object ('rich', blue circles) or another object ('sparse', red squares). Results showed faster responses for rich target trials both during training and testing. (D) The same plot for Experiment 1b, which replicated Experiment 1a in 18 new participants with images of different objects. Error bars represent within-subject 95% confidence intervals as recommended in Morey (2008).

Results

Because accuracy was nearly at ceiling (95% in Experiment 1a; 97% in 1b), our main dependent measure was response time (RT) on accurate trials. We investigated whether responses were faster when the frequent ('rich') target object was the target relative to the infrequent target objects ('sparse') separately for the training phase and the testing phase. We grouped data from all blocks in each phase to increase statistical power (and because exploratory analyses suggested that effects were present even in the first block of trials, see Supplementary Materials). RT was analyzed separately for each experiment (1a and 1b) and each phase (training and testing) with dependent-samples *t*-tests comparing rich versus sparse targets (Figure 1C-D). Our primary question in Experiment 1 was whether prior experience shapes target selection during the testing phase; this phase is more clearly interpretable as any biases towards the rich target cannot be explained by inter-trial priming or other short-term effects, instead reflecting evidence of persistent changes in attentional tuning.

Experiment 1a. During training, participants found rich targets faster than sparse ones, mean difference = 258 ms, 95% CI [163, 354], $t(25) = 5.58$, $p < .00001$, $d_z = 1.10$. The same pattern was found during testing, M = 125 ms, 95% CI [66, 184], $t(25) = 4.38$, $p < .001$, $d_z = 0.86$.

Experiment 1b. Experiment 1b replicated Experiment 1a's findings. Response times were faster for rich versus sparse targets during training, mean difference = 363 ms, 95% CI [239, 488], $t(17) = 6.15$, $p < .0001$, $d_z = 1.45$, and during testing, M = 237 ms, 95% CI [96, 378], $t(17) = 3.55$, $p = .002$, $d_z = 0.84$.

Experiment 1 Discussion

We found that selection history sped search for a frequent target object, which persisted throughout a neutral testing phase in which all objects were targets equally often (as in past research for simple visual features; Jiang et al., 2013; Sha et al., 2017). This shows that past experience shapes future search behavior for real-world object images. But what are participants learning about the frequent target objects? Are they prioritizing object-specific visual features or something about the broader category of the object? And do response time benefits arise from more efficient attentional guidance (i.e., change in search slope) or from speeded object identification or orientation discrimination for the frequent target object? We tested these questions in the following experiments, which address whether facilitation is specific to the learned object or generalizes to other within-category objects (Exp. 2-4) and whether response time differences are due to more efficient attentional guidance or post-selective processes (Exp. 5). Collectively, we found clear evidence that the observed response time effects reflect differences in attentional guidance that vary in specificity based on the training environment.

Experiment 2: Learning to attend to one object fails to transfer to a within-category exemplar

Experiment 2 tested whether learning to search for one object facilitates search for other objects from the same category. We tested this with a training phase just like Experiment 1's and a testing phase including objects from the training phase and eight new objects, each matching a different one of the training object categories. The critical question was whether, in the testing phase, search would be biased not only to the previously frequent object but also to the new within-category exemplar.

Method

We preregistered Experiment 2's design, number of participants, exclusion criteria, and analysis plan prior to data collection at https://aspredicted.org/M6L_HVX.

Participants. To ensure an equal number of participants was assigned to each image as the rich object, a sample size divisible by 16 was required. Thus, we collected data until we had 32 participants who passed our pre-registered inclusion criteria (which were identical to those of Experiment 1 with one predetermined adjustment: epochs used in exclusion for accuracy and outlier RTs consisted of two, rather than 3, 56-trial blocks). 47 participants completed Experiment 2, of which 15 were excluded (2 for taking over 90 minutes, 12 for low accuracy, and 1 for too many outlier RTs). Experiment 2 included 23 women, 7 men, and 2 participants who declined to report their gender, with a mean age of 20 (range: 18-26).

Stimuli. Search arrays were identical to those in Experiment 1, but with different object images. Images were manually selected from internet searches to have two exemplars from each of eight image categories: cars, clocks, lamps, grills, plants, shoes, sinks, and trophies, with one image from each category present on a search trial. The 16 images were split into two object sets (set A and set B), with each set containing one exemplar from each of the eight categories, and participants trained on one of the two sets and tested on both sets intermixed. All Experiment 2 images are shown in Figure 2A.

Procedure. The procedure was similar to Experiment 1, with the following differences. Participants completed twelve 56-trial blocks: six training blocks and six testing blocks. In training, participants saw either object set A or object set B (16 participants were trained on each set), with one of the eight training objects appearing more frequently as the target (on ⅜ of trials, reduced from ½ in Experiment 1). This resulted in two training target identities ('rich old' and 'sparse old'). We balanced the rich target image across participants to ensure that any advantages for the target-rich object aren't due to certain objects being intrinsically faster to find.

During the testing phase, the two object sets were randomly intermixed for all participants and all objects appeared equally often. On a given trial, one image from each category was randomly presented, and counterbalancing ensured that each category was the target equally often. Each trial mixed images from the training object set (including the 'rich old' image and 'sparse old' images) and the previously unseen testing object set (which could include an exemplar from the same category as the rich image, 'rich new' and new exemplars from the sparse categories, 'sparse new'), and on each trial all images were equally likely to be the target. Thus, in the test phase, specific distractor images varied across trials, disrupting the possibility of contextual learning and making any persistence of the effect in the test phase unambiguously attributable to learning about the target object and not its visual context. Figure 2A is a schematic of Experiment 2's probability structure.

Results

Accuracy was high (97%) as expected, so analyses were conducted on RTs from accurate trials.

Training phase. We first verified that the probability manipulation induced an effect on visual search response times, as expected from Experiment 1, and found that participants were faster for rich targets relative to sparse targets, mean difference = 329 ms, 95% CI [239, 420], $t(31) = 7.40$, $p < 1*10^{-7}$, $d_z = 1.31$.

Testing phase. Our critical analyses investigated RTs in the testing phase for four target identities using a 2x2 ANOVA with factors of object set (old versus new) and target identity (rich versus sparse), followed by planned post-hoc *t*-tests of the effects of target identity within each object set (i.e., ‘rich old’ vs. ‘sparse old’ and ‘rich new’ vs. ‘sparse new’; see Figure 2B, right). All effects were significant: rich target RT was faster than sparse target RT, $F(1, 31) = 18.43$, $p < .001$, $\eta_p^2 = .37$; RT was faster for images from the old set than the new one, $F(1, 31) = 10.56$, $p = .003$, $\eta_p^2 = .25$, and the two factors interacted, $F(1, 31) = 6.47$, $p = .016$, $\eta_p^2 = .17$. This interaction showed that the advantage for targets from the rich category was driven entirely by the ‘rich old’ image: while ‘rich old’ targets were found faster than ‘sparse old’ targets, mean difference = 200 ms, 95% CI [117, 284], $t(31) = 4.88$, $p < .0001$, $d_z = 0.86$, ‘rich new’ targets were not found any faster than ‘sparse new’ ones, mean difference = 18 ms, 95% CI [-96, 132], $t(31) = 0.32$, $p = .750$, $d_z = 0.06$. To test for evidence in favor of the null hypothesis of no difference between rich new and sparse new images, we conducted an unplanned Bayesian paired *t*-test with a default prior of 0.707. It showed moderate evidence for no difference between the ‘rich new’ and ‘sparse new’ targets, $BF_{10} = 0.19$ (reflecting about 5 times more evidence for the null than the alternative hypothesis). A final unplanned comparison showed that rich old RT was also faster than rich new RT, mean difference = 250 ms, 95% CI [109, 390], $t(31) = 3.62$, $p = .001$, $d_z = 0.64$. Overall, we found no evidence that experience finding a specific object as a search target transferred to benefits searching for another object from the same object category.

Experiment 2 Discussion

The results of Experiment 2 revealed that a learned response time advantage for one specific object during a training phase did not transfer to a previously unseen object from the same category during a testing phase. This indicates that participants acquired a narrow search bias to the exact frequent target object, with no spreading to other within-category objects.

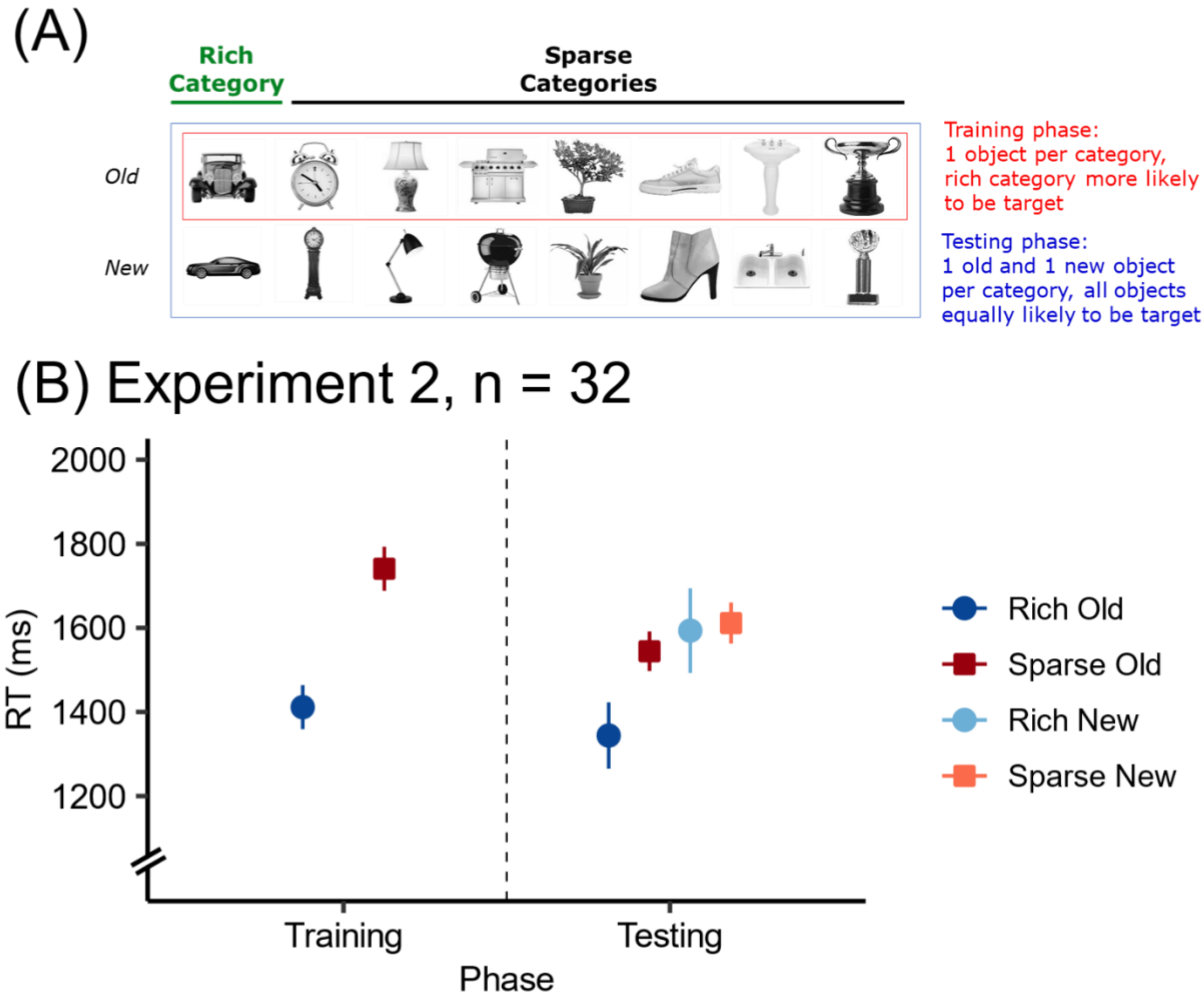

Figure 2. Experiment 2 design and results. (A) A schematic of Experiment 2's design. The 336-trial training phase was similar to that of Experiment 1, where one of two sets of 8 categorically distinct objects was used in search, with one of these objects appearing more often as the target (on 37.5% of trials). In the 336-trial testing phase, a mixture of the old object set and eight new objects from the same categories as the training set were shown, with each object equally likely to be the target. (B) Mean response times from 32 participants in Experiment 2 for the training and testing phases, separated by whether the target was the frequent target object from the training set ('rich old', dark blue circles), a rare target object from the training set ('sparse old', dark red squares), the testing set object from the rich object category ('rich new', light blue circle), or a testing set object from a sparse object category ('sparse new', light red square). Results showed faster responses for the rich object from the training set, but no transfer of this advantage to the categorically matched 'rich new' object. Error bars represent 95% within-subject confidence intervals.

Experiment 3: Learning to attend to multiple objects from a category transfers to new objects from that category

Experiment 2 showed no transfer of an experience-based search advantage for a specific target object to other objects from the same category. Experiment 3 tested whether this is a fundamental limitation of experience-based attention or a result of the rigid training environment containing only one frequent target object. We tested this by exposing participants to two, four, or six category-matched frequent target objects during training and tested if learning of a variety of objects from one category would transfer to new objects from the frequent target category.

Method

We preregistered the sample size, design, and analysis plan of Experiment 3 prior to data collection at https://aspredicted.org/H52_HTG. The preregistration of statistical analyses erroneously omitted the factor of category variability (whether participants saw 2, 4, or 6 images in training) and thus did not describe our intended analyses accurately. In the present analyses we include the between-subjects factor of category variability, as originally planned and appropriate to the research question we are asking.

Participants. We required a sample size divisible by 24 to ensure that an equal number of participants were assigned to each combination of frequent target category (8 options) and number of training images per category (2, 4, or 6). We collected data until we had 72 participants who passed our pre-registered inclusion criteria (identical to those of Experiment 2). 38 were excluded (16 for taking over 90 minutes, 19 for low accuracy, and 3 for too many outlier RTs). These exclusion rates are typical for one-hour online studies, and our strict exclusion criteria are important for ensuring that people are engaged enough in the task to acquire learned biases in the training phase. Experiment 3 included 46 women, 20 men, 3 non-binary people, and 3 participants who declined to report their gender, with a mean age of 21 (range: 18-29).

Stimuli. Search arrays were similar to those in Experiment 1 and 2, except we used new images and the target was always an upright object that participants clicked on with the mouse on each trial (see *Procedure*). Images were manually selected from internet searches to have twelve exemplars from each of eight image categories: coffeemakers, dresses, chairs, lightbulbs, plants, bags, pianos, and bottles, with one image from each category present on a search trial. Each search array still had eight items, one from each of the eight categories. A selection of Experiment 3 objects is shown in Figure 3A (all object images are publicly available on OSF at https://osf.io/e3sfz).

Procedure. The procedure was similar to Experiment 2 with the following exceptions. We adjusted the task to now require a two-stage response, with participants first pressing the spacebar as soon as they found the (now always-upright) target object, at which point the search array was replaced with circular placeholders and participants were asked to take as long as they needed to accurately click on the placeholder at the target's location. Response time was measured as the time to press the spacebar on accurate trials in which participants clicked on the correct location.

Experiment 3 had a training phase (five 56-trial blocks) in which the target was from a specific one of the eight object categories on 37.5% of trials, with the rich category counterbalanced across participants. Unlike previous experiments, training included multiple exemplars from each category: the variability of images in category was now a between-subjects factor, with participants assigned either to the low variability (2 training objects per category), medium variability (4 training objects per category), or high variability (6 training objects per category) condition. On each trial, a random object from each of the eight categories was chosen to be presented at a random location, and then the target was chosen based on the likelihood of its category being the target: ⅜ for the rich category and 5/56 for each of the sparse categories. This meant that the probability that the target will be from the rich category was constant for all participants, but as category variability increased, the probability that any one object would be the target on a given trial decreased. There were still two target identity conditions, 'rich old' (when the target was any of the training set images from the frequent target category) and 'sparse old' (other training set images). In training, we expected to replicate the results of the previous experiments, with faster responses for all frequent target images.

During the testing phase (another five 56-trial blocks), the number of objects participants could see from each category doubled, using random images from the full set of 12 per category. For example, the low variability condition included the two training set images and two new images for each category. There was still one object from each category per trial, and all objects were equally likely to appear as the target (1/8 per category; the variable number of objects per category meant that each image had a 1/32, 1/64, or 1/96 chance of being the target for the low, medium, and high variability conditions respectively). This design led to the same four testing target identities as in Experiment 2: 'rich old' (any training set object from the previously rich category), 'sparse old' (training set objects from non-rich categories), 'rich new' (previously unseen objects from the previously rich category), and 'sparse new' (previously unseen objects from non-rich categories). We were interested in whether increasing variability in the number of images per category increased the spread of attentional biases from the specifically trained images to the category-matched 'rich new' images. Figure 3A is a schematic of Experiment 3's probability structure.

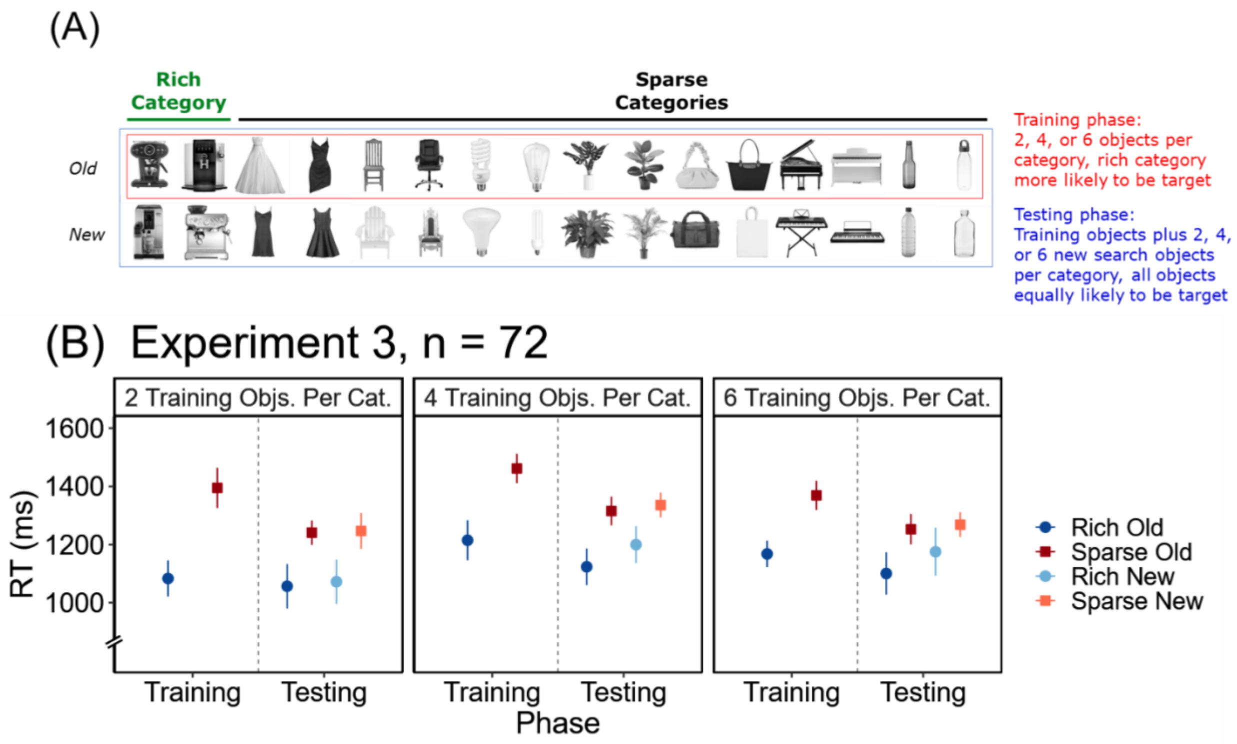

Figure 3. Experiment 3 design and results. (A) A schematic of Experiment 3's design. The 280-trial training phase still consisted of a visual search task among eight objects, one from each category on each trial, but the specific objects varied across trials. Participants were assigned to see two, four, or six objects per category during training, with one of the categories more likely the target category (37.5%; each exemplar from this category was equally likely to be the target). In the 280-trial testing phase, the number of objects per category was doubled, adding two, four, or six objects per category depending on the participant's condition. Each object was equally likely to be the target. The example condition shown has two objects per category in training, adding two more in testing. (B) Mean response times in Experiment 3 (N = 72), split by the between-subjects number of training objects per category (two, four, or six); mean RTs are also separated into training and testing phases and by whether the target was the frequent target object from the training set ('rich old', dark blue circles), a rare target object from the training set ('sparse old', dark red squares), the testing set object from the rich object category ('rich new', light blue circle), or a testing set object from a sparse object category ('sparse new', light red square). All three conditions, unlike the one-object condition of Experiment 2, resulted in spread of attentional biases from the frequently experienced target objects to new objects from the rich category. Error bars represent 95% within-subject confidence intervals.

Results

Accuracy in localizing the target was high (98%), so analyses were conducted on response time to press the spacebar in the first stage of the response on accurate localization trials.

Training phase. A 3x2 ANOVA with a between-subjects factor of category variability (2, 4, or 6 objects per category in training) and a within-subjects factor of target identity (‘rich old’ vs. ‘sparse old’) was conducted on RTs during the training phase (see Figure 3B). RT was faster for ‘rich old’ objects than ‘sparse old’ objects during training, mean difference = 253 ms, 95% CI [205, 301], $F(1, 69) = 112.07$, $p < 1*10^{-15}$, $\eta_p^2 = .62$. There was no main effect of category variability, $F(2, 69) = 0.94$, $p = .395$, $\eta_p^2 = .03$, and no interaction, $F(2, 69) = 1.78$, $p = .176$, $\eta_p^2 = .05$. Overall, training data showed that, although participants were much faster at finding frequent target images, this effect did not significantly differ when there was more variety in the specific objects from the frequent target category being shown.

Testing phase. A 3x2x2 ANOVA with category variability (2, 4, or 6), target identity (rich vs. sparse), and object set (old vs. new) as factors was conducted on RTs in the testing phase (see Figure 3B). RTs to rich category targets were faster than RTs to sparse category targets, mean difference = 161 ms, 95% CI [118, 206], $F(1, 69) = 47.43$, $p < 1x10^{-8}$, $\eta_p^2 = .41$, indicating that learning for a particular category persisted throughout the testing phase. There was also some evidence for faster RTs for the ‘old’ object set than the ‘new’ object set, mean difference = 20 ms, 95% CI [-2, 41], $F(1, 69) = 5.18$, $p = .026$, $\eta_p^2 = .07$, likely reflecting a difference in overall familiarity with the image sets. No other effects were significant, including the effect of category variability ($p = .436$, $\eta_p^2 = .02$) and all interactions (*p*s > .162, η_p^2s < .03). Thus, RTs were not only faster for the familiar ‘rich old’ target objects but also for the new target objects belonging to the frequent category. Follow-up t-tests demonstrated that indeed there was a reliable advantage for ‘rich new’ objects over ‘sparse new’ objects in all three levels of category variability (even though ‘rich new’ objects were never themselves more frequently the target), $t(71) = 4.84$, $p < .00001$, $d_z = 0.57$, indicating that learning transferred to the new objects from the same category. Additional follow-up *t*-tests showed that RTs remained faster during testing for ‘rich old’ objects than ‘sparse old’ ones, mean difference = 176 ms, 95% CI [125, 227] $t(71) = 6.89$, $p < 1x10^{-8}$, $d_z = 0.81$. An unplanned comparison also showed that ‘rich old’ was marginally faster than ‘rich new,’ $t(71) = 2.02$, $p = .047$, $d_z = 0.24$. Interestingly, this pattern did not differ significantly with more variability in objects in each category, as indicated by the lack of interaction between category variability and either target identity or object set.

Experiment 3 Discussion

Experiment 3 showed that experience-driven attention can spread to new objects of the same category, as long as prior experience includes multiple exemplars from the same object category. Surprisingly, this effect was comparable regardless of whether the frequent target category included two, four, or six objects.

Experiment 4: Differences in spread of learning to new objects are due to variability of learning objects, not task differences

We attributed differences between the specificity of learning in Experiment 2 and the spread of learning to new objects from the same category in Experiment 3 to people learning about one versus multiple objects from the rich category, but the two experiments also used different image sets and slightly different tasks: Experiment 2 involved an upright/upside-down discrimination, whereas Experiment 3 involved a two-stage response involving localizing an upright target. To ensure that these task differences or specifics about the image sets did not drive the different results across Experiment 2 and 3, we ran an additional control Experiment 4, which replicates Experiment 2's learning structure with Experiment 3's images and task.

Method

Experiment 4's method was identical to Experiment 3, except with participants learning a single frequent target image (as in Experiment 2). As in Experiment 3, the specific images – taken from the set of 12 exemplars from each of eight categories – used in training and testing were selected at random for each participant (but consistent for each participant), with one object per category during training (and one of these objects the target on 37.5% of trials) and an additional one image per category during testing (and all 16 objects equally likely to be the target).

Experiment 4's sample size, design, and analysis plan were preregistered at https://aspredicted.org/K1F_191. We collected data until we had 24 participants who passed our pre-registered inclusion criteria, which were the same as in Experiments 1-3. We excluded data from 8 participants (4 for taking longer than 90 minutes and 4 for low accuracy). The final 24 participants included 13 women, 7 men, 2 non-binary people, and 2 participants who declined to report their gender, with a mean age of 21 (range: 18-27).

Results

Accuracy in localizing the target was high (96%), so analyses were conducted on response time to press the spacebar in the first stage of the response on accurate localization trials.

Training Phase. A dependent-samples *t*-test of the effect of target identity ('rich old' vs. 'sparse old') was conducted on RTs during the training phase (see Figure 4). RT was faster for 'rich old' objects than 'sparse old' objects during training, mean difference = 236 ms, 95% CI [205, 301], $t(23) = 4.10$, $p < .001$, $d_z = 0.84$.

Testing Phase. We evaluated RTs in the testing phase using a 2x2 repeated measures ANOVA with factors of object set (old versus new) and target identity (rich versus sparse), followed by planned post-hoc *t*-tests of the effects of target identity within each object set (i.e., 'rich old' vs. 'sparse old' and 'rich new' vs. 'sparse new'; see Figure 4, right). A significant main effect of image set indicated that participants responded faster to old images than new ones, $F(1, 23) = 19.91$, $p < .001$, $\eta_p^2 = 0.46$, with no main effect of target identity (rich vs. sparse), $F(1, 23) = 2.43$, $p = .133$, $\eta_p^2 = .10$. These main effects were qualified by a significant image set by target identity interaction, $F(1, 23) = 6.89$, $p = .015$, $\eta_p^2 = .23$. As in Experiment 2, this interaction showed that the rich old object was the only object that people found faster than others. Participants were faster at finding the rich old object than the sparse old objects, mean difference = 167 ms, 95% CI [35, 299], $t(23) = 2.62$, $p = .015$, $d_z = 0.53$. In contrast, they were no faster at finding the rich new image than the sparse new images, mean difference = 8 ms, 95% CI [-124, 140], $t(23) = 0.13$, $p = .900$, $d_z = 0.03$. To test for evidence in favor of the null hypothesis of no difference between rich new and sparse new images, we conducted an unplanned Bayesian paired *t*-test that showed moderate evidence for no difference, $BF_{10} = 0.22$. As in Experiment 2, we also conducted an unplanned comparison to show that rich old RT was faster than rich new RT, mean difference = 215 ms, 95% CI [98, 331], $t(23) = 3.81$, $p < .001$, $d_z = 0.78$.

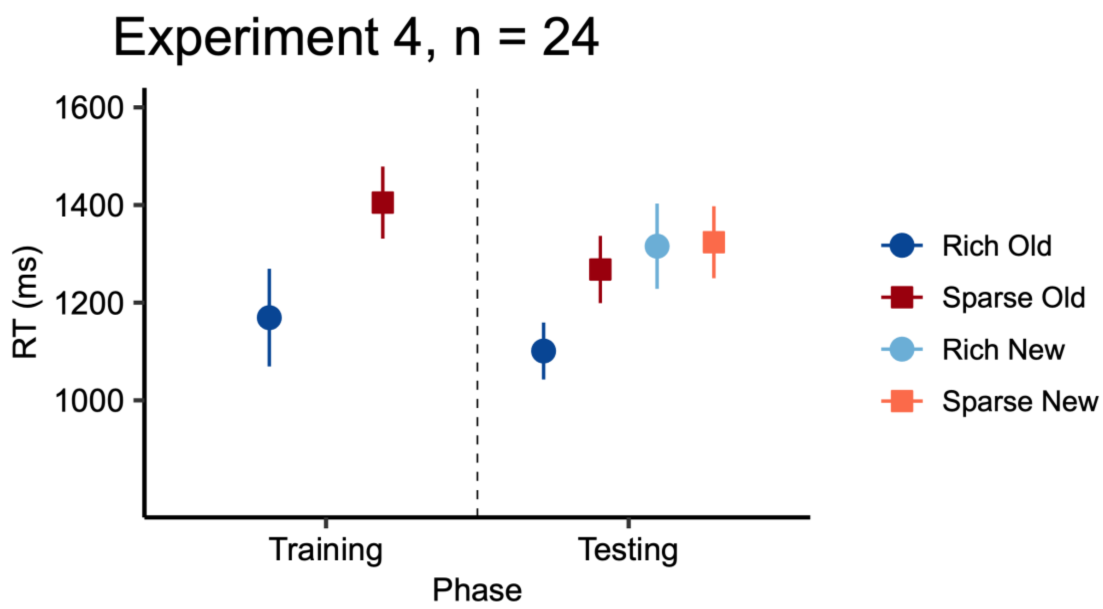

Figure 4. Mean response times in Experiment 4 (N = 24), in which participants learned about a single target object (as in Experiment 2), but with the image set and 2-stage response of Experiment 3. Participants learned to prioritize the frequent target image in training, an effect which persisted for the learned image during testing but did not spread to new objects from the rich category. This is consistent with spread of learning beyond the learned images requiring at least a small amount of variability in the frequent target objects. Error bars represent 95% within-subject confidence intervals.

Experiment 4 Discussion

The results of Experiment 4 show that the difference we found in the spread of learning between Experiment 2 (no spread beyond the learned object) and Experiment 3 (category-wide spread) was due to differences in the variability of the training objects, not the specific task or the image set we used. This confirms that learning to search for an object category is specific to the learned image when people find a single frequent target (Exp. 4 and Exp. 2) but generalizes to new objects from the learned category when people learn about two or more frequent targets (Exp. 3).

Experiment 5: Learning affects attentional guidance, not just post-selection processes

Experiments 1 through 4 show that people learn to respond faster to frequent target objects, and that this learning generalizes to new objects from a learned category when people experienced at least two frequent target objects. However, these experiments can’t disambiguate whether the RT benefits are due to attentional guidance or to faster post-selection processes, for example making the decision to respond to a rich category item more quickly than a sparse category one (i.e., a lower response threshold). In Experiment 5, we tested whether learning affects attentional guidance using a set size manipulation (cf. Donders, 1969/1868; Sternberg, 1969; Wolfe et al., 1989). If faster mean RTs are a result of more efficient search, these effects should increase as a function of set size, leading to shallower search slopes for the faster conditions (e.g., for rich targets than sparse ones). To test whether any guidance induced by learning also spread to new objects differently as a function of training variability, Experiment 5 intermixed trials with eight and with four search items, with each participant in one of two levels of category variability during a training phase. One group of participants learned about a single frequent target (no-variability group; as in Exps. 1, 2 and 4), and another group learned about four frequent targets from a single category (medium-variability group; as in the intermediate variability condition of Exp. 3).

Method

Experiment 5 was similar to Experiment 3, with the following exceptions. In each experimental block, half of the trials presented four search items, and the remaining half of the trials presented eight search items (see Figure 5A). We ran participants in two category variability conditions (learning about 1 object, as in Exp. 4, or 4 objects, as the medium-variability group in Exp. 3) to see if search slopes differed between the two conditions. The target frequency manipulation was counterbalanced within each set size, with targets being from the rich category on 37.5% of training trials. For counterbalancing purposes, number of trials and block structure differed from the other experiments: Participants completed two 112-trial training blocks and two 112-trial testing blocks, all of which intermixed set size four and set size eight trials.

Because we were interested in interactions between set size and target identity, we increased the number of participants per condition to 48 (from 24 in Experiments 3 and 4), for a total of 96 participants in Experiment 5. Experiment 5's sample size, design, and analysis plan were preregistered at https://aspredicted.org/Z9L_HGL. We collected data until we had 48 participants in each category variability condition who passed our pre-registered inclusion criteria, which were the same as previous experiments. We excluded data from 22 participants (8 for taking longer than 90 minutes, 13 for low accuracy, and 1 for too many outlier RTs). The final 96 participants included 78 women, 13 men, and 5 people who declined to report their gender, with a mean age of 21 (range: 18-33).

Results

Accuracy in localizing the target was high (98%), so analyses were conducted on response time to press the spacebar in the first stage of the response on accurate localization trials.

Training Phase. We conducted a 2x2x2 ANOVA with a between-subjects factor of category variability (no-variability or medium-variability) and within-subjects factors of target identity (rich old or sparse old) and set size (4 or 8) on RT (see Figures 5B and 5C). There were significant main effects of target identity, $F(1, 94) = 48.64$, $p < 1\times10^{-9}$, $\eta_p^2 = .23$, and set size, $F(1, 94) = 878.88$, $p < 1 \times 10^{-48}$, $\eta_p^2 = .90$, and a significant interaction between target identity and set size, $F(1, 94) = 24.17$, $p < 1 \times 10^{-5}$, $\eta_p^2 = .21$. This interaction reflects that participants' search slopes were shallower for the rich objects than the sparse ones. Figures 5D and 5E visualize these slopes directly by plotting search slope – the average added time per search item calculated as the difference between set size eight RT and set size four RT divided by four (the number of additional items in set size eight) – on the vertical axis. No other effects, including the three-way interaction, were significant (*p*s > .25).

Testing Phase. In line with our preregistration, we first conducted a 2x2x2x2 ANOVA with a between-subjects factor of category variability (no-variability or medium-variability) and within-subjects factors of target identity (rich or sparse), image set (old or new), and set size (4 or 8) on RT (see Figure 5). There was no significant main effect of category variability, $F < 1$, but significant main effects of target identity (rich faster than sparse), image set (old faster than new), and set size (4 faster than 8), *p*s < .01. There was also a significant two-way interaction between target identity and category variability, $F(1, 94) = 4.62$, $p = .034$, $\eta_p^2 = .05$, indicating that the testing phase RT advantage for the rich targets was greater for the medium-variability group than for the no-variability group. There was also a significant interaction between image set and set size, $F(1, 94) = 6.59$, $p = .012$, $\eta_p^2 = .05$, indicating significantly shallower search slopes for old targets than new ones. Finally, there was a small but statistically unreliable interaction between target identity and set size, $F(1, 94) = 3.88$, $p = .052$, $\eta_p^2 = .04$, in the direction of shallower slopes for rich targets than sparse targets. There was no reliable three-way interaction of target identity, set size, and category variability, $F(1, 94) = 2.84$, $p = .095$, $\eta_p^2 = .03$. All other effects were statistically non-significant (*p*s > .1).

To further explore effects of learning that might differ between variability groups, we ran additional unplanned, repeated-measures ANOVA tests of the effects of image set, target identity, and set size on testing phase RT separately for each variability group.

For the no-variability group, there were unexpectedly no significant main effects or interactions involving target identity (rich vs sparse; Fs < 1). This was somewhat surprising given that all other experiments found consistent effects of target identity during testing (rich old was always faster than sparse old during testing; all *p*'s <0.05, $d_z > 0.5$). Despite the non-significant result, participants' average response times were numerically faster for rich old relative to sparse old targets (48-ms advantage for set size 8 and 23-ms advantage for set size 4), thus following the same pattern we observed in the other experiments, just with a much smaller effect size. Therefore, we think this non-significant result is likely due to statistical noise and should not be interpreted as theoretically meaningful. We did find reliable effects of image set and set size: RTs were faster for old relative to new targets, $F(1, 47) = 7.42$, $p = .009$, $\eta_p^2 = .14$, and for set size four trials than set size eight trials, $F(1, 47) = 456.74$, $p < 1 \times 10^{-15}$, $\eta_p^2 = .91$.

For the medium-variability group, RTs were faster for rich than sparse targets, $F(1, 47) = 15.44$, $p < .001$, $\eta_p^2 = .25$, and an interaction of target identity and set size indicated that this effect was larger for set size eight than set size four, $F(1, 47) = 7.57$, $p = .008$, $\eta_p^2 = .14$. These effects did not differ for old and new images: There was no significant two-way interaction of target identity and image set, $F(1, 47) = 2.03$, $p = .161$, $\eta_p^2 = .04$, and no three-way interaction of target identity, image set, and set size, $F < 1$. There was, however, a significant interaction between image set and set size, $F(1, 47) = 9.20$, $p = .004$, $\eta_p^2 = .16$, indicating that search slopes were shallower for old images than new images (regardless of whether they were rich or sparse images). Furthermore, RTs were faster for old than new targets, $F(1, 47) = 14.07$, $p < .001$, $\eta_p^2 = .23$, and for set size four than set size eight, $F(1, 47) = 698.76$, $p <$ 1*10^-29, $\eta_p^2 = .94$.

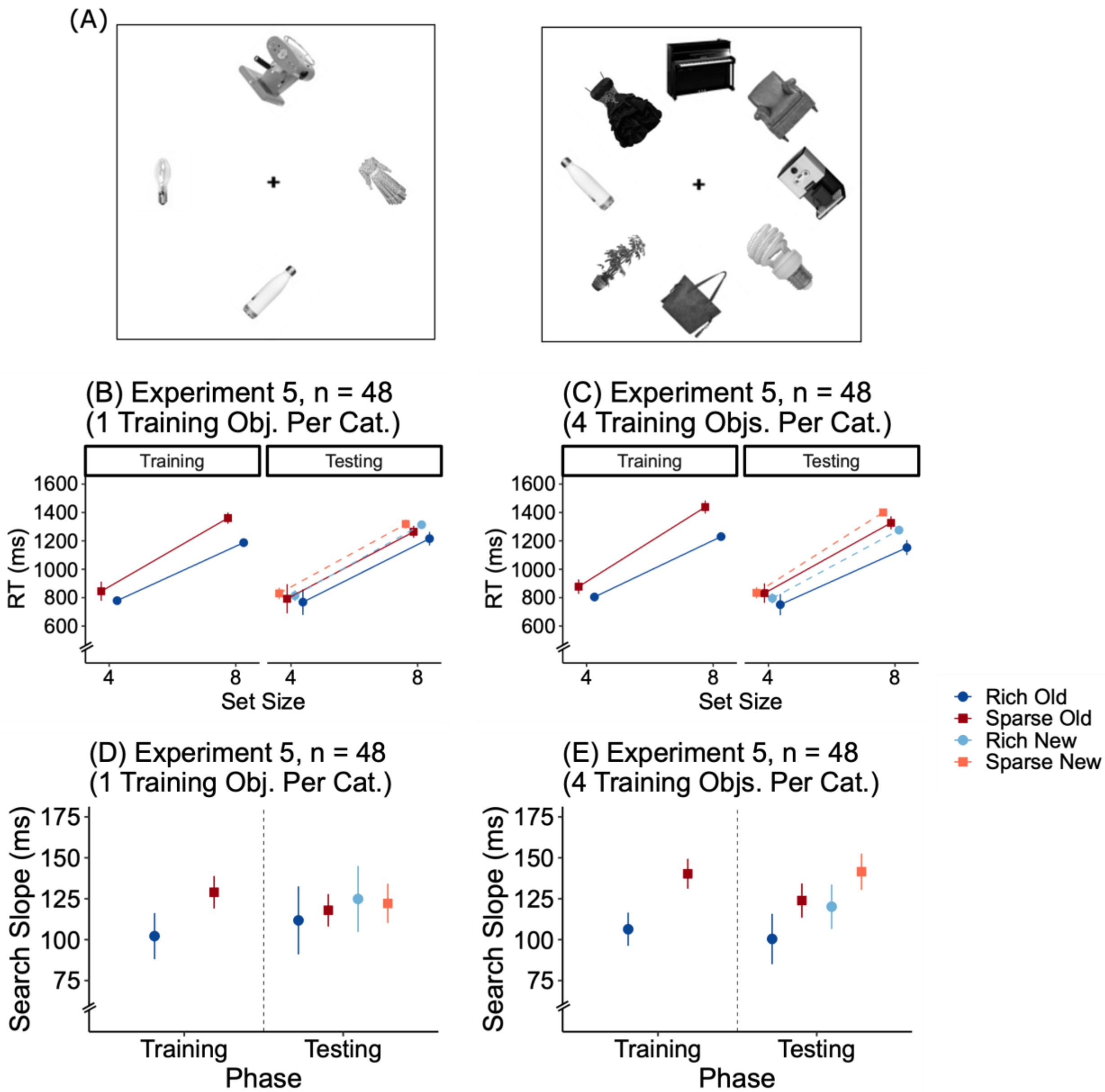

Figure 5. Experiment 5 design and results. (A) Example set size four and set size eight displays. (B) Mean RTs for the no-variability group (N = 48), separately for training and testing phases, plotting set size on the x-axis and split by target identity (rich vs sparse) and image set (old vs new). In training, participants found rich targets faster than sparse ones, and this effect was larger at set size eight than set size four, indicating that participants guided their search towards rich items. In testing, there was no reliable advantage for previously rich objects. Note that this result diverges from all other experiments (see Exp. 5's Discussion for more details). (C) Mean RTs for the medium-variability group (N = 48). Participants found rich targets faster than sparse targets, and this led to larger RT benefits in set size eight trials than set size four trials. In testing, participants persisted in prioritizing the previously rich item, an effect which affected attentional guidance as measured by search slopes. This guidance benefit spread to rich new images. (D) Search slopes for the no-variability group. Slopes indicated that people guided attention towards rich targets during training, with no differences between slopes during testing. (E) Search slopes for the medium-variability group, showing greater search efficiency for rich targets that

persisted into the testing phase and spread to the new images from the rich category. Error bars represent 95% within-subject confidence intervals.

Experiment 5 Discussion

Experiment 5 demonstrates that learning can influence attentional guidance and not just later post-selection processing stages. In the training phase, search slopes were significantly shallower for the rich relative to the sparse category items, regardless of the variability condition. In testing, the initially planned statistical tests failed to find reliable evidence for persistent benefits of learning on search slopes, which was driven by the no-variability group (1 training object per category) that, unexpectedly, did not show reliable persistence of learning in the testing phase. Given that we've found persistence of learning for the frequent target object ('rich old' vs. 'sparse old') across all other experiments, we think the lack of significance in this Exp. 5 is likely a spurious null result due to chance rather than a theoretically important difference from our other experiments. This is particularly likely because the results are qualitatively similar to those from our other no-variability experiments (Exps. 2 and 4), only with smaller effect sizes. That said, one possible (though speculative) theoretical explanation for the smaller effect in Experiment 5 could be that longer-term learning effects may be weaker at set size 4 than 8 because of overall lower task difficulty, which would make it more difficult for learning to introduce reliable RT differences between conditions. Regardless, the most important finding of Experiment 5 – based on our results within each variability condition – is that the medium-variability group (4 training objects per category) showed persistence of learning into testing, with both learned rich images and new category-matched images having faster overall RTs and shallower slopes than sparse items. In fact, despite overall less efficient search for new images than old ones – likely because of practice finding old objects during training – the advantage for rich old versus sparse old objects was numerically comparable to, and not significantly different from, the advantage for rich new objects over sparse new ones (see Fig. 5E). This suggests that people search more efficiently for an entire category of objects after learning that a few exemplars from that category are frequent search targets.

General Discussion

How do people know what to attend when searching for real-world objects? We found that recent experience adaptively shapes attention to focus on either individual objects or object categories. When people frequently selected a specific object during a training phase, this biased their attention persistently to that particular target object but provided no advantage in finding other within-category objects. However, when participants experienced finding multiple frequent targets from one category, this attentional bias rapidly spread to new objects from the same category. Importantly, as shown by Experiment 5, these biases operated at an early attentional selection stage (guidance), rather than later post-selection stages. This suggests that people can readily tune attention to object categories if prior experience in a search environment provides sufficient information about the relevant category, which appears to be possible with as few as two exemplars from the same category.

Selection history influences real-world object search

Our study demonstrates that selection history can rapidly and persistently guide visual search towards real-world object identities or categories, and not only towards lower-level features like location and color as past work has shown (Addleman & Störmer, 2023; Geng & Behrmann, 2005; Jiang et al., 2013; Sha et al., 2017). In Experiment 5, we used a set size manipulation and found that incidental learning affects how people guide attention to relevant objects during search and not only later post-selection processes, such as lowering decision thresholds. This is consistent with demonstrations that learning about frequent target locations guides attention as well (Geng & Behrmann, 2005; Golan & Lamy, 2023; Jiang et al., 2015). Our results reveal that incidental experience is even more powerful in its effects on attention than previously assumed, as the search templates required to direct attention to certain real-world objects, or object categories, are considerably more complex than those for attending to low-level visual features. They also suggest that current mechanistic explanations of how experience influences attention to non-spatial features – which primarily involve gain modulations of attentional priority maps (Ferrante et al., 2018; Stilwell et al., 2019) – need to be expanded to account for how experience can bias attention at higher levels of complexity than simple features.

What might support learning to search for complex objects and object categories? There is evidence that templates for objects or object categories rely on low-and mid-level visual feature combinations common to the searched-for objects or categories (Godwin, Hout, & Menneer, 2014; Schmidt & Zelinsky, 2009; for a review, see Wolfe & Horowitz, 2017). For instance, people can guide search relatively efficiently for categorically unique targets even when these have been scrambled to be semantically uninterpretable as long as they retain the object's basic form and texture (Long et al., 2017). These effects can be explained by object categories having perceptual correlates that are frequently shared among objects of the same category (Long et al., 2016; Long et al., 2017; Rosch et al., 1976; Torralba & Oliva, 2003; this also has implications for how the brain processes semantic content of natural images, Lescroart et al., 2015). Thus, we hypothesize that guidance towards these low- and mid-level feature bundles that constitute common perceptual differences between object categories drives our effects.

On this account, learning about a single frequent target image tunes attention to the features of that image in particular (e.g., orientations, spatial frequencies, shapes, or textures that distinguish that image from other search items), but learning about multiple targets broadens attentional tuning to the common perceptual correlates of that entire object category. The breadth of this tuning would also depend on the variability of perceptual features within the learned category: Specifically, for categories of highly similar objects (e.g., donuts), attention would be tuned to a narrower set of visual features (and might operate more efficiently) compared to low-similarity object categories (e.g., leaves from many different species of trees). To what degree this 'perceptual compactness' of objects within a category modulates the experience-based effects on guidance, and how this may relate to semantic distinctiveness within a category, is an important question for future research. In the present set of experiments, target object categories were counterbalanced across participants to ensure that the effects would not be due to one particular object category.[1] However, future studies could more systematically investigate the roles of perceptual and conceptual distinctiveness in guiding search towards real-world objects via experience, perhaps using quantitative metrics of categorical and perceptual similarity generated by human observers or computer models.

[1] At the same time, this counterbalancing did not provide sufficient statistical power to test for category-specific effects.

Attentional tuning to object categories

We found a striking divergence in the breadth of attentional tuning across our experiments, with attention completely failing to spread to new within-category objects when experiencing only one frequent target object in Experiments 2 and 4, and spreading almost completely to new within-category objects when experiencing two, four, or six frequent target objects in Experiments 3 and 5, with no difference between these variability conditions. One might have instead expected a more gradual emergence of attentional spreading to other objects as the number of learned objects increased during training. The lack of such a gradual effect suggests that participants, rather than employing fine-grained variations in attentional tuning, largely adopted one of two ‘modes’: searching based on features specific to one object or based on the features defining an entire object category (leading to almost complete spread to all objects from that category). What is particularly surprising, however, is that participants appeared to switch to this category-level search strategy abruptly, with as few as two exemplars from the frequent target category.

While potentially counterintuitive, this tendency to switch abruptly between these two modes could be adaptive in many real-world situations. The need to search for either one specific object or any object from a category may be more common than searching for any of a few specific objects, making people inclined to interpret experience with a handful of search targets from a category as evidence that the entire category is task-relevant. Biasing search towards multiple potential targets from a category could also be costly, as differentiating several targets from other within-category objects would require a highly complex search template, resulting in a preference for a less effortful tuning of attention to features defining the searched-for objects’ category. Regardless of why people abruptly transitioned between attending to one object or its category, the lack of transfer from one frequent target object to another in Experiments 2 and 4 contrasts with evidence that explicitly cueing a specific object leads to some spread of attention to other within-category objects (Hout & Goldinger, 2015; Moores et al., 2003; Yeh & Peelen, 2022). These differences may arise from how attentional biases are induced, as there are many established differences between experience-driven and goal-driven attention (e.g., Addleman et al., 2018; for a review, see Jiang, 2019). Based on the current findings, we think that studying how experience-driven attention influences search for realistic objects is a promising direction for researchers trying to understand the contents of attentional templates.

The flexibility with which participants tuned their attention more or less broadly in our experiments builds on findings about attention to ranges of simple visual features (e.g., attending to multiple warm colors). People can track the predictability of an upcoming target feature (i.e., color or orientation) and adjust their attentional templates to approximate the range of likely target features, or even shift attention towards or away from whole feature dimensions based on which feature dimension was previously more diagnostic in distinguishing targets from distractors (Lee & Geng, 2020; Witkowski & Geng, 2019; Witkowski & Geng, 2022; see also Müller et al., 1995). Attention can also be biased away from recently irrelevant feature distributions, even when these change throughout the course of an experiment (Chetverikov et al., 2016, 2017). We’ve shown here that this flexibility is not limited to simple visual features, but also applies to search for complex, real-world objects and even object categories.

Constraints on Generality

We recruited adults from the University of California San Diego community between the ages of 18 and 35. Selective attention is known to change throughout healthy aging, but we would expect the general pattern of results to be comparable for younger or older adults, even if overall performance and effect size of learning varies with age. Participants were representative in terms of race, culture, and ethnicity of the and University of California San Diego community. Although we would not expect large differences in experience-driven attention across cultures, visual categorization of object images may differ, which could lead to differences in our measures of the breadth of attentional tuning for a given set of object stimuli.

Conclusion

The present experiments begin to bridge a gap between laboratory research that studies extremely simple visual features, such as orientations and locations, and real-world cognition, adding to a growing literature investigating the central role of experience in supporting attention to naturalistic stimuli (for reviews, see Brady et al., 2019; Wolfe et al., 2011). Our data reveal that existing knowledge about objects and categories can scaffold learned attentional biases, recruiting a lifetime of learning (about object categories) to influence how attention selects realistic stimuli based on recent experience (about search targets). This relates to other research demonstrating how experience can optimize naturalistic attentional behaviors, for example work showing that people can learn color-object associations to support feature-based attention (Bahle et al., 2021; Kershner & Hollingworth, 2022, 2023), learn to search for different real-world objects in different locations (Zhang & Carlisle, 2023), use domain expertise to more efficiently search through familiar categories relative to unfamiliar ones (Hershler & Hochstein, 2009), and attend to recently rewarded object categories (Hickey et al., 2015). The present study contributes to this venture of scaling up attention research to more realistic contexts by showing how recent experience – with search target identities – can combine with a lifetime of learning – about object categories – to adaptively tune attention to various levels of breadth during search for real-world objects.

Acknowledgements

This work was supported by an NSF grant (BCS-1829434) to V.S.S. and by a Dartmouth Leave Term grant to R. R. Thanks to members of the Brady Lab at UCSD and the Störmer Lab at Dartmouth, as well as T. Horowitz, J. Wolfe and three anonymous reviewers for helpful comments. Preliminary versions of this work were presented at the 2022 and 2023 Vision Sciences Society meetings.

Transparency and Openness

Experiment 1 was not preregistered but followed the same preregistered exclusion criteria as Experiments 2 and 3. The design, sample size, and exclusion criteria were formally preregistered at AsPredicted for Experiments 2 through 5 prior to data collection and can be viewed at https://aspredicted.org/M6L_HVX (Exp. 2), https://aspredicted.org/H52_HTG (Exp. 3), https://aspredicted.org/K1F_191 (Exp. 4), and https://aspredicted.org/Z9L_HGL (Exp. 5). Raw data, experimental code, and analysis code for all experiments are available on the Open Science Framework (Addleman, 2023; https://osf.io/e3sfz).

References

Addleman, D. A., & Störmer, V. S. (2023). Distractor ignoring is as effective as target enhancement when incidentally learned but not when explicitly cued. *Attention, Perception & Psychophysics*, *85*(3), 834–844. https://doi.org/10.3758/s13414-022-02588-y

Addleman, D. A., Tao, J., Remington, R. W., & Jiang, Y. V. (2018). Explicit goal-driven attention, unlike implicitly learned attention, spreads to secondary tasks. *Journal of Experimental Psychology. Human Perception and Performance*, *44*(3), 356–366. https://doi.org/10.1037/xhp0000457

Addleman, D. A. (2023, January 23). Attention to object categories: Selection history determines the breadth of attentional tuning during real-world object search. Retrieved from osf.io/e3sfz

Alexander, R. G., & Zelinsky, G. J. (2011). Visual similarity effects in categorical search. *Journal of Vision*, *11*(8). https://doi.org/10.1167/11.8.9

Bahle, B., Kershner, A. M., & Hollingworth, A. (2021). Categorical cuing: Object categories structure the acquisition of statistical regularities to guide visual search. *Journal of Experimental Psychology: General*. https://doi.org/10.1037/xge0001059

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

Brady, T. F., Störmer, V. S., Shafer-Skelton, A., Williams, J. R., Chapman, A. F., & Schill, H. M. (2019). Scaling up visual attention and visual working memory to the real world. In K. D. Federmeier & D. M. Beck (Eds.), *Psychology of Learning and Motivation* (Vol. 70, pp. 29–69). Academic Press. https://doi.org/10.1016/bs.plm.2019.03.001

Chetverikov, A., Campana, G., & Kristjánsson, Á. (2016). Building ensemble representations: How the shape of preceding distractor distributions affects visual search. *Cognition*, *153*, 196–210. https://doi.org/10.1016/j.cognition.2016.04.018

Chetverikov, A., Campana, G., & Kristjánsson, Á. (2017). Rapid learning of visual ensembles. *Journal of Vision*, *17*(2), 21. https://doi.org/10.1167/17.2.21

Donders, F. C. (1969). On the speed of mental processes (W. G. Koster, trans.). *Acta Psychologica*, *30*, 412–431. https://doi.org/10.1016/0001-6918(69)90065-1 (Original work published 1868)

Ferrante, O., Patacca, A., Di Caro, V., Della Libera, C., Santandrea, E., & Chelazzi, L. (2018). Altering spatial priority maps via statistical learning of target selection and distractor filtering. *Cortex*, *102*, 67–95. https://doi.org/10.1016/j.cortex.2017.09.027

Geng, J. J., & Behrmann, M. (2005). Spatial probability as an attentional cue in visual search. *Perception & Psychophysics*, *67*(7), 1252–1268. https://doi.org/10.3758/BF03193557

Godwin, H. J., Hout, M. C., & Menneer, T. (2014). Visual similarity is stronger than semantic similarity in guiding visual search for numbers. *Psychonomic Bulletin & Review*, *21*(3), 689–695. https://doi.org/10.3758/s13423-013-0547-4

Golan, A., & Lamy, D. (2023). Attentional guidance by target-location probability cueing is largely inflexible, long-lasting, and distinct from inter-trial priming. *Journal of Experimental Psychology. Learning, Memory, and Cognition*. https://doi.org/10.1037/xlm0001220

Hershler, O., & Hochstein, S. (2009). The importance of being expert: Top-down attentional control in visual search with photographs. *Attention, Perception & Psychophysics*, *71*(7), 1478–1486. https://doi.org/10.3758/APP.71.7.1478

Hickey, C., Kaiser, D., & Peelen, M. V. (2015). Reward guides attention to object categories in real-world scenes. *Journal of Experimental Psychology. General*, *144*(2), 264–273. https://doi.org/10.1037/a0038627

Hout, M. C., & Goldinger, S. D. (2015). Target templates: The precision of mental representations affects attentional guidance and decision-making in visual search. *Attention, Perception & Psychophysics*, *77*(1), 128–149. https://doi.org/10.3758/s13414-014-0764-6

Jiang, Y. V. (2018). Habitual versus goal-driven attention. *Cortex*, *102*, 107–120. https://doi.org/10.1016/j.cortex.2017.06.018

Jiang, Y. V., Sha, L. Z., & Sisk, C. A. (2018). Experience-guided attention: Uniform and implicit. *Attention, Perception & Psychophysics*, *80*(7), 1647–1653. https://doi.org/10.3758/s13414-018-1585-9

Jiang, Y. V., Swallow, K. M., Rosenbaum, G. M., & Herzig, C. (2013). Rapid acquisition but slow extinction of an attentional bias in space. *Journal of Experimental Psychology. Human Perception and Performance*, *39*(1), 87–99. https://doi.org/10.1037/a0027611

Jiang, Y. V., Swallow, K. M., Won, B.-Y., Cistera, J. D., & Rosenbaum, G. M. (2015). Task specificity of attention training: the case of probability cuing. *Attention, Perception & Psychophysics*, *77*(1), 50–66. https://doi.org/10.3758/s13414-014-0747-7

Kershner, A. M., & Hollingworth, A. (2022). Real-world object categories and scene contexts conjointly structure statistical learning for the guidance of visual search. *Attention, Perception & Psychophysics*, *84*(4), 1304–1316. https://doi.org/10.3758/s13414-022-02475-6

Kershner, A. M., & Hollingworth, A. (2023). Category-specific learning of color, orientation, and position regularities guide visual search. *Journal of Experimental Psychology. Human Perception and Performance*, *49*(6), 907–922. https://doi.org/10.1037/xhp0001098

Lee, J., & Geng, J. J. (2020). Flexible weighting of target features based on distractor context. *Attention, Perception & Psychophysics*, *82*(2), 739–751. https://doi.org/10.3758/s13414-019-01910-5

Lescroart, M. D., Stansbury, D. E., & Gallant, J. L. (2015). Fourier power, subjective distance, and object categories all provide plausible models of BOLD responses in scene-selective visual areas. *Frontiers in Computational Neuroscience*, *9*, 135. https://doi.org/10.3389/fncom.2015.00135

Levin, D. T., Takarae, Y., Miner, A. G., & Keil, F. (2001). Efficient visual search by category: Specifying the features that mark the difference between artifacts and animals in preattentive vision. *Perception & Psychophysics*, *63*(4), 676–697. https://doi.org/10.3758/bf03194429

Long, B., Konkle, T., Cohen, M. A., & Alvarez, G. A. (2016). Mid-level perceptual features distinguish objects of different real-world sizes. *Journal of Experimental Psychology: General*, *145*(1), 95.

Long, B., Störmer, V. S., & Alvarez, G. A. (2017). Mid-level perceptual features contain early cues to animacy. *Journal of Vision*, *17*(6), 20. https://doi.org/10.1167/17.6.20

Moores, E., Laiti, L., & Chelazzi, L. (2003). Associative knowledge controls deployment of visual selective attention. *Nature Neuroscience*, *6*(2), 182–189. https://doi.org/10.1038/nn996

Müller, H. J., Heller, D., & Ziegler, J. (1995). Visual search for singleton feature targets within and across feature dimensions. *Perception & Psychophysics*, *57*(1), 1–17. https://doi.org/10.3758/bf03211845

Potter, M. C. (1975). Meaning in visual search. *Science*, *187*(4180), 965–966. https://doi.org/10.1126/science.1145183

Reeder, R. R., & Peelen, M. V. (2013). The contents of the search template for category-level search in natural scenes. *Journal of Vision*, *13*(3), 13. https://doi.org/10.1167/13.3.13

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, *8*(3), 382–439. https://doi.org/10.1016/0010-0285(76)90013-x

Schmidt, J., & Zelinsky, G. J. (2009). Search guidance is proportional to the categorical specificity of a target cue. *The Quarterly Journal of Experimental Psychology*, *62*(10), 1904–1914. https://doi.org/10.1080/17470210902853530

Sha, L. Z., Remington, R. W., & Jiang, Y. V. (2017). Short-term and long-term attentional biases to frequently encountered target features. *Attention, Perception & Psychophysics*, *79*(5), 1311–1322. https://doi.org/10.3758/s13414-017-1317-6

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica*, *30*, 276–315. https://doi.org/10.1016/0001-6918(69)90055-9

Torralba, A., & Oliva, A. (2003). Statistics of natural image categories. *Network*, *14*(3), 391–412. https://www.ncbi.nlm.nih.gov/pubmed/12938764

Vickery, T. J., King, L.-W., & Jiang, Y. (2005). Setting up the target template in visual search. *Journal of Vision*, *5*(1), 81–92. https://doi.org/10.1167/5.1.8

Witkowski, P., & Geng, J. J. (2019). Learned feature variance is encoded in the target template and drives visual search. *Visual Cognition*, *27*(5-8), 487–501. https://doi.org/10.1080/13506285.2019.1645779

Witkowski, P. P., & Geng, J. J. (2022). Attentional priority is determined by predicted feature distributions. *Journal of Experimental Psychology. Human Perception and Performance*. https://doi.org/10.1037/xhp0001041

Wolfe, J. M., Horowitz, T. S., Kenner, N., Hyle, M., & Vasan, N. (2004). How fast can you change your mind? The speed of top-down guidance in visual search. *Vision Research*, *44*(12), 1411–1426. https://doi.org/10.1016/j.visres.2003.11.024

Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology. Human Perception and Performance*, *15*(3), 419–433. https://doi.org/10.1037/0096-1523.15.3.419

Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search. *Nature Human Behaviour*, *1*(3), 1–8. https://doi.org/10.1038/s41562-017-0058

Wolfe, J. M., Võ, M. L.-H., Evans, K. K., & Greene, M. R. (2011). Visual search in scenes involves selective and nonselective pathways. *Trends in Cognitive Sciences*, *15*(2), 77–84. https://doi.org/10.1016/j.tics.2010.12.001

Yeh, L.-C., & Peelen, M. V. (2022). The time course of categorical and perceptual similarity effects in visual search. *Journal of Experimental Psychology. Human Perception and Performance*. https://doi.org/10.1037/xhp0001034

Zhang, Z., & Carlisle, N. B. (2023). Explicit attentional goals unlock implicit spatial statistical learning. *Journal of Experimental Psychology. General*. https://doi.org/10.1037/xge0001368