

Journal of Experimental Psychology: Human Perception and Performance

Failures of Perception in the Low-Prevalence Effect: Evidence From Active and Passive Visual Search

Michael C. Hout, Stephen C. Walenchok, Stephen D. Goldinger, and Jeremy M. Wolfe
Online First Publication, April 27, 2015. <http://dx.doi.org/10.1037/xhp0000053>

CITATION

Hout, M. C., Walenchok, S. C., Goldinger, S. D., & Wolfe, J. M. (2015, April 27). Failures of Perception in the Low-Prevalence Effect: Evidence From Active and Passive Visual Search. *Journal of Experimental Psychology: Human Perception and Performance*. Advance online publication. <http://dx.doi.org/10.1037/xhp0000053>

Failures of Perception in the Low-Prevalence Effect: Evidence From Active and Passive Visual Search

Michael C. Hout
New Mexico State University

Stephen C. Walenchok and Stephen D. Goldinger
Arizona State University

Jeremy M. Wolfe
Brigham and Women's Hospital, Cambridge, Massachusetts, and Harvard Medical School

In visual search, rare targets are missed disproportionately often. This low-prevalence effect (LPE) is a robust problem with demonstrable societal consequences. What is the source of the LPE? Is it a perceptual bias against rare targets or a later process, such as premature search termination or motor response errors? In 4 experiments, we examined the LPE using standard visual search (with eye tracking) and 2 variants of rapid serial visual presentation (RSVP) in which observers made present/absent decisions after sequences ended. In all experiments, observers looked for 2 target categories (*teddy bear* and *butterfly*) simultaneously. To minimize simple motor errors, caused by repetitive *absent* responses, we held overall target prevalence at 50%, with 1 low-prevalence and 1 high-prevalence target type. Across conditions, observers either searched for targets among other real-world objects or searched for specific bears or butterflies among within-category distractors. We report 4 main results: (a) In standard search, high-prevalence targets were found more quickly and accurately than low-prevalence targets. (b) The LPE persisted in RSVP search, even though observers never terminated search on their own. (c) Eye-tracking analyses showed that high-prevalence targets elicited better attentional guidance and faster perceptual decisions. And (d) even when observers looked directly at low-prevalence targets, they often (12%–34% of trials) failed to detect them. These results strongly argue that low-prevalence misses represent failures of perception when early search termination or motor errors are controlled.

Keywords: visual search, eye movements, target prevalence

When searching for things in the environment, people are affected (knowingly or not) by learned expectations regarding the likelihoods that specific objects will appear in specific places. These expectations can implicitly govern decisions to persist in searching or to give up after some period of failed searching. Sometimes, contextual expectations are very strong (e.g., milk is highly likely, and keys are highly unlikely, to be found in a refrigerator) and will powerfully affect search persistence. For likely target locations, a person will likely search thoroughly, visually revisiting locations despite prior checking. And for unlikely locations, a person may scarcely glance around. Of course,

such learned expectations comprise a spectrum from “surely there” to “surely not there” and may be either explicitly appreciated by searchers or representative of implicit but unconscious biases (see Ishibashi, Kita, & Wolfe, 2012).

The well-documented *low-prevalence effect* (LPE) arises from such learned expectations. The LPE is a robust finding that when targets are rare, they are missed disproportionately often relative to the same objects when target prevalence is high (Wolfe, Horowitz, & Kenner, 2005). This pattern arises even when observers merely think that targets are rare (Reed, Chow, Chew, & Brennan, 2014a, 2014b; Reed, Ryan, McEntee, Evanoff, & Brennan, 2011). In everyday search, the LPE may hinder search efficiency (e.g., a person might repeatedly overlook his or her keys, even though they are visible, because they are in a low-prevalence location), but the costs, beyond mild frustration, are minimal. However, there are some societally important search tasks (e.g., in radiology, in airport baggage screening) that are characterized by low-prevalence targets. Certain tumors are extremely rare, and few travelers try to sneak guns through airport security. In such cases, the LPE becomes especially problematic, because the costs of missing targets can be catastrophic. Expert searchers also experience the LPE (Evans, Birdwell, & Wolfe, 2013; Evans, Tambouret, Evered, Wilbur, & Wolfe, 2011; Wolfe, Brunelli, Rubinstein, & Horowitz, 2013). Importantly, the LPE does not arise because searchers grow lazy over time. Instead, the LPE appears to reflect an implicit biasing toward expectations that rare targets will be absent that is

Michael C. Hout, Department of Psychology, New Mexico State University; Stephen C. Walenchok and Stephen D. Goldinger, Department of Psychology, Arizona State University; Jeremy M. Wolfe, Brigham and Women's Hospital, Cambridge, Massachusetts, and Department of Ophthalmology, Harvard Medical School.

This work was supported by National Institutes of Health Grant R01 HD075800-02 to Stephen D. Goldinger. We thank Jordan Mitman, Roxanne Crosswell, Guadalupe Preciado, Samer Naseredden, Amanda Godinez, and Deanna Masci for assistance in data collection. We thank Jeremy Schwark and Kyle Cave for comments on an earlier version of this article.

Correspondence concerning this article should be addressed to Michael C. Hout, Department of Psychology, New Mexico State University, P.O. Box 30001/MSC 3452, Las Cruces, NM 88003. E-mail: mhout@nmsu.edu

ecologically sensible and likely reflects deeply rooted statistical learning principles that arise across species, as in animals that learn optimal foraging patterns (e.g., Biernaskie, Walker, & Gegear, 2009; Charnov, 1976; see Cain, Vul, Clark, & Mitroff, 2012). This means that searchers will typically be fast and reasonably accurate, at the cost of occasional misses. Although the LPE may be a reasonable response to the environment under most circumstances, the high-stakes nature of certain searches makes the LPE troubling (Mitroff & Biggs, 2014). Given that it is not readily overcome by instruction or sheer willpower (Kunar, Rich, & Wolfe, 2010; Wolfe et al., 2007), it would be useful to have a better understanding of its causes.

What Causes the LPE?

What are the factors that foster elevated miss rates for low-prevalence targets? There are three well-known principles that produce the LPE in laboratory settings: (a) perceptual/decision failures in the presence of ambiguous stimuli, (b) early search termination errors, and (c) motor errors that reflect prepotent responses. We discuss each of these in turn.

Visual search can be conceptualized as a series of paired two-alternative forced-choice decisions (Wolfe & Van Wert, 2010). Each time the searcher encounters a new object (or location), two decisions are required. First, “is this item a target or not?” And, second, “should I continue or terminate search?” Assuming that the searcher’s perceptual sensitivity is stable over time, errors can therefore be influenced by changes in decision criteria at either stage. In most laboratory tasks, the first decision is typically straightforward (e.g., “Is this item a *T* or an *L*?”). In contrast, real-world targets are often ambiguous (e.g., “Is this cancer or not?”). Experts—be they radiologists, airport security officers, or intelligence community image analysts—are highly trained in their specific tasks, but, even so, discrimination of target from nontarget items is imperfect.

A recent study by Schwark, MacDonald, Sandry, and Dolgov (2013) found that when perceptual decision making is difficult, searchers are prone to make prevalence-based decisions. In the authors’ paradigm, participants searched for letter stimuli in cluttered arrays with as many as 300 distractor letters (increasing to 700 in later studies). In these highly jumbled search displays, letters often overlapped, and the juxtapositions of letter fragments could create ambiguous targets. Participants were given three options: Respond by clicking on the target, by clicking on a *target-absent* button, or by clicking on a *target-present* button. The last option would indicate that the searcher was reasonably confident that the target was there but could not perfectly identify it (i.e., a prevalence-based decision rather than a perceptual decision). Participants used the *target-present* button more frequently when search was difficult (indexed by the number of distractors) than when target prevalence was high, suggesting that they relied on cognition when they could not rely on perception. In a similar study, Schwark, Sandry, MacDonald, and Dolgov (2012) used false feedback to manipulate participants’ perceptions of miss errors. When participants were led to believe that they were committing many misses, they shifted their response criteria and began committing more false alarms. In both studies, when participants could not adequately rely on perception, they relied

instead on memory for prior search outcomes (see also Schwark, Sandry, & Dolgov, 2013).

Returning to the multiple-decision nature of visual search, it is clear that the second decision (“Should I continue or terminate search?”) is more complex (see Chun & Wolfe, 1996). How does an observer decide that he or she has accumulated enough evidence to safely conclude that a target is absent? One account of the LPE suggests that it reflects criterion shifts regarding search termination. Wolfe and Van Wert (2010) had participants look through realistic baggage stimuli, searching for weapons. In their first experiment, when target prevalence was high (98%), participants committed more false alarms than they did when prevalence was relatively lower (50%). With respect to reaction time (RT), under high prevalence, target-present RTs were not shortened, but the rare target-absent RTs were greatly lengthened. In Wolfe and Van Wert’s second experiment, prevalence was varied sinusoidally over trials. Target-absent RTs and criteria (as calculated using signal detection theory; Macmillan & Creelman, 1990) tracked the relative prevalence rate across trials, but target-present RTs and sensitivity (d') did not.

Wolfe and Van Wert (2010) simulated their data using a modified drift-diffusion model, varying an internal decision criterion parameter (e.g., “Is this item the target?”) and the model’s quitting threshold. For high-prevalence rates, the criterion was set to increase *yes* responses, and the quitting threshold made target-absent RTs longer (for low-prevalence rates, these parameters were moved in the opposite directions). The model successfully recreated the response patterns that Wolfe and Van Wert observed, suggesting that prevalence rates affect both perceptual decisions and search termination. These findings suggest that in low-prevalence conditions, quitting thresholds may be adjusted too liberally, terminating search before an observer has attended to the target object. Indeed, Rich et al. (2008) tracked the eye movements of participants searching for rotated *T*s among rotated *L*s and found that when searchers missed rare targets, they had often failed to gaze on them at all.

The final contributing factor to the LPE may be an artifact of laboratory procedures; specifically, rhythmically repeated motor responses (e.g., “no-no-no-no . . .”) may cause people to mistakenly respond with a *target-absent* key press even when they actually find targets. Fleck and Mitroff (2007) suggested that low-prevalence conditions lull participants into a motor pattern wherein they prepare target-absent responses in every trial. Fleck and Mitroff replicated prior low-prevalence miss rates during noncorrectable search, but when participants were given the opportunity to correct potential mistakes, miss rates dropped back to very low levels (between 4% and 10%) and did not vary as a function of prevalence rate. Subsequent work has suggested that the LPE cannot be entirely attributed to motor errors (Rich et al., 2008), but such errors do contribute to the effect.

The Current Investigation

In Schwark et al.’s (2012, 2013) studies, participants relied on cognition when they could not trust their perception. In the current study, we investigated whether perceptual errors would contribute to the LPE even if perception was generally trustworthy. Returning to the multiple-decision framework of visual search, we asked this question: When targets are rare, do people miss them because of

Stage 1, perceptual failures? To examine this, we conducted experiments wherein early termination and motor errors were not possible. To preview our results, we found a persistent LPE for rare items, even when overall prevalence was high (50%), making it unlikely that motor errors or early termination drove the effect (cf. Mitroff & Biggs, 2014). We also found a robust LPE when quitting early was not possible and that participants often missed rare targets that they directly examined. Together, we conclude that although motor errors and early quitting may contribute to the LPE, they are not sufficient to explain it. We argue that any account of the LPE must include failures in perceptual decision making.

In the present experiments, participants always searched for two targets at once (e.g., Godwin, Menneer, Cave, Helman, Way, & Donnelly, 2010; Menneer et al., 2012), making target-absent versus target-present responses in every trial. Only one target could appear in any given trial. Overall prevalence was always 50%, which ensured that participants maintained equivalent attentional vigilance (Mackworth, 1970) across conditions and minimized any tendency to favor either response. Crucially, across participants, we varied the relative prevalence rates of the two targets, from an approximately balanced condition (wherein one target type appeared in 20% of trials and the other in 30%) to more extreme conditions (wherein one target type appeared in only 5% of trials, and its counterpart appeared in 45%). Participants always searched for teddy bears and butterflies, with categories counterbalanced in assignment to prevalence conditions. Across experiments, people either searched categorically for any bear or butterfly among varied distractors, or they searched for specific bears or butterflies among other bear or butterfly distractors. In all cases, the targets and distractors were unoccluded, real-world pictures, making the targets entirely unambiguous.

To further reduce the likelihood of motor errors, in the standard oculomotor search tasks, participants ended each trial by pressing the space bar on a keyboard, after which they made an unsped-up present/absent decision. Thus, the speeded response was separated from information about its intended meaning. In Experiments 2 and 3, we used passive rapid serial visual presentation (RSVP) tasks in which stimuli were quickly presented in the center of the screen, one item (or a few items) at a time. Only after the entire stream was presented did participants indicate their decision, thus eliminating their ability to terminate search on their own. Our prior work (Hout & Goldinger, 2010, 2012) has shown that many standard visual search phenomena can be replicated using an RSVP procedure (see also Williams, 2010). Importantly, this technique simplifies participants' decision making, isolating search to Stage 1, perceptual decisions.

Experiment 1

In Experiment 1, participants performed standard oculomotor search. In Experiment 1a, they looked for the categories *teddy bear* and *butterfly* and responded *target-present* when they found an item matching either category among assorted real-world distractors (pictures of, e.g., a computer, a lamp, a dog). In Experiment 1b, task difficulty was increased by having participants search for specific teddy bear and butterfly exemplars presented among similar distractors (i.e., teddy bears and butterflies that differed in

appearance from the cued targets). We predicted that, in both experiments, we would find prevalence effects in accuracy (more miss errors) and RTs (longer target-present RTs). Moreover, we predicted that we would observe a Relative Prevalence Group \times Target Prevalence interaction (in accuracies and RTs) such that the LPE was more pronounced as the difference in prevalence rates between targets became more extreme.

Method

Participants. Sixty-four students from Arizona State University participated in Experiment 1a, and 44 participated in Experiment 1b, in partial fulfillment of a course requirement. Three participants were excluded from analysis for having mean accuracies that were >2.5 standard deviations below their group mean or mean RTs >2.5 standard deviations above their group mean. All participants reported normal or corrected-to-normal vision and normal color vision.

Design. Three levels of relative prevalence (near balanced, unbalanced, extreme) were manipulated between participants. In all conditions, participants searched for two categories simultaneously, and overall prevalence was 50%. In the *near-balanced* condition, one target category appeared in 20% of trials (giving us a slightly low-prevalence category), and the other appeared in 30% of trials (the high-prevalence category). In the *unbalanced* condition, the relative prevalence rates were 10% and 40%, and in the *extreme* condition, the rates were 5% and 45%. One baseline block (of 40 trials) was also administered in which overall prevalence was 50%, but relative prevalence was not manipulated (each category of target appeared in 10 trials). The baseline block was followed by three blocks of 100 experimental trials, during which relative prevalence was manipulated. The experiment always began with 12 practice trials during the instructions phase; these trials were not analyzed.

Stimuli. All stimuli came from the *Massive Memory* database (<http://cvcl.mit.edu/MM/stimuli.html>; Brady, Konkle, Alvarez, & Oliva, 2008; Konkle, Brady, Alvarez, & Oliva, 2010; see also Hout, Goldinger, & Brady, 2014) and were photographs of real-world objects, resized (maintaining original proportions) to a maximum of 2.5° of visual angle (horizontal or vertical), seen from a viewing distance of 55 cm. Images were no smaller than 2.0° of visual angle along either dimension. The pictures contained no background, and a single object or entity was shown per image.

Apparatus. Data were collected on up to 12 computers simultaneously (all had identical hardware and software profiles). Dividing walls separated each viewing station, and experimental sessions were monitored at all times by one or more research assistants. The PCs were Dell (Round Rock, TX) Optiplex 380 systems (3.06 GHz, 3.21 GB RAM) operating at $1,366 \times 768$ resolution on Dell E1912H 18.5-in. monitors (operated at a 60 Hz refresh rate). Displays were controlled by an Intel (Santa Clara, CA) G41 Express chipset, and the operating system was Windows XP (Microsoft, Redmond, WA). E-Prime Version 2.0 software (Psychology Software Tools, Inc., Sharpsburg, PA; Schneider, Eschman, & Zuccolotto, 2002) was used to control the experiment.

Procedure: Visual search. At the beginning of each trial, participants were shown a target cue. When they were ready to begin, they pressed the space bar on the keyboard, which triggered a fixation cross for 500 ms. Afterward, the visual search display

appeared, which remained onscreen until the participant responded or 10 s elapsed (in which case, the trial was marked as inaccurate). Participants rested their fingers on the space bar during search. As soon as they found the target (only one target could appear in any given trial, and participants were informed of this) or determined that both targets were absent, they pressed the space bar; RTs were measured from the onset of the search display to the space bar press. The images were then cleared from view and replaced with a prompt asking the participant to indicate a *present* or *absent* response via the keyboard. Instructions asked participants to respond as quickly as possible (to the search display, not the decision prompt) while still retaining the highest possible accuracy. (No feedback was given.) The set size was 32 objects, with the teddy bear and butterfly categories assigned to low- and high-prevalence conditions in counterbalanced fashion across participants.

In Experiment 1a, the target cue simply reminded the participant to try to find any teddy bear or butterfly (see Figure 1). Targets were randomly selected from a pool of 17 teddy bears or 17 butterflies; distractors were real-world objects from other categories. In Experiment 1b, the target cues were specific exemplars from each category (held constant throughout an entire block of trials). The distractors were other exemplars from the same categories, presented in equal proportions.

Procedure: Search array organization. A search array algorithm was used to create spatial configurations with pseudorandom

organization (see Hout & Goldinger, 2010, 2012, 2015). Eight objects appeared in each quadrant of the display. Each quadrant was broken down into nine equal-sized cells (effectively making the entire display a 6×6 grid). In each trial, images were placed in random cells (per quadrant); specific locations were jittered within these cells to ensure a minimum of 1.5° of visual angle between adjacent images and between any image and the edges of the screen. The target appeared in each quadrant of the display equally often.

Procedure: Distractor selection. In Experiment 1a, distractors were selected quasirandomly from among 236 other categories. Distractors were chosen such that no more than one exemplar per semantic category was represented in any given trial; across trials, no category was repeated until each had been used at least once. The image set contains nearly 4,000 distinct exemplars; no picture was used more than twice over the entire experiment. In Experiment 1b, the distractors were the other 16 exemplars from each category.

Results

The most important results of Experiments 1a and 1b are shown in Figure 2: Low-prevalence targets were found less often and more slowly than high-prevalence targets, especially in the unbalanced and extreme conditions. In the interest of brevity, we report

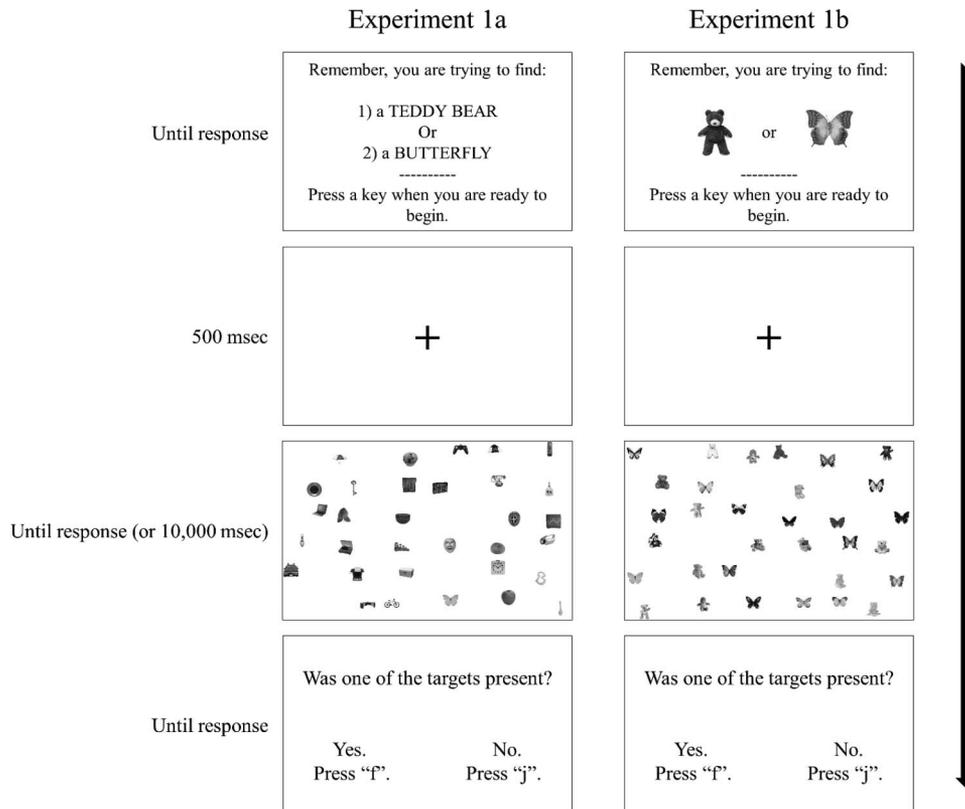


Figure 1. Trial progression in Experiments 1a and 1b. Participants were first shown a target cue, followed by a brief fixation cross, and then the search display. When a participant reached a decision, a space bar press cleared the search array, after which he or she indicated a decision on the keyboard.

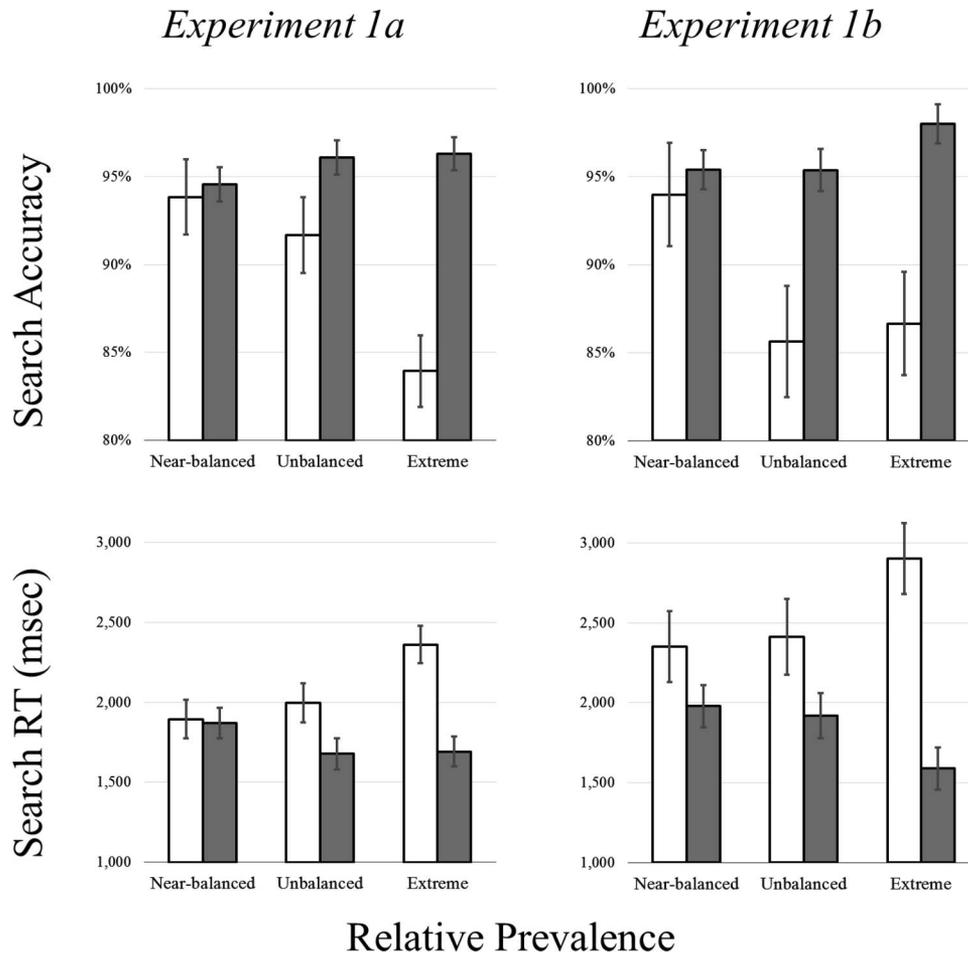


Figure 2. Experimental target-present search accuracy and reaction time (RT) in Experiments 1a and 1b. Results are plotted separately for each between-participants condition of relative prevalence. Light-gray and dark-gray bars show results for the low-prevalence and high-prevalence items, respectively. Error bars represent ± 1 standard error of the mean.

only the most theoretically interesting analyses—specifically, target-present trials in experimental blocks. Target-present results are shown in figures; all other means are reported in tables.

Experimental target-present trials were analyzed using 3 (relative prevalence group) \times 2 (target prevalence) \times 3 (block) mixed-model, repeated measures analyses of variances. Only correct trial RTs were analyzed (see Tables 1 and 2 for Experiment 1 accuracy and RT means, respectively).¹

Accuracy. In Experiment 1a, overall accuracy was 93%. There was no effect of relative prevalence group ($F < 3.00$), but there was a main effect of target prevalence, $F(1, 59) = 33.50$, $p < .01$, $\eta_p^2 = .36$, with more hits to high-prevalence targets (96%) than low-prevalence targets (90%). There was also a Relative Prevalence Group \times Target Prevalence interaction, $F(2, 59) = 11.91$, $p < .01$, $\eta_p^2 = .29$, indicating that the prevalence effect was larger in the less balanced groups (see Figure 2). No other interactions were reliable ($F_s < 2.00$).

In Experiment 1b, overall accuracy was 93%. There was no effect of relative prevalence group ($F < 2.00$). There was a main effect of target prevalence, $F(1, 40) = 18.82$, $p < .01$, $\eta_p^2 = .32$,

with more hits to the high-prevalence category (96%) than the low-prevalence category (89%). There was a Relative Prevalence Group \times Target Prevalence interaction, $F(2, 40) = 3.30$, $p < .05$, $\eta_p^2 = .14$, indicating larger prevalence effects in the less balanced groups (see Figure 2). No other interactions were reliable ($F_s < 3.00$).

RTs. In Experiment 1a, there was no effect of relative prevalence group ($F < 2.00$). There was, however, a main effect of target prevalence, $F(1, 59) = 41.83$, $p < .01$, $\eta_p^2 = .42$, with

¹ Because relative prevalence was not manipulated in the baseline trials, examining these trials allowed us to determine whether there were any preexisting performance differences among the participants who were randomly assigned to each group. If relative prevalence group or target prevalence were shown to have an effect prior to the experimental manipulation, that would suggest potential problems with participant or stimulus selection. We found no effects of relative prevalence group or target prevalence in the baseline trials for Experiments 1a and 1b, suggesting that our findings were not attributable to preexisting group differences or stimulus idiosyncrasies. As such, we did not analyze the baseline trials in Experiments 2–4.

Table 1
Mean Visual Search Accuracies (in Percentages) in Experiments 1a and 1b in Each Between-Participants Condition of Relative Prevalence

Relative prevalence	Trial type	Experiment 1a				Experiment 1b			
		Baseline	Experimental			Baseline	Experimental		
			Block 1	Block 2	Block 3		Block 1	Block 2	Block 3
Near balanced	Target absent	98	98	99	98	99	99	99	99
	Low prevalence	95	92	95	95	97	96	94	91
	High prevalence	96	93	95	96	99	96	96	95
Unbalanced	Target absent	98	99	99	99	98	99	99	99
	Low prevalence	90	90	92	94	91	88	89	79
	High prevalence	96	96	96	96	98	95	95	97
Extreme	Target absent	99	99	99	99	99	100	100	100
	Low prevalence	95	84	83	85	95	89	88	83
	High prevalence	91	96	97	97	99	98	98	98

shorter RTs to the high-prevalence category (1,747 ms) relative to the low-prevalence category (2,084 ms). There was a main effect of block, $F(2, 58) = 11.88, p < .01, \eta_p^2 = .29$, indicating shorter RTs over time (2,120, 1,929, and 1,698 ms for Blocks 1–3, respectively). The Relative Prevalence Group \times Target Prevalence interaction was reliable, $F(2, 59) = 13.00, p < .01, \eta_p^2 = .31$, indicating that the prevalence effect was larger in the less balanced groups. No other interactions were reliable ($F_s < 1.00$).

In Experimental 1b, there was no effect of relative prevalence group ($F < 1.00$). We found a main effect of target prevalence, $F(1, 40) = 41.36, p < .01, \eta_p^2 = .51$, with shorter RTs to the high-prevalence category (1,829 ms) relative to the low-prevalence category (2,555 ms). There was a reliable effect of block, $F(2, 39) = 25.62, p < .01, \eta_p^2 = .57$, indicating shorter RTs over time (2,688, 2,013, and 1,874 ms for Blocks 1–3, respectively). The Relative Prevalence Group \times Target Prevalence interaction was reliable, $F(2, 40) = 7.05, p < .01, \eta_p^2 = .26$, indicating that the effect of target prevalence was exacerbated in the less balanced groups (see Figure 2). No other interactions were reliable ($F_s < 1.00$).

Discussion

The findings from Experiment 1 are easily summarized in three points: (a) Using real-world picture stimuli and a relative preva-

lence manipulation in dual-target search, we found consistently better performance for high-prevalence targets. (b) The differences in accuracy and RT increased as prevalence disparity increased. And (c) these findings were reliable both when participants searched categorically (Experiment 1a) and when they searched using specific target templates (Experiment 1b).

It is unlikely that motor errors contributed to the observed prevalence effects. The two-stage procedure we implemented (wherein participants first quickly indicated search termination and then slowly indicated their decision) discourages misguided button presses. More important, because overall prevalence was 50% in all conditions, participants should not have been biased to respond in either direction, at least not in any differential manner across relative prevalence groups. Early search termination errors, however, were possible in this design. When participants missed targets, they may have failed to place attention on them prior to ending search (a premature-termination error). Alternatively, they may have attended targets but failed to appreciate them as targets (a perceptual decision-making error). To investigate this question, in Experiments 2 and 3, we implemented a passive RSVP procedure in which participants looked directly at every item and indicated target presence/absence decisions after viewing all items. In effect, this made it nearly impossible for participants to miss targets simply because they were not fixated.

Table 2
Mean Visual Search Reaction Times (in Milliseconds) in Experiments 1a and 1b in Each Between-Participants Condition of Relative Prevalence

Relative prevalence	Trial type	Experiment 1a				Experiment 1b			
		Baseline	Experimental			Baseline	Experimental		
			Block 1	Block 2	Block 3		Block 1	Block 2	Block 3
Near balanced	Target absent	6,198	5,240	4,933	4,389	8,061	6,755	5,694	4,876
	Low prevalence	2,403	2,021	1,886	1,779	3,276	2,849	2,094	2,112
	High prevalence	2,336	2,103	1,885	1,624	2,983	2,444	1,872	1,620
Unbalanced	Target absent	5,925	5,097	4,420	3,891	7,622	6,071	4,824	4,358
	Low prevalence	2,291	2,223	1,946	1,819	3,436	2,877	2,212	2,150
	High prevalence	2,245	1,895	1,680	1,459	3,679	2,404	1,753	1,601
Extreme	Target absent	6,423	5,410	4,897	4,416	8,902	7,028	5,509	4,805
	Low prevalence	2,330	2,522	2,496	2,063	3,654	3,567	2,652	2,483
	High prevalence	2,603	1,955	1,679	1,443	2,927	1,991	1,497	1,280

Turning to RTs, there are three possible explanations for the RT benefit for high-prevalence targets. First, observers might have responded more quickly to high-prevalence targets because they were able to locate them more efficiently in space. A second (but not mutually exclusive) possibility is that when high-prevalence targets are examined, they are more quickly identified. Finally, observers might have searched sequentially, first for high-prevalence targets and then for low-prevalence targets if necessary. We consider these hypotheses further in Experiment 4, after we more closely examine whether the LPE persists when search termination is removed from the observer's control.

Experiment 2

In Experiment 2, we investigated whether prevalence effects would arise when participants had to examine every item in the search display and when they could not decide to stop searching of their own accord. Participants searched through RSVP streams in which different images were rapidly presented, one item at a time, in the center of the screen. Once all images were shown, participants indicated whether they had found a designated target. Although RSVP search is inherently different from standard oculomotor search, many of the basic findings in visual search are replicated with such a modified procedure (e.g., Hout & Goldinger, 2010). Recently, we have shown that RSVP search tasks are ideal for investigating object-identification processing during search and that this procedure is amenable to multiple-target search paradigms (Godwin, Walenchok, Houpt, Hout, and Goldinger, in press). Moreover, although the presentation rate in our experiments (100 ms per item) was shorter than typical fixations during search, it has been shown that people can identify target categories in RSVP search with presentations as short as 13 ms (Potter, Wyble, Hagmann, & McCourt, 2014). Thus, this procedure seemed an ideal way to investigate the LPE under conditions in which early termination errors were not possible. As in Experiment 1, there were two versions: In Experiment 2a, participants searched categorically, and in Experiment 2b, they looked for specific bears and butterflies. We predicted, as in Experiment 1, that we would find a prevalence effect in accuracy, with low-prevalence targets missed more frequently than high-prevalence targets, and that this effect would be exacerbated in the more extreme relative prevalence conditions.

Method

The design, stimuli, and apparatus were identical to those in Experiment 1. Again, three levels of relative prevalence (near balanced, unbalanced, extreme) were manipulated between participants, but overall target prevalence remained at 50%. Participants searched for two target categories simultaneously within an RSVP search stream. After completing 12 practice trials, a baseline block was administered, with 40 trials. During this block, overall prevalence was 50%, but relative prevalence was not manipulated. Next, there were three blocks of 100 experimental trials, during which relative prevalence was manipulated using the same proportions as in Experiment 1.

Participants. Fifty-six and 45 new students from Arizona State University participated in Experiments 2a and 2b, respectively, in partial fulfillment of a course requirement. Two participants were excluded from analysis for having mean accuracies >2.5 standard deviations below their group means. All participants had normal or corrected-to-normal vision, and all reported normal color vision.

Procedure. The procedures were largely identical to those in Experiment 1, with two exceptions: Now, following the fixation cross, participants were shown a single-item RSVP stream rather than a standard visual search display (with all items displayed at once). Each image was centrally displayed for 100 ms, followed by a colorful backward mask for 50 ms. Just as there were 32 items in the spatial set size in Experiment 1, there were 32 items per sequence in Experiment 2. Different images were shown in every sequence, and image orders were randomized in each trial. After the search stream terminated, participants were asked whether they had found either target, making a present/absent response as before. Instructions emphasized accuracy; speed was not mentioned, because participants could not respond until all items were shown.

Results

The data were analyzed in identical fashion to the data in Experiment 1, but only accuracy was examined. As before, we report only the results from experimental target-present trials here but present the target-absent data in a table (see Table 3) for completeness. As shown in Figure 3 (mean accuracy for experi-

Table 3
Mean Visual Search Accuracies (in Percentages) in Experiments 2a and 2b in Each Between-Participants Condition of Relative Prevalence

Relative prevalence	Trial type	Experiment 2a				Experiment 2b			
		Baseline	Experimental			Baseline	Experimental		
			Block 1	Block 2	Block 3		Block 1	Block 2	Block 3
Near balanced	Target absent	96	97	97	98	88	91	90	92%
	Low prevalence	98	97	96	96	75	77	82	89
	High prevalence	98	99	96	97	69	78	85	85
Unbalanced	Target absent	96	97	97	98	88	92	93	93
	Low prevalence	97	93	96	97	70	79	91	91
	High prevalence	93	97	97	98	71	84	93	94
Extreme	Target absent	98	95	97	97	88	89	94	96
	Low prevalence	93	95	94	97	77	83	76	71
	High prevalence	94	95	95	95	74	83	94	97

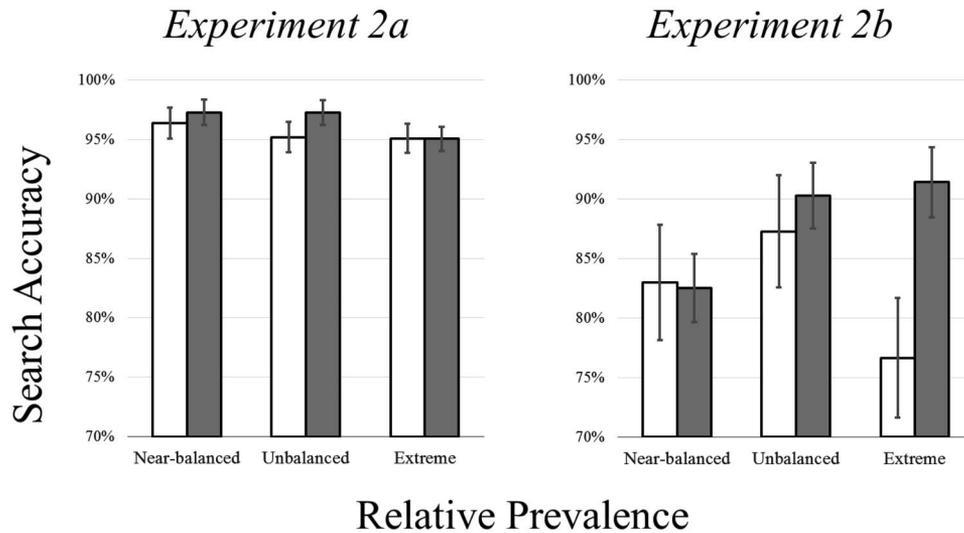


Figure 3. Experimental target-present search accuracy in Experiments 2a and 2b. Results are plotted separately for each between-participants condition of relative prevalence. Light-gray and dark-gray bars show results for the low-prevalence and high-prevalence items, respectively. Error bars represent ± 1 standard error of the mean.

mental, target-present trials), there was no sign of the LPE in Experiment 2a, but the effect emerged in Experiment 2b. In Experiment 2a, overall accuracy was 96%. We found no effects of relative prevalence group, prevalence item, or block ($F_s < 2.00$). None of the interactions were reliable ($F_s < 1.00$). In Experiment 2b, overall accuracy was 85%. There was no effect of relative prevalence group ($F < 2.00$), but there was a main effect of target prevalence, $F(1, 42) = 4.50, p < .05, \eta_p^2 = .10$, with more hits to the high-prevalence category (88%) than the low-prevalence category (82%). There was a main effect of block, $F(2, 41) = 6.98, p < .01, \eta_p^2 = .25$, indicating that performance got better over time (81%, 87%, and 88% for Blocks 1–3, respectively). The Relative Prevalence Group \times Target Prevalence interaction was marginal, $F(2, 42) = 2.77, p = .07, \eta_p^2 = .12$, suggesting a stronger LPE in the more extreme conditions. The three-way Relative Prevalence Group \times Target Prevalence \times Block interaction was reliable, $F(4, 82) = 3.51, p < .05, \eta_p^2 = .15$.

Discussion

In Experiment 2, participants were unable to make early termination errors and were unlikely to make motor response errors. When participants searched categorically (Experiment 2a), we found no prevalence effect, but in the more difficult task of looking for specific teddy bear and butterfly exemplars (Experiment 2b), the prevalence effect emerged. The null prevalence effect in Experiment 2a likely reflected the near-ceiling performance shown in Figure 3. In Experiment 3, we investigated this further by slightly increasing task difficulty.

Experiment 3

In Experiment 3, we conceptually replicated Experiment 2, using another passive RSVP procedure. Now, however, the task was made more difficult by dividing participants' attention across a few centrally presented items. Participants searched through

rapidly presented clusters of images, each containing 2–4 images at a time (similar to the procedure in Shiffrin & Schneider, 1977). The individual images were of sufficient size and shown close together, such that all four images could be simultaneously processed without moving one's eyes off of central fixation. As in Experiment 2, search decisions were indicated after the entire stream was completed. Our predictions were identical to those of Experiment 2: We anticipated a prevalence effect in accuracy, with a larger effect in the more extreme relative prevalence conditions.

Method

The design, stimuli, and apparatus were identical to those in Experiment 2, with the only change being that items were shown in sets of 2–4 images rather than being shown individually. Participants again searched for two target categories simultaneously in a novel RSVP task, and three levels of relative prevalence (near balanced, unbalanced, extreme) were manipulated between participants, with overall target prevalence held at 50%. Participants again completed 12 practice trials, then 40 baseline trials with no prevalence manipulation, then three experimental blocks of 100 trials with differing degrees of prevalence asymmetry.

Participants. Forty-eight and 49 new students from Arizona State University participated in Experiments 3a and 3b, respectively, in partial fulfillment of a course requirement. Six participants were excluded from analysis for reporting anomalous color vision. All other participants reported normal or corrected-to-normal vision and normal color vision.

Procedure. The procedure was identical to that in Experiment 2, with the exception of stimulus presentation. Following the fixation cross, participants were shown an RSVP search stream with either two, three, or four items presented at once. Each image was randomly presented at one of four locations, equidistant from the center of the screen; images were positioned close together such that all could be seen when looking in the display center. Each image set was displayed for 150 ms followed by a colorful

backward mask for 50 ms. In every trial, there were 27 total images, distributed across nine sets, with three sets apiece containing two, three, and four images. Different sizes were presented in random order, and individual images were randomly assigned to each set. As in Experiment 2, participants indicated their response after all images were shown, and instructions emphasized accuracy. In Experiment 3a, participants searched categorically, and in Experiment 3b, they looked for specific bears and butterflies.

Results

The data were analyzed in identical fashion to the data in Experiment 2; results from the critical target-present trials are shown in Figure 4. As shown, there was stronger evidence for the LPE in this experiment relative to the Experiment 2. In Experiment 3a, overall accuracy was 87%. There were no effects of relative prevalence group or block ($F_s < 2.00$). There was a main effect of target prevalence, $F(1, 44) = 18.51, p < .01, \eta_p^2 = .30$, with more hits to the high-prevalence category (91%) than the low-prevalence category (83%). The Relative Prevalence Group \times Target Prevalence interaction was reliable, $F(2, 44) = 3.90, p < .05, \eta_p^2 = .15$, as the prevalence effect increased in more extreme groups. The Target Prevalence \times Block interaction was also reliable, $F(2, 43) = 3.34, p < .05, \eta_p^2 = .13$, indicating that low-prevalence items were missed more frequently over time, despite steady performance to the high-prevalence items (see Table 4). The three-way Relative Prevalence Group \times Target Prevalence \times Block interaction was also reliable, $F(4, 86) = 2.54, p < .05, \eta_p^2 = .11$. No other interactions were reliable ($F_s < 2.00$).

In Experiment 3b, overall accuracy was 58%. There was no effect of relative prevalence group ($F < 3.00$). There was a main effect of target prevalence, $F(1, 41) = 32.33, p < .01, \eta_p^2 = .44$, with more hits to the high-prevalence category (70%) than the low-prevalence category (47%). There was a main effect of block, $F(2, 40) = 3.37, p < .05, \eta_p^2 = .14$, indicating that performance

improved over time (55%, 60%, and 60% for Blocks 1–3, respectively). The Relative Prevalence Group \times Target Prevalence interaction was reliable, $F(2, 41) = 17.74, p < .01, \eta_p^2 = .46$, indicating the larger prevalence effect in the less balanced groups. The three-way Prevalence Group \times Target Prevalence \times Block interaction was reliable, $F(4, 80) = 3.03, p < .05, \eta_p^2 = .13$. No other interactions were reliable ($F_s < 3.00$).

Discussion

In Experiment 3, we again found consistent prevalence effects when early termination errors were not possible. Importantly, when task difficulty was increased, the prevalence effect was restored in categorical search (Experiment 3a). When participants looked for specific exemplars (Experiment 3b), the prevalence effects were of the largest magnitude that we had observed. Overall accuracy in both groups hovered above chance performance. When the two target categories were balanced, they were each found a little more than half the time, but when they were unbalanced, high-prevalence targets were found nearly 20% more often than low-prevalence targets. And in the most extreme condition, high-prevalence targets were more than three times more likely to be found relative to low-prevalence targets. Together with the results of Experiment 2, these findings verify that prevalence effects persist when early search termination and motor errors are removed as possible contributing factors.

Experiment 4

Experiments 2 and 3 established that prevalence effects sometimes occur as a result of perceptual errors. In Experiment 4, we further examined the basis of prevalence effects (including the RT effects noted in Experiment 1) by tracking participants' eye movements during standard visual search. We tested whether high-prevalence targets would be found more often and more quickly

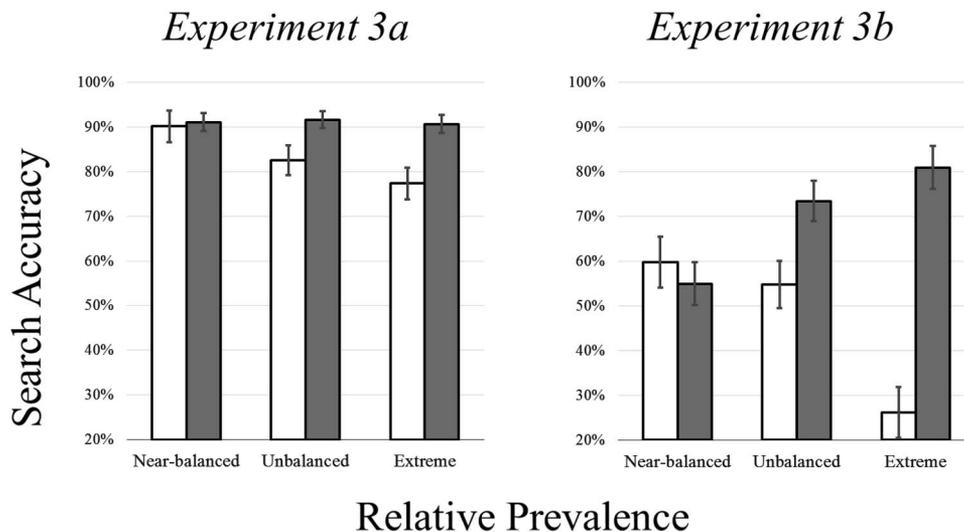


Figure 4. Experimental target-present search accuracy from Experiments 3a and 3b. Results are plotted separately for each between-participants condition of relative prevalence. Light-gray and dark-gray bars show results for the low-prevalence and high-prevalence items, respectively. Error bars represent ± 1 standard error of the mean.

Table 4
Mean Visual Search Accuracies (in Percentages) in Experiments 3a and 3b in Each Between-Participants Condition of Relative Prevalence

Relative prevalence	Trial type	Experiment 3a				Experiment 3b			
		Baseline	Experimental			Baseline	Experimental		
			Block 1	Block 2	Block 3		Block 1	Block 2	Block 3
Near balanced	Target absent	94	94	94	95	81	87	89	84
	Low prevalence	86	90	87	93	59	53	62	64
	High prevalence	87	90	92	92	52	53	55	56
Unbalanced	Target absent	92	95	93	95	82	84	85	90
	Low prevalence	86	83	82	82	58	53	60	52
	High prevalence	87	91	93	91	55	67	72	82
Extreme	Target absent	92	95	96	95	79	90	92	95
	Low prevalence	84	88	75	69	51	33	26	20
	High prevalence	87	89	90	93	62	70	86	87

because observers direct attention to them more efficiently or because, once attended, high-prevalence targets are more rapidly appreciated as being targets. Specifically, we deconstructed search behavior into two functionally separate phases: *scanning*, defined as the time from search initiation to target fixation, and *decision making*, defined as the time from first target fixation to the overt button response (see Malcolm & Henderson, 2009, 2010).

We used two dependent measures to characterize search behavior during scanning and decision making. *Scan-path ratios* (SPRs) were obtained by summing the amplitude of all saccades (in degrees of visual angle) prior to target fixation and dividing that value by the shortest distance from central fixation directly to the target. Thus, an SPR of 1 indicates that the participant's eye moved straight from central fixation to the target without deviating. Ratios greater than 1 indicate imperfect attentional guidance (i.e., that other locations were visited prior to the target being fixated). *Decision times* (DTs) were obtained by calculating the time between initial target fixation and the space bar press indicating identification. Finally, and most relevant to potential mechanisms underlying the LPE, we examined the likelihood that targets would be missed despite being directly examined. Our first prediction for this experiment was that we would replicate Experiment 1, observing prevalence effects in both accuracy and RT. Second, with respect to the eye-movement data, we predicted that similar prevalence effects and interactions would arise in SPRs and DTs, wherein SPRs would be longer to low-prevalence targets (indicating weakened attentional guidance), and DTs would be longer to low-prevalence targets (indicating poorer perceptual decision making) relative to high-prevalence targets. Third, we predicted that when low-prevalence targets were directly examined, they would be more likely to be missed relative to high-prevalence targets.

Method

The materials and design were identical to those in Experiment 1, using the standard (spatial) search procedure. The only change, relative to Experiment 1, was that we collected eye-movement data in addition to accuracy and RT data.

Participants. Twenty-nine and 31 new students from Arizona State University participated in Experiments 4a and 4b, respectively, in partial fulfillment of a course requirement. All partici-

pants reported normal or corrected-to-normal vision and normal color vision.

Apparatus. Data were collected using a Dell Optiplex 755 PC (2.66 GHz, 3.25 GB RAM). The display was a 21-in. NEC (Itasca, IL) FE21111 CRT monitor, with resolution set to 1,280 × 1,024 and refresh rate of 60 Hz. E-Prime Version 2.0 software was used to control stimulus presentation and collect responses. Eye movements were recorded using an Eyelink 1000 eye tracker (SR Research Ltd., Mississauga, Ontario, Canada) mounted on the desktop. Temporal resolution was 1000 Hz, and spatial resolution was 0.01°. An eye movement was classified as a saccade when its distance exceeded 0.5° and its velocity reached 30°/s (or acceleration reached 8,000°/s²). Viewing was binocular, but only the left eye was recorded.

Procedure. The procedure was identical to that in Experiment 1, with the exception of details pertaining to eye tracking. Participants used a chin rest during all trials and were initially calibrated to ensure accurate tracking. The chin rest was adjusted so each participant's gaze landed centrally on the computer screen when the participant looked straight ahead. The calibration procedure establishes a map of the participant's known gaze position relative to the tracker's coordinate estimate of that position. The routine proceeded by having participants fixate a black circle as it moved to nine different positions (randomly) on the screen. Calibration was accepted if the mean error was less than 0.5° of visual angle, with no error exceeding 1.0° of visual angle. Periodic drift correction and recalibrations ensured accurate recording of gaze position throughout the experiment. Interest areas were defined as the smallest rectangular area that encompassed any given image.

The trial procedure was modified to include a gaze-contingent fixation cross. When the fixation cross appeared, participants had to maintain gaze on it for 500 ms, which triggered the search display to appear. In (rare) circumstances wherein this did not occur within 10 s, because of human error or calibration problems, the trial was discarded, and recalibration was performed before the next trial.

Results

The results of principal interest (accuracy and RTs for target-present trials) are shown in Figures 5 and 6; both figures reveal

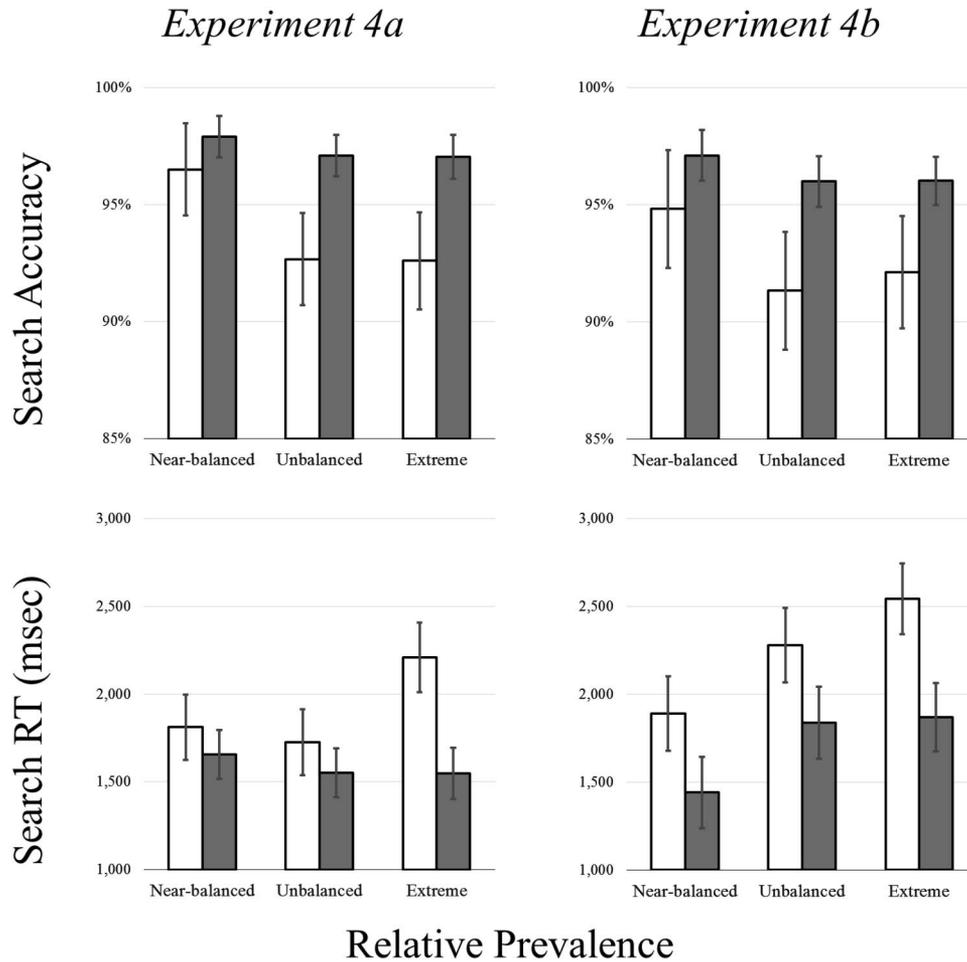


Figure 5. Experimental target-present search accuracy and reaction time (RT) in Experiments 4a and 4b. Results are plotted separately for each between-participants condition of relative prevalence. Light-gray and dark-gray bars show results for the low-prevalence and high-prevalence items, respectively. Error bars represent ± 1 standard error of the mean.

clear LPEs in both Experiments 4a and 4b. The data were analyzed in identical fashion to the data in Experiment 1, although we now included two new dependent measures for eye tracking: SPRs and DTs. RTs, SPRs, and DTs were only analyzed from trials with correct responses. Because people can occasionally detect target images without direct fixation, SPRs and DTs were not analyzed for trials in which targets were not directly fixated (this occurred in only a small percentage of trials: 3% and 6% for Experiments 4a and 4b, respectively). We also performed a nonparametric chi-square test for independence of observations; specifically, we were interested in determining whether low-prevalence targets were more likely to be fixated, and yet still missed, relative to high-prevalence targets (see Tables 5, 6, 7, and 8 for accuracy, RT, SPR, and DT means, respectively).

Accuracy. In Experiment 4a, overall accuracy was 96%. There were no main effects of relative prevalence group or block ($F_s < 2.00$). There was a main effect of target prevalence, $F(1, 26) = 8.82, p < .01, \eta_p^2 = .25$, with more hits to high-prevalence categories (97%) than to low-prevalence categories (94%). No interactions were reliable ($F_s < 2.00$).

In Experiment 4b, overall accuracy was 95%. There were no main effects of relative prevalence group or block ($F_s < 1.00$). There was a main effect of target prevalence, $F(1, 28) = 5.99, p < .05, \eta_p^2 = .18$, with more hits to high-prevalence categories (96%) than to low-prevalence categories (93%). No interactions were reliable ($F_s < 2.00$).

RTs. In Experiment 4a, there was no effect of relative prevalence group ($F < 1.00$). There was a main effect of target prevalence, $F(1, 26) = 13.29, p < .01, \eta_p^2 = .34$, with shorter RTs to the high-prevalence category (1,585 ms) relative to the low-prevalence category (1,915 ms). There was an effect of block, $F(2, 25) = 6.32, p < .01, \eta_p^2 = .34$, indicating shorter RTs over time (1,882, 1,780, and 1,588 ms for Blocks 1–3, respectively). The Relative Prevalence \times Target Prevalence interaction was marginally significant, $F(2, 26) = 3.25, p = .06, \eta_p^2 = .20$, suggesting that the prevalence effect increased in the less balanced groups (see Figure 5). No other interactions were reliable ($F_s < 1.00$).

In Experiment 4b, there was no effect of relative prevalence group ($F < 2.00$). There was a main effect of target prevalence, $F(1, 28) = 24.52, p < .01, \eta_p^2 = .47$, with shorter RTs to the high-prevalence

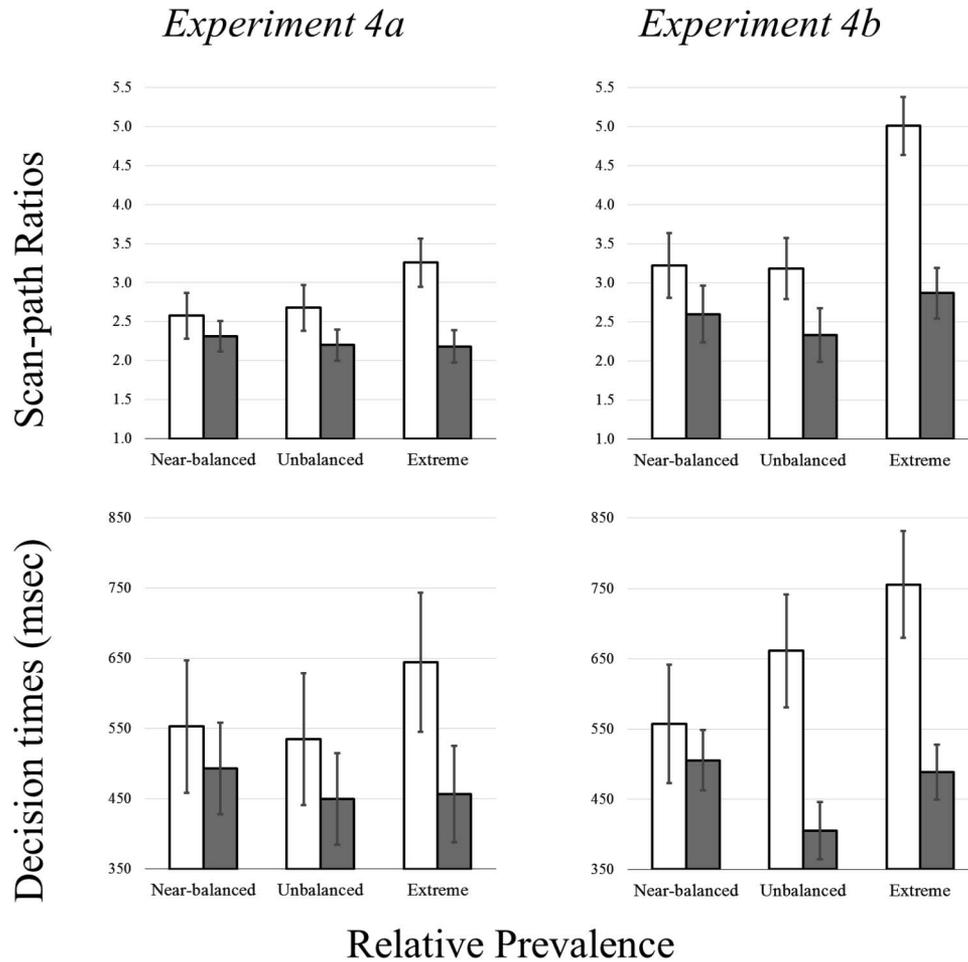


Figure 6. Experimental scan-path ratios and decision times in Experiments 4a and 4b. Results are plotted separately for each between-participants condition of relative prevalence. Light-gray and dark-gray bars show results for the low-prevalence and high-prevalence items, respectively. Error bars represent ± 1 standard error of the mean.

category (1,717 ms) relative to the low-prevalence category (2,239 ms). There was an effect of block, $F(2, 27) = 18.42, p < .01, \eta_p^2 = .58$, indicating shorter RTs over time (2,337, 1,879, and 1,717 ms for Blocks 1–3, respectively). No interactions were reliable ($F_s < 2.00$).

SPRs. In both Experiments 4a and 4b, the eye-tracking measures were of primary interest. As shown in Figure 6, there were clear prevalence effects in both the SPR and DT measures. In Experiment 4a, there was no main effect of relative prevalence group ($F < 1.00$), but we observed a robust effect of target prevalence, $F(1, 26) = 17.36, p < .01, \eta_p^2 = .40$, with more optimal SPRs to high-prevalence targets (2.23) relative to low-prevalence targets (2.84). There was a main effect of block, $F(2, 25) = 6.30, p < .01, \eta_p^2 = .34$, indicating that search got more efficient over time (2.68, 2.63, and 2.29 for Blocks 1–3, respectively). None of the interactions were reliable ($F_s < 3.00$).

In Experiment 4b, there was a main effect of relative prevalence group, $F(2, 27) = 4.22, p < .05, \eta_p^2 = .24$, with less optimal SPRs in the extreme group relative to more balanced groups (2.91, 2.76, and 3.94 for near balanced, unbalanced, and extreme, respec-

tively). There was a main effect of target prevalence, $F(1, 27) = 33.46, p < .01, \eta_p^2 = .55$, with more optimal SPRs to high-prevalence targets (2.60) relative to low-prevalence targets (3.81). There was an effect of block, $F(2, 26) = 6.61, p < .01, \eta_p^2 = .34$, indicating that search got more efficient over time (3.56, 3.02, and 3.04 for Blocks 1–3, respectively). The Relative Prevalence Group \times Target Prevalence interaction was reliable, $F(2, 27) = 5.35, p < .05, \eta_p^2 = .28$, indicating that the prevalence effect increased in the less balanced groups (see Figure 6). The Target Prevalence \times Block interaction was also reliable, $F(2, 26) = 5.55, p < .05, \eta_p^2 = .30$, indicating that observers got better over time at finding high-prevalence targets but not low-prevalence targets (see Table 7). No other interactions were reliable ($F_s < 2.00$).

DTs. In Experiment 4a, there was no effect of relative prevalence group ($F < 1.00$), but we did observe a main effect of target prevalence, $F(1, 26) = 4.31, p < .05, \eta_p^2 = .14$, with shorter DTs to the high-prevalence category (466 ms) relative to the low-prevalence category (577 ms). There was a block effect, $F(2, 25) = 7.43, p < .01, \eta_p^2 = .37$, although no clear pattern was

Table 5

Mean Visual Search Accuracies (in Percentages) in Experiments 4a and 4b in Each Between-Participants Condition of Relative Prevalence

Relative prevalence	Trial type	Experiment 4a				Experiment 4b			
		Baseline	Experimental			Baseline	Experimental		
			Block 1	Block 2	Block 3		Block 1	Block 2	Block 3
Near balanced	Target absent	100	99	99	99	99	100	100	100
	Low prevalence	97	96	98	96	98	94	96	95
	High prevalence	96	97	99	98	97	97	96	98
Unbalanced	Target absent	99	97	100	98	99	100	99	100
	Low prevalence	86	94	90	94	92	91	93	90
	High prevalence	95	98	97	97	86	96	95	97
Extreme	Target absent	99	97	99	98	98	100	100	99
	Low prevalence	98	96	87	96	95	95	89	93
	High prevalence	98	97	97	97	92	95	97	96

observed (444, 568, and 553 ms for Blocks 1–3, respectively). No interactions were reliable ($F_s < 2.00$).

In Experiment 4b, there was no effect of relative prevalence group ($F < 2.0$), but there was a large effect of target prevalence, $F(1, 27) = 22.28, p < .01, \eta_p^2 = .45$, with shorter DTs to high-prevalence targets (466 ms) relative to low-prevalence targets (658 ms). There was an effect of block, $F(2, 26) = 11.78, p < .01, \eta_p^2 = .48$, indicating that DTs got shorter over time (723, 508, and 455 ms for Blocks 1–3, respectively). None of the interactions were reliable ($F_s < 3.00$).

Perceptual failures. In Experiment 4a, participants missed a total of 53 low-prevalence targets. Among these misses, 18 (34%) of the targets were directly fixated. In comparison, 90 high-prevalence targets were missed, but only 11 (12%) were directly fixated. Thus, when rare targets were fixated, they were more likely to be missed relative to common targets, $\chi^2(1, N = 29) = 9.75, p < .01$. This pattern was not observed in Experiment 4b, which entailed more challenging perceptual decisions. Participants in Experiment 4b missed 67 low-prevalence targets and directly fixated 11 (16%) of them. Collectively, they also missed 132 high-prevalence targets, of which 17 (13%) were directly fixated, $\chi^2(1, N = 31) = 0.46, ns$.

Discussion

The findings from Experiment 4 can be quickly summarized. First, we replicated the results from Experiment 1, finding

consistent prevalence effects in accuracy and RTs. Second, by tracking eye movements, we discovered that the RT benefit for high-prevalence targets was a result of more efficient scanning and faster decision making. Simply put, relative to low-prevalence targets, high-prevalence items were both found and appreciated more quickly (see Hon & Tan, 2013). And, finally, in Experiment 4a, even when rare targets were directly fixated, they were still more likely to be missed relative to common targets (for similar results in different paradigms, see Godwin, Menneer, Riggs, Cave, & Donnelly, 2015; Solman, Cheyne, & Smilek, 2012). In Experiment 4b, this effect was not seen, most likely because, in this more challenging search, all targets (of both high and low prevalence) required fairly close scrutiny. In contrast, in the easier search condition (Experiment 4a), participants likely learned to rely on common targets popping out with fairly superficial examination, which would have a detrimental effect on appreciating rare targets.

General Discussion

The present experiments were intended to capture aspects of the challenges facing professional visual searchers in settings such as airport security and radiology. In tasks such as these, observers need to build perceptual expertise, which is achieved by amassing great experience examining X-rays, for example, and learning how

Table 6

Mean Visual Search Reaction Times (in Milliseconds) in Experiments 4a and 4b in Each Between-Participants Condition of Relative Prevalence

Relative prevalence	Trial type	Experiment 4a				Experiment 4b			
		Baseline	Experimental			Baseline	Experimental		
			Block 1	Block 2	Block 3		Block 1	Block 2	Block 3
Near balanced	Target absent	6,042	5,487	4,750	4,560	6,993	5,273	3,925	3,495
	Low prevalence	2,235	1,942	1,834	1,657	2,811	2,182	1,845	1,647
	High prevalence	2,146	1,860	1,684	1,424	3,339	1,876	1,330	1,120
Unbalanced	Target absent	6,541	5,079	4,701	3,835	6,727	5,665	4,260	3,854
	Low prevalence	2,377	1,751	1,878	1,545	2,794	2,649	2,197	1,996
	High prevalence	2,437	1,633	1,661	1,364	3,271	2,433	1,556	1,530
Extreme	Target absent	7,231	6,678	5,864	5,662	6,862	5,987	5,151	4,504
	Low prevalence	2,259	2,306	2,143	2,175	2,664	2,632	2,537	2,462
	High prevalence	2,437	1,803	1,479	1,362	3,243	2,252	1,811	1,548

Table 7

Mean Scan-Path Ratios in Experiments 4a and 4b in Each Between-Participants Condition of Relative Prevalence

Relative prevalence	Trial type	Experiment 4a				Experiment 4b			
		Baseline	Experimental			Baseline	Experimental		
			Block 1	Block 2	Block 3		Block 1	Block 2	Block 3
Near balanced	Low prevalence	3.06	2.60	2.68	2.45	3.48	3.67	2.92	3.08
	High prevalence	3.26	2.71	2.43	1.80	4.57	3.45	2.22	2.13
Unbalanced	Low prevalence	3.37	2.63	2.86	2.54	4.13	3.24	3.22	3.11
	High prevalence	3.23	2.30	2.36	1.94	4.43	3.14	2.04	1.81
Extreme	Low prevalence	3.09	3.32	3.42	3.02	4.75	4.22	5.06	5.75
	High prevalence	3.58	2.53	2.05	1.97	6.22	3.61	2.64	2.36

to spot important anomalies. As expertise grows, however, it comes with an ironic side effect, that of implicitly teaching observers what to anticipate seeing even before they look at a display. Such learning is typically accurate (by definition) and allows search to proceed rapidly, as would be necessary for a baggage screener with a line of travelers or a radiologist with a heavy caseload. The price for such efficiency, however, is an implicit bias against noticing rare targets. Yet, in these contexts, missed targets may have dire consequences. Moreover, professional searchers must typically be vigilant for multiple potential targets. Radiologists may look for multiple types of abnormalities, and baggage screeners must identify any potential weapon in addition to miscellaneous prohibited items, such as water bottles or lighters. Even under conditions without prevalence asymmetries, human observers struggle when searching for multiple potential targets (e.g., Godwin, Menneer, Cave, & Donnelly, 2010; Menneer, Donnelly, Godwin, & Cave, 2010; Hout & Goldinger, 2010, 2012).

In many natural situations, looking for multiple potential targets is complicated by their differing rates of occurrence. As noted in the introduction, rare targets are more likely to be missed, and screeners typically have little more than personal experience (or intuition) to guide their expectations for different targets. In their Experiments 3 and 4, Wolfe et al. (2007) found robust prevalence effects in situations in which participants searched for three or four targets with different rates of occurrence, ranging from 1% to 44%. A recent study by Mitroff and Biggs (2014) extended this result to much lower prevalence, examining millions of trials from the Airport Scanner smartphone application (Kedlin Company, Bellevue, WA [http://www.airportscannergame.com]), which allowed them to investigate detection of what they termed *ultrarare items*.

Thirty of their potential targets were present in fewer than 0.15% trials, with the rarest item present in just 0.078% of trials (rates on par with the prevalence of cancer rates in mammography; Breast Cancer Surveillance Consortium, 2009). Mitroff and Biggs found that the relationship between target frequency and target detection was logarithmic, leading to disastrously low performance with targets that seldom appeared.

Visual search can be conceptualized as a two-stage process. On each new allocation of attention, the searcher asks, "Is this the target?" And, if not, the next question is, "Should I continue searching?" It is reasonable to expect that, in many cases, searchers will miss rare targets because of Stage 2 failures, missing items merely because they stopped searching prior to placing attention on them (owing to time pressure, reaching an internal quitting threshold, or both). In the current investigation, we examined whether low-prevalence misses might also occur as result of Stage 1 perceptual errors by using passive RSVP procedures and tracking participants' eye movements in standard oculomotor search. Our findings strongly suggest that low-prevalence misses sometimes reflect failures to appreciate targets once they have received attention.

Prevalence Effects in Multiple-Target Search

Routine visual searches often involve looking for more than one item at a time (sometimes many more; Boettcher, Drew, & Wolfe, 2013; Drew & Wolfe, 2014; Wolfe, 2012). When scanning the refrigerator for necessary dinner ingredients, it is not terribly difficult to search for broccoli, carrots, and butter simultaneously. Although it is possible to maintain more than one target representation at a time (Beck, Hollingworth, & Luck, 2012), multiple-

Table 8

Mean Decision Times (in Milliseconds) in Experiments 4a and 4b in Each Between-Participants Condition of Relative Prevalence

Relative prevalence	Trial type	Experiment 4a				Experiment 4b			
		Baseline	Experimental			Baseline	Experimental		
			Block 1	Block 2	Block 3		Block 1	Block 2	Block 3
Near balanced	Low prevalence	448	491	620	548	1140	746	522	403
	High prevalence	481	420	524	535	1230	654	479	383
Unbalanced	Low prevalence	691	483	574	548	1238	874	561	548
	High prevalence	762	389	520	440	1457	515	376	323
Extreme	Low prevalence	849	449	694	790	1221	929	667	670
	High prevalence	738	435	476	459	1483	620	441	404

target search incurs costs in both accuracy and RT (Hout & Goldinger, 2010) that arise because of decreased efficiency in both attentional guidance and perceptual decision making (Hout & Goldinger, 2012, 2014). For professional searchers, it is typically necessary to seek multiple targets simultaneously. Under some circumstances, it is more effective to conduct several consecutive single-target searches (Menneer, Barrett, Phillips, Donnelly, & Cave, 2007; Menneer, Cave, & Donnelly, 2009; Menneer, Phillips, Donnelly, Barrett, & Cave, 2004; Stroud, Menneer, Cave, & Donnelly, 2012); however, it is unclear whether professionals do this when the number of possible targets is small, and it is unlikely to be practical when many target types are possible.

Godwin and colleagues investigated the interaction between prevalence rates and the costs of multiple-target search (Godwin, Menneer, Cave, & Donnelly, 2010; Menneer et al., 2010) using a dual-target relative prevalence manipulation, as in the current study. Their participants searched through X-rays of threatening and nonthreatening items, looking for metal threats (e.g., guns, knives) and explosive devices. In some blocks, observers looked for one target category or the other; in others, they looked for both categories concurrently. The relative prevalence of the categories was varied, occurring either with equal frequency or with one category nine times more likely than its counterpart. Godwin, Menneer, Cave, and Donnelly (2010) found a dual-target cost in performance, indexed by higher error rates, longer RTs, and decreased sensitivity (d') relative to single-target baselines (see also Menneer et al., 2004, 2007). Moreover, high-prevalence targets were detected more quickly and accurately than low-prevalence targets.

Our current findings extend Godwin, Menneer, Cave, Helman, et al.'s (2010) results, answering a question they posed in their General Discussion: "When searching for targets of varied prevalence, do participants show some form of preferential guidance for the higher-prevalence target?" (p. 84). Our eye-movement results offer an affirmative answer to this query, but they also show that, once guidance is complete, perceptual appreciation of targets is also modulated by their relative prevalence. Once an observer has learned not to expect certain targets, it becomes more difficult for those targets to attract attention and more difficult for them to resonate with search intentions held in memory. Indeed, our findings are also in alignment with a more recent investigation by Godwin and colleagues (Godwin, Menneer, Riggs, et al., 2015; see also Godwin, Menneer, Cave, et al., 2015); their participants searched for two rotated T s (that differed considerably in color) among rotated L s (of assorted colors). As in our experiments, one target color was present in 45% of trials and its counterpart in 5% of trials. The experimenters tracked eye movements, finding that prevalence rates influenced failures of attentional selection (i.e., low-prevalence items tended to be fixated more slowly and less frequently than high-prevalence targets) and failures of perceptual identification (i.e., low-prevalence items that were fixated were missed more often than high-prevalence items that had been examined). More important, in the second experiment (wherein the targets were of very similar colors), it was found that even when perceptual selection was equal across low- and high-prevalence targets, perceptual identification errors were more likely to occur for the low-prevalence item.

Mental Representation of Rare Targets

In Experiment 4, we found that high-prevalence targets were initially located more quickly (indexed by SPRs) and their identities confirmed more rapidly (indexed by DTs) relative to low-prevalence targets. Why might this be the case? One hypothesis is that such benefits reflect "better" mental representations for targets that are more frequently encountered. In a recent study examining search-template precision, we (Hout & Goldinger, 2015) found a similar pattern of findings to those reported here. We examined how the precision of people's target templates (i.e., their working memory representations of targets; Bravo & Farid, 2009, 2012; Malcolm & Henderson, 2009; Tinbergen, 1960; Wolfe, Horowitz, Kenner, Hyle, & Vasan, 2004) affects their ability to search effectively. We did this in two ways—first by corrupting searchers' templates with erroneous features and, second, by introducing superfluous features to the template that were unhelpful. Participants' eye movements revealed the same patterns for poor versus good templates as we currently observed for low-versus high-prevalence targets. Specifically, better mental representations led to more efficient attentional guidance and faster perceptual decisions. It therefore seems possible that the benefits observed for high-prevalence targets may reflect searchers having more effective template representations for them.

One reason that target templates may be enhanced for high-prevalence items is the self-reinforcing nature of finding those items more often. Each time an observer finds a target, he or she is given another opportunity to perceptually encode its features, which will promote perceptual priming for that same target in subsequent trials (Theeuwes, 2013). Such template tuning, however, cannot provide a complete account. Wolfe et al. (2007) made several attempts to overcome the low-prevalence effect. In one of their experiments, participants searched for 2,000 trials, looking for targets with 1% prevalence. Interspersed within those 2,000 trials were six 60-trial bursts of high target prevalence. Prevalence rose in each burst from 1% to 50%, and then it returned to 1% again, and participants were given feedback after every trial. If simply seeing targets more often was enough to strengthen the target template, then searchers in this experiment should have overcome the LPE, but they did not: Miss rates averaged 41%, consistent with other low-prevalence findings using the same stimuli. Only in a modified task, when feedback was provided only during the high-prevalence bursts, was the LPE overcome (and then only temporarily).

Focusing on the current results, in half of our experiments, participants saw categorical (verbal) cues and, thus, could not create perceptually detailed target templates. The other half of our experiments involved having participants search for specific exemplars, shown before each trial. If having a clear template was the key to overcoming low target prevalence, we should have observed minimal LPEs in the experiments with image cues, but this did not occur: The LPE was robust, despite (presumably) perfect search templates.

Considering the eye-movement data, an alternative explanation for the LPE is that participants gradually learned that one target occurred more often and decided (explicitly or implicitly) to search for that item first and then perform a second scan for the lower prevalence target if needed. This seems an unlikely explanation for our results, for several reasons. First, if participants only found rare

targets on a second search through a display, we might expect SPRs to be on the order of double those for high-prevalence targets. Although Figure 6 shows that low-prevalence SPRs were longer than high-prevalence SPRs, they were not two times longer, suggesting that participants searched for both targets at once. Note that this doubling hypothesis assumes that there is no carryover of information from the first search (for the high-prevalence target) to the second. Suppose that there was some carryover, such that participants could avoid some items on the second search. This would imply that they had determined that those items were not low-prevalence targets on the first pass, thus indicating that they had searched for both items during the first pass. The SPR data do not illuminate whether participants could search simultaneously for both targets or if they rapidly switched back and forth between target types during search. Wolfe's (2012) data on search for up to 100 target items at once argues against a sequential-switching account.

Taken together, the current findings suggest that high-prevalence targets receive privileged status in a searcher's mind. Considering the results from Experiment 4, another potential account regards the strength of top-down matching, an idea that is closely related to our hypothesis about template precision. Per this hypothesis, an observer may have equally robust templates in visual working memory for rare and common targets (especially when search is immediately preceded by a target image), but one of the targets receives greater implicit weight as the searcher scans the display (see Vö & Wolfe, 2012). If the more likely target were given an attentional boost, it would take priority in seeking resonance with objects in the display (i.e., attentional guidance) and would be more readily available to resonate with fixated items (i.e., perceptual decisions). When an image is foremost in mind, it becomes salient in the environment, as seen in numerous anxiety disorders that are characterized by excessive perceptual vigilance to threat signals (e.g., Fox, 1993; MacLeod, Mathews, & Tata, 1986).

An implicit, prevalence-based weighting mechanism would naturally explain our findings in standard and RSVP search. Moreover, it would explain the most curious finding in our results. Specifically, in Experiment 4a, participants searched for broad categories (butterflies or teddy bears) among other random objects. In this experiment, when people directly fixated rare targets, they often failed to perceptually appreciate them as being targets, missing them three times more often than similarly fixated common targets. Per a weighting hypothesis, despite looking directly at the image, a participant prioritizing a likely target is simply not prepared to appreciate a rare target and, thus, the target fails to achieve perceptual resonance. In contrast to this result, in Experiment 4b, when participants had to search for specific objects among similar distractors (e.g., a butterfly among other butterflies), we no longer found any difference in miss rates for rare or common items that were directly attended. Participants occasionally looked at target items and still failed to detect them, but this did not vary as a function of target prevalence. We suggest that, in this condition, observers cannot rely on categorical matching to perform search. Because the targets and distractors are always visually similar, observers should optimally adopt a more laborious search strategy, relying on specific features to identify targets. (Note that search RTs were far longer in Experiment 4b relative to Experiment 4a.) In this case, every item should receive closer

perceptual examination, which would allow even unexpected targets a reasonable chance of resonating with templates in memory.

It is also worth noting that such an explanation would naturally mesh well with LPE findings during single-target search. We suggest that when observers search for a single rare target, their mental representation is less pronounced (relative to search for a single, more prevalent target) not because they have learned to favor the representation of one category over another. Rather, the low regularity with which rare targets are located makes their attentional status in visual working memory less effective, leaving the observer more susceptible to low-prevalence errors.

Conclusion

The LPE is a persistent problem that has important societal implications (e.g., Mitroff & Biggs, 2014). There are at least three contributing factors to the LPE: (a) perceptual errors, (b) early search termination, and (c) motor errors. By controlling for the latter two factors, we were able to show clear evidence for the first—a perceptual basis for the LPE. Eye-tracking analyses showed that common targets were found more efficiently and identified more quickly than rare targets. Our findings strongly align with recent work by Godwin and colleagues (Godwin, Menneer, Cave, et al., 2015; Godwin, Menneer, Riggs, et al., 2015) and extend their work in two important ways. First, our RSVP experiments (Experiments 3 and 4) demonstrate that LPEs arise even when early search termination errors are not possible and when reflexive motor errors are extremely unlikely. Second, the LPEs shown across picture- and word-cued search suggest that prevalence rates influence search behavior not only when templates are precise (i.e., when participants are given picture cues) but also when the specific features of the target are somewhat unpredictable (as in categorical, word-cued search). Taking these findings together, we conclude that perceptual identification errors are a contributing factor to the LPE and suggest that searchers maintain superior and more available mental representations for more likely items in visual search.

References

- Beck, V. M., Hollingworth, A., & Luck, S. J. (2012). Simultaneous control of attention by multiple working memory representations. *Psychological Science*, *23*, 887–898. <http://dx.doi.org/10.1177/0956797612439068>
- Biernaskie, J. M., Walker, S. C., & Gegear, R. J. (2009). Bumblebees learn to forage like Bayesians. *American Naturalist*, *174*, 413–423. <http://dx.doi.org/10.1086/603629>
- Boettcher, S. E. P., Drew, T., & Wolfe, J. M. (2013). Hybrid search in context: How to search for vegetables in the produce section and cereal in the cereal aisle. *Visual Cognition*, *21*, 678–682. <http://dx.doi.org/10.1080/13506285.2013.844959>
- Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences of the United States of America*, *105*, 14325–14329. <http://dx.doi.org/10.1073/pnas.0803390105>
- Bravo, M. J., & Farid, H. (2009). The specificity of the search template. *Journal of Vision*, *9*(1), 34, 1–9. <http://dx.doi.org/10.1167/9.1.34>
- Bravo, M. J., & Farid, H. (2012). Task demands determine the specificity of the search template. *Attention, Perception, & Psychophysics*, *74*, 124–131. <http://dx.doi.org/10.3758/s13414-011-0224-5>
- Breast Cancer Surveillance Consortium. (2009). *Cancer rate (per 1,000 examinations) and cancer detection rate (per 1,000 examinations) for*

- 1,960,150 screening mammography examinations from 2002 to 2006 by age. Retrieved from http://breastscreening.cancer.gov/data/performance/screening/2009/rate_age.html
- Cain, M. S., Vul, E., Clark, K., & Mitroff, S. R. (2012). A Bayesian optimal foraging model of human visual search. *Psychological Science*, 23, 1047–1054. <http://dx.doi.org/10.1177/0956797612440460>
- Charnov, E. L. (1976). Optimal foraging, the marginal value theorem. *Theoretical Population Biology*, 9, 129–136. [http://dx.doi.org/10.1016/0040-5809\(76\)90040-X](http://dx.doi.org/10.1016/0040-5809(76)90040-X)
- Chun, M. M., & Wolfe, J. M. (1996). Just say no: How are visual searches terminated when there is no target present? *Cognitive Psychology*, 30, 39–78. <http://dx.doi.org/10.1006/cogp.1996.0002>
- Drew, T., & Wolfe, J. M. (2014). Hybrid search in the temporal domain: Evidence for rapid, serial logarithmic search through memory. *Attention, Perception, & Psychophysics*, 76, 296–303. <http://dx.doi.org/10.3758/s13414-013-0606-y>
- Evans, K. K., Birdwell, R. L., & Wolfe, J. M. (2013). If you don't find it often, you often don't find it: Why some cancers are missed in breast cancer screening. *PLoS ONE*, 8, e64366. <http://dx.doi.org/10.1371/journal.pone.0064366>
- Evans, K. K., Tambouret, R. H., Evered, A., Wilbur, D. C., & Wolfe, J. M. (2011). Prevalence of abnormalities influences cytologists' error rates in screening for cervical cancer. *Archives of Pathology & Laboratory Medicine*, 135, 1557–1560. <http://dx.doi.org/10.5858/arpa.2010-0739-OA>
- Fleck, M. S., & Mitroff, S. R. (2007). Rare targets are rarely missed in correctable search. *Psychological Science*, 18, 943–947. <http://dx.doi.org/10.1111/j.1467-9280.2007.02006.x>
- Fox, E. (1993). Allocation of visual attention and anxiety. *Cognition and Emotion*, 7, 207–215. <http://dx.doi.org/10.1080/02699939308409185>
- Godwin, H. J., Menneer, T., Cave, K. R., Helman, S., Way, R. L., & Donnelly, N. (2010). The impact of relative prevalence on dual-target search for threat items from airport X-ray screening. *Acta Psychologica*, 134, 79–84. <http://dx.doi.org/10.1016/j.actpsy.2009.12.009>
- Godwin, H. J., Menneer, T., & Cave, K. R., Thaisyah, M., & Donnelly, N. (2015). The effects of increasing target prevalence on information processing during visual search. *Psychonomic Bulletin & Review*, 22, 469–475.
- Godwin, H. J., Menneer, T., Cave, K. R., & Donnelly, N. (2010). Dual-target search for high and low prevalence X-ray threat targets. *Visual Cognition*, 18, 1439–1463. <http://dx.doi.org/10.1080/13506285.2010.500605>
- Godwin, H. J., Menneer, T., Riggs, C. A., Cave, K. R., & Donnelly, N. (2015). Perceptual failures in the selection and identification of low-prevalence targets in relative prevalence visual search. *Attention, Perception, & Psychophysics*, 77, 150–159.
- Godwin, H. J., Walenchok, S., Houpt, J. W., Hout, M. C., & Goldinger, S. D. (in press). Faster than the speed of rejection: Object identification processes during visual search for multiple targets. *Journal of Experimental Psychology: Human Perception and Performance*.
- Hon, N., & Tan, C.-H. (2013). Why rare targets are slow: Evidence that the target probability effect has an attentional locus. *Attention, Perception, & Psychophysics*, 75, 388–393. <http://dx.doi.org/10.3758/s13414-013-0434-0>
- Hout, M. C., & Goldinger, S. D. (2010). Learning in repeated visual search. *Attention, Perception, & Psychophysics*, 72, 1267–1282. <http://dx.doi.org/10.3758/APP.72.5.1267>
- Hout, M. C., & Goldinger, S. D. (2012). Incidental learning speeds visual search by lowering response thresholds, not by improving efficiency: Evidence from eye movements. *Journal of Experimental Psychology: Human Perception and Performance*, 38, 90–112. <http://dx.doi.org/10.1037/a0023894>
- 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, 128–149.
- Hout, M. C., Goldinger, S. D., & Brady, K. J. (2014). MM-MDS: A multidimensional scaling database with similarity ratings for 240 object categories from the *Massive Memory* picture database. *PLoS ONE*, 9, e112644. <http://dx.doi.org/10.1371/journal.pone.0112644>
- Ishibashi, K., Kita, S., & Wolfe, J. M. (2012). The effects of local prevalence and explicit expectations on search termination times. *Attention, Perception, & Psychophysics*, 74, 115–123. <http://dx.doi.org/10.3758/s13414-011-0225-4>
- 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, 558–578. <http://dx.doi.org/10.1037/a0019165>
- Kunar, M. A., Rich, A. N., & Wolfe, J. M. (2010). Spatial and temporal separation fails to counteract the effects of low prevalence in visual search. *Visual Cognition*, 18, 881–897. <http://dx.doi.org/10.1080/13506280903361988>
- Mackworth, J. (1970). *Vigilance and attention*. Harmondsworth, England: Penguin.
- MacLeod, C., Mathews, A., & Tata, P. (1986). Attentional bias in emotional disorders. *Journal of Abnormal Psychology*, 95, 15–20. <http://dx.doi.org/10.1037/0021-843X.95.1.15>
- Macmillan, N. A., & Creelman, C. D. (1990). Response bias: Characteristics of detection theory, threshold theory, and “nonparametric” indexes. *Psychological Bulletin*, 107, 401–413. <http://dx.doi.org/10.1037/0033-2909.107.3.401>
- Malcolm, G. L., & Henderson, J. M. (2009). The effects of target template specificity on visual search in real-world scenes: Evidence from eye movements. *Journal of Vision*, 9(11): 8, 1–13. <http://dx.doi.org/10.1167/9.11.8>
- Malcolm, G. L., & Henderson, J. M. (2010). Combining top-down processes to guide eye movements during real-world scene search. *Journal of Vision*, 10(2): 4, 1–11. <http://dx.doi.org/10.1167/10.2.4>
- Menneer, T., Barrett, D. J. K., Phillips, L., Donnelly, N., & Cave, K. R. (2007). Costs in searching for two targets: Dividing search across target types could improve airport security screening. *Applied Cognitive Psychology*, 21, 915–932. <http://dx.doi.org/10.1002/acp.1305>
- Menneer, T., Cave, K. R., & Donnelly, N. (2009). The cost of search for multiple targets: Effects of practice and target similarity. *Journal of Experimental Psychology: Applied*, 15, 125–139. <http://dx.doi.org/10.1037/a0015331>
- Menneer, T., Donnelly, N., Godwin, H. J., & Cave, K. R. (2010). High or low target prevalence increases the dual-target cost in visual search. *Journal of Experimental Psychology: Applied*, 16, 133–144. <http://dx.doi.org/10.1037/a0019569>
- Menneer, T., Phillips, L., Donnelly, N., Barrett, D. J. K., & Cave, K. R. (2004). Search efficiency for multiple targets. *Cognitive Technology*, 9, 22–25.
- Menneer, T., Stroud, M. J., Cave, K. R., Li, X., Godwin, H. J., Liversedge, S. P., & Donnelly, N. (2012). Search for two categories of target produces fewer fixations to target-color items. *Journal of Experimental Psychology: Applied*, 18, 404–418. <http://dx.doi.org/10.1037/a0031032>
- Mitroff, S. R., & Biggs, A. T. (2014). The ultra-rare-item effect: Visual search for exceedingly rare items is highly susceptible to error. *Psychological Science*, 25, 284–289. <http://dx.doi.org/10.1177/0956797613504221>
- Potter, M. C., Wyble, B., Haggmann, C. E., & McCourt, E. S. (2014). Detecting meaning in RSVP at 13 ms per picture. *Attention, Perception, & Psychophysics*, 76, 270–279. <http://dx.doi.org/10.3758/s13414-013-0605-z>
- Reed, W. M., Chow, S. L. C., Chew, L. E., & Brennan, P. C. (2014a). Assessing the impact of prevalence expectations on radiologists' behav-

- ior. *Academic Radiology*, 21, 1220–1221. <http://dx.doi.org/10.1016/j.acra.2014.06.001>
- Reed, W. M., Chow, S. L. C., Chew, L. E., & Brennan, P. C. (2014b). Can prevalence expectations drive radiologists' behavior? *Academic Radiology*, 21, 450–456. <http://dx.doi.org/10.1016/j.acra.2013.12.002>
- Reed, W. M., Ryan, J. T., McEntee, M. F., Evanoff, M. G., & Brennan, P. C. (2011). The effect of abnormality-prevalence expectation on expert observer performance and visual search. *Radiology*, 258, 938–943. doi:radiol.10101090
- Rich, A. N., Kunar, M. A., Van Wert, M. J., Hidalgo-Sotelo, B., Horowitz, T. S., & Wolfe, J. M. (2008). Why do we miss rare targets? Exploring the boundaries of the low prevalence effect. *Journal of Vision*, 8(15): 15, 1–17. <http://dx.doi.org/10.1167/8.15.15>
- Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime user's guide*. Pittsburgh, PA: Psychology Software Tools.
- Schwark, J. D., MacDonald, J., Sandry, J., & Dolgov, I. (2013). Prevalence-based decisions undermine visual search. *Visual Cognition*, 21, 541–568. <http://dx.doi.org/10.1080/13506285.2013.811135>
- Schwark, J., Sandry, J., & Dolgov, I. (2013). Evidence for a positive relationship between working-memory capacity and detection of low-prevalence targets in visual search. *Perception*, 42, 112–114. <http://dx.doi.org/10.1068/p7386>
- Schwark, J., Sandry, J., Macdonald, J., & Dolgov, I. (2012). False feedback increases detection of low-prevalence targets in visual search. *Attention, Perception, & Psychophysics*, 74, 1583–1589. <http://dx.doi.org/10.3758/s13414-012-0354-4>
- Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review*, 84, 127–190. <http://dx.doi.org/10.1037/0033-295X.84.2.127>
- Solman, G. J. F., Cheyne, J. A., & Smilek, D. (2012). Found and missed: Failing to recognize a search target despite moving it. *Cognition*, 123, 100–118. <http://dx.doi.org/10.1016/j.cognition.2011.12.006>
- Stroud, M. J., Menneer, T., Cave, K. R., & Donnelly, N. (2012). Using the dual-target cost to explore the nature of search target representations. *Journal of Experimental Psychology: Human Perception and Performance*, 38, 113–122. <http://dx.doi.org/10.1037/a0025887>
- Theeuwes, J. (2013). Feature-based attention: It is all bottom-up priming. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences*, 368, 20130055. <http://dx.doi.org/10.1098/rstb.2013.0055>
- Tinbergen, N. (1960). The natural control of insects in pinewoods. *Archives Neerlandaises de Zoologie*, 13, 265–343. <http://dx.doi.org/10.1163/036551660X00053>
- Võ, M. L.-H., & Wolfe, J. M. (2012). When does repeated search in scenes involve memory? Looking at versus looking for objects in scenes. *Journal of Experimental Psychology: Human Perception and Performance*, 38, 23–41. <http://dx.doi.org/10.1037/a0024147>
- Williams, C. C. (2010). Not all visual memories are created equal. *Visual Cognition*, 18, 201–228. <http://dx.doi.org/10.1080/13506280802664482>
- Wolfe, J. M. (2012). Saved by a log: How do humans perform hybrid visual and memory search? *Psychological Science*, 23, 698–703. <http://dx.doi.org/10.1177/0956797612443968>
- Wolfe, J. M., Brunelli, D. N., Rubinstein, J., & Horowitz, T. S. (2013). Prevalence effects in newly trained airport checkpoint screeners: Trained observers miss rare targets, too. *Journal of Vision*, 13(3): 33, 1–9. <http://dx.doi.org/10.1167/13.3.33>
- Wolfe, J. M., Horowitz, T. S., & Kenner, N. M. (2005, May 26). Cognitive psychology: Rare items often missed in visual searches. *Nature*, 435, 439–440. <http://dx.doi.org/10.1038/435439a>
- 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, 1411–1426. <http://dx.doi.org/10.1016/j.visres.2003.11.024>
- Wolfe, J. M., Horowitz, T. S., Van Wert, M. J., Kenner, N. M., Place, S. S., & Kibbi, N. (2007). Low target prevalence is a stubborn source of errors in visual search tasks. *Journal of Experimental Psychology: General*, 136, 623–638. <http://dx.doi.org/10.1037/0096-3445.136.4.623>
- Wolfe, J. M., & Van Wert, M. J. (2010). Varying target prevalence reveals two dissociable decision criteria in visual search. *Current Biology*, 20, 121–124. <http://dx.doi.org/10.1016/j.cub.2009.11.066>

Received December 15, 2014

Revision received February 27, 2015

Accepted February 28, 2015 ■