Published in final edited form as:

Nature. 2005 May 26; 435(7041): 439–440. doi:10.1038/435439a.

Rare targets are often missed in visual search

Jeremy M Wolfe^{1,2}, Todd S Horowitz^{1,2}, and Naomi M Kenner¹

¹Visual Attention Lab, Brigham and Women's Hospital, Boston, MA

²Dept. of Ophthalmology, Harvard Medical School, Boston, MA

Abstract

Our society relies on accurate performance of visual screening tasks (e.g. for knives in luggage or tumors in mammograms). These are visual searches for rare targets. We report that target rarity leads to disturbingly inaccurate performance.

Visual search is the subject of a voluminous laboratory literature ¹. Typically, observers perform several hundred searches and targets are presented on 50% of trials. Target prevalence in baggage screening or cancer screening is much lower (~0.3% in routine mammography ²). We compared performance on high and low prevalence versions of an artificial baggage-screening task. Observers looked for “tools” among objects drawn from other categories. Semi-transparent objects were presented on noisy backgrounds and could overlap (Fig 1). The number of objects in a display was 3, 6, 12, or 18. Target prevalence was 1%, 10% or 50%. At 1% prevalence, 12 paid volunteer observers had to be tested for 2000 trials each (broken into 250 trial blocks) to obtain a mere 20 target-present trials each. Each observer was tested for 200 trials in the 10% and 50% conditions. Observers were given feedback on their performance, including a point system designed to emphasize the importance of finding the target (see supplementary methods). Low prevalence search has some similarity to vigilance tasks where observers wait for fleeting signals ^{3,4}. However our search stimuli are continuously visible until observers choose to respond.

Figure 2a shows error rates as a function of number of objects. 50% prevalence produced 7% miss errors, typical for laboratory search tasks of this sort. However, errors increased dramatically (and reliably) as prevalence decreased. 10% prevalence produced 16% errors, while at 1% prevalence errors soared to 30%. Errors were primarily “misses” (failing to notice a target). “False alarms” (saying “yes” when targets are absent) were vanishingly rare (0.03%), despite incentives to produce the opposite behavior (see supplementary methods). Simply changing prevalence produced a fourfold increase in error rate. If similar effects occur in socially important searches, this could have significant consequences.

Why does this happen? The reaction time (RT) data (Fig 2b&c) provide some clues. Observers require a threshold for quitting when no target has been found. This threshold is constantly adjusted; observers slow down after mistakes and speed up after successes⁵.

64 Sidney St. Cambridge, MA 02139, wolfe@search.bwh.harvard.edu.

Supplementary information accompanies the communication on www.nature.com/nature.

When targets are frequent, fast “no” responses will often lead to mistakes. As a result, “no” RTs are slower than “yes” RTs in high prevalence search (2b). With infrequent targets, observers can successfully say “no” almost all of the time, driving down the quitting threshold. As seen in Fig 2c, the result is too rapid target-absent search. Observers abandon search in less than the average time required to find a target.

The problem cannot be solved simply by adding pseudotargets to increase prevalence (e.g. search baggage for iPods *and* weapons) In a second experiment, we mixed common (44% prevalence), rarer (10%), and very rare (1%) targets such that *some* target was present on 50% of trials. Here, observers missed just 11% of common targets but 25% of rarer targets and 52% of very rare targets (see supplementary methods & results). Is the prevalence effect just a by-product of naive Os unfamiliarity with the targets? In a separate investigation, we compared the miss error rate for 4000 trials at 1% prevalence (40 targets, 41% miss errors) to the miss error rate for the first 100 trials at 34% prevalence (34 targets, 11% miss errors). It appears to be prevalence, not just number of targets presented that is critical (See Supplementary Methods).

Visual search is a ubiquitous human signal detection task ⁶. Heuristics that produce acceptable performance over a wide range of target prevalence may betray us at low prevalence. Because the experiments are burdensome, we do not have a clear idea whether these effects occur in the field ^{7,8}. A scoring system in the lab cannot duplicate the motivation to find a gun or a tumor nor the motivation to move the check-in line along. The training of lab volunteers differs from that of professionals. Nevertheless, there are sufficient similarities between lab and field to strongly suggest that we should find out if the massive increases in error shown here occur in socially important search tasks.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

1. Wolfe, JM. Attention. Pashler, H., editor. Psychology Press Ltd; Hove, East Sussex, UK: 1998. p. 13-74.
2. Gur D, et al. Changes in breast cancer detection and mammography recall rates after the introduction of a computer-aided detection system. J Natl Cancer Inst. 2004; 96:185–90. [PubMed: 14759985]
3. Warm, JS. Workload transition: Implications for individual and team performance. Wicken, BMHCD., editor. National Academy Press; Washington, DC: 1993. p. 139-170.
4. Mackworth, J. Vigilance and Attention. Penguin Books; Harmondsworth, England: 1970.
5. Chun MM, Wolfe JM. Just say no: How are visual searches terminated when there is no target present? Cognitive Psychology. 1996; 30:39–78. [PubMed: 8635311]
6. Palmer J, Verghese P, Pavel M. The psychophysics of visual search. Vision Res. 2000; 40:1227–1268. [PubMed: 10788638]
7. Kundel, HL. Medical Imaging 2000: Image Perception and Performance. Krupinski, EA., editor. 2000. p. 135-144.
8. Gur D, Rockette HE, Warfel T, Lacomis JM, Fuhrman CR. From the laboratory to the clinic: the “prevalence effect”. Acad Radiol. 2003; 10:1324–6. [PubMed: 14626308]

Figure 1.
Stimuli: Observers searched for tools in displays with semi-transparent objects placed randomly on a noisy background.

Figure 2.

Figure Two: the effects of target prevalence on search performance.

2a: Error data: When targets were rare (1% prevalence - black bars) observers made more than 4X times the errors made when targets were common (50% prevalence - white bars). Data are averages of 12 observers. Error bars show ± 1 s.e. for those 12 error rates. Gray bars show 10% prevalence results.

2b: Left: Reaction Times for 50% prevalence. Typical reaction times are longer when the target is absent (open symbols) than when targets are present (closed). Miss error RTs are shown by diamonds.

2c: Right: Reaction Times for 1% prevalence. However, when prevalence is low, observers make “absent” responses that are faster than the “present”. This leads to elevated error rates. For 1c&d, error bars (s.e.) fall within data points.