Homo economicus in visual search

Vidhya Navalpakkam
Division of Biology, California Institute of Technology, Pasadena, CA, USA

Christof Koch
Division of Biology, Division of Engineering and Applied Science, Computation and Neural Systems, California Institute of Technology, Pasadena, CA, USA

Pietro Perona
Division of Engineering and Applied Science, Computation and Neural Systems, California Institute of Technology, Pasadena, CA, USA

How do reward outcomes affect early visual performance? Previous studies found a suboptimal influence, but they ignored the non-linearity in how subjects perceived the reward outcomes. In contrast, we find that when the non-linearity is accounted for, humans behave optimally and maximize expected reward. Our subjects were asked to detect the presence of a familiar target object in a cluttered scene. They were rewarded according to their performance. We systematically varied the target frequency and the reward/penalty policy for detecting/missing the targets. We find that 1) decreasing the target frequency will decrease the detection rates, in accordance with the literature. 2) Contrary to previous studies, increasing the target detection rewards will compensate for target rarity and restore detection performance. 3) A quantitative model based on reward maximization accurately predicts human detection behavior in all target frequency and reward conditions; thus, reward schemes can be designed to obtain desired detection rates for rare targets. 4) Subjects quickly learn the optimal decision strategy; we propose a neurally plausible model that exhibits the same properties. Potential applications include designing reward schemes to improve detection of life-critical, rare targets (e.g., cancers in medical images).

Keywords: computational modeling, detection, learning, search, visual cognition, reward, target rarity

Citation: Navalpakkam, V., Koch, C., & Perona, P. (2009). Homo economicus in visual search. *Journal of Vision, 9*(1):31, 1–16, http://journalofvision.org/9/1/31/, doi:10.1167/9.1.31.

Introduction

The behavioral and neural mechanisms of reward or value-based economic decision making (Kahneman & Tversky, 2000; Sugrue, Corrado, & Newsome, 2005) and sensory-based decision making (Gold & Shadlen, 2007; Green & Swets, 1966) have been extensively studied in humans and animals. In comparison, much less is known about how reward combines with sensory information (likelihood and prior) to guide decisions. Here, we investigate sensory-economic decision making in the context of a visual search task where a familiar target object must be found in a cluttered scene with several distracting objects. Humans and animals rely on visual search to detect food, mates, and predators. Much research in visual search behavior has focused on the role of sensory information such as salience of the target (Itti & Koch, 2001; Nothdurft, 1992); amount of background clutter in terms of number (Treisman & Gelade, 1980), heterogeneity (Duncan & Humphreys, 1989; Rosenholtz, 2001) of distracting objects; and knowledge of target and distractor statistics (Navalpakkam & Itti, 2007; Vickery, King, & Jiang, 2005). In contrast, the role of reward outcomes in visual search behavior is relatively unknown.

A recent study (Wolfe, Horowitz, & Kenner, 2005) shows that target frequency plays an important role in visual search performance. While earlier research was limited to stimuli with high target frequencies (typically 50%; Palmer, Verghese, & Pavel, 2000; Treisman & Gelade, 1980), this study tested subjects on a range of target frequencies as low as 1% and found an alarming drop in target detection rates as target frequency decreased. Several attempts to improve the detection rates, such as encouraging subjects to slow down, or by increasing the frequency of the target category, failed (Wolfe et al., 2007). This raises concern for life-critical searches such as detecting rare diseases in medical images (some types of cancer have 0.3% frequency; Gur et al., 2004) and detecting weapons in airline passenger luggage (Rubenstein, 2001). Here, we investigate whether changing the reward outcomes (e.g., increasing the reward received upon finding the target and increasing penalties upon missing the target) can improve detection rates.

doi: 10.1167/9.1.31 Received March 18, 2008; published January 23, 2009 ISSN 1534-7362

Earlier studies in signal detection considered reward payoff manipulations (Green & Swets, 1966; Healy & Kubovy, 1981; Kubovy & Healy, 1977; Lee & Janke, 1964, 1965; Lee & Zentall, 1966; Maddox, 2002) but only for high or medium target frequencies (mostly 50%, some 10–25%). Two robust findings emerged:

1. the decision criterion was more conservative than the optimal decision criterion (i.e., shifted toward fewer target detections),
2. changing the target frequency, but not reward, lead to near-optimal shifts in the decision criterion.

These studies found, contrary to the assumption in standard economic theories, that subjects do not maximize expected reward. These studies tested subjects on simple tasks such as whether the stimulus at a single location contained a target or not. Here, we test subjects' performance in more complex tasks such as whether a multi-item display contains a target or not. In Experiment 1, we replicate the previous findings that changing sensory priors (target frequency) causes an optimal shift in the decision criterion, while reward has a suboptimal effect. However, this result is confounded by the subject's non-linear utility function (Kahneman & Tversky, 1979; von Neumann & Morgenstern, 1953), rendering subject's perceived reward different from the objective reward value. In Experiments 2 and 3, we redesign the visual search task as a contest between subjects, with an attractive cash prize for the winner. Under such competitive settings that encourage a literal interpretation of the reward scheme, we find that subjects maximize expected reward per trial and operate at the optimal decision criterion. In Learning the optimal decision criterion section, we report that when confronted with any new target frequency and reward scheme, subjects learn the optimal decision criterion rapidly. We propose and simulate neurally plausible models of such rapid reward-based learning. We conclude in the Discussion section, by illustrating how reward schemes may be designed to improve, and to yield desired target detection rates for any target frequency, and outline potential applications of our findings.

Effect of reward and target frequency

Experiment 1

We asked subjects to detect an oddly oriented line in briefly flashed pictures containing a number of parallel lines (Figure 1a). Detecting a target when none is present is called a 'false alarm' error, while failing to detect a target when one is present is called a 'miss' error. We measured the frequency of these errors as a function of the target frequency (50%, 10%, 2%) and of the reward structure. As indicated in Table 1, we experimented with two reward schemes: in Experiment 1A, the reward was 'Neutral', i.e., both types of errors were equally penalized and both types of correct responses were equally rewarded. In Experiment 1B, missing a target was penalized much more seriously than generating a false alarm (we call this the 'Airport' reward scheme as missing a bomb hidden in a suitcase is prohibitively expensive compared to false alarms that result in a relatively quick manual inspection). We also tested a third reward scheme 'Gain' (Table 1c) where subjects gained much more when they correctly detected the target than when they correctly rejected it.

Four naive subjects (Caltech students and postdocs, with normal or corrected vision) participated with informed consent and IRB approval. They received paid compensation of $15.50/hour for their participation. In addition, in Experiment 1, subjects were paid one cent for each point earned during the experiment ($0 if the total points earned was negative). The display size was 28 × 21 degrees and each stimulus was 2 × 0.2 degrees in size. In Experiments A and B, each trial began with a central fixation cross (for 250 ms), followed by the search display (for 50 ms), followed by a blank (until SOA), and the mask (until keypress). The search display contained 12 oriented bars arranged in a ring around the center (eccentricity 6 degrees) with additional spatial jitter up to 0.5 degrees. The target was always oriented at 70 degrees (from the horizontal) and the distractors were oriented at 60 degrees. The mask contained the target and distractor superimposed at each location. Each trial was a yes/no task where the subject had to respond either target present or absent. The trial timed out at 4 s with the wrong response, to prevent subjects from avoiding response. At the end of each trial, subjects received feedback on the reward earned on the trial, as well aggregate reward earned on the block.

Experiments 1A and 1B were divided into multiple sessions (lasting up to an hour a day), each consisting of blocks of 100 trials each. Subjects completed all blocks in one experimental condition (a particular combination of target frequency and reward scheme), before proceeding to the next. Each subject performed 3 blocks (300 trials) in the 50% target frequency condition, and 3 blocks (300 trials) in the 10% condition. The sequence of experimental conditions was randomized across subjects, and subjects saw different random sequences of target present and absent displays (an average of 150 targets in the 50% target frequency condition, and 30 targets in the 10% condition).

The experiments began with a procedure to calibrate SOA. To become familiar with the task, subjects were trained on the 50% target frequency trials in the Neutral reward scheme with decreasing SOAs (starting from 500 ms and decreasing in steps of 50 ms as subjects

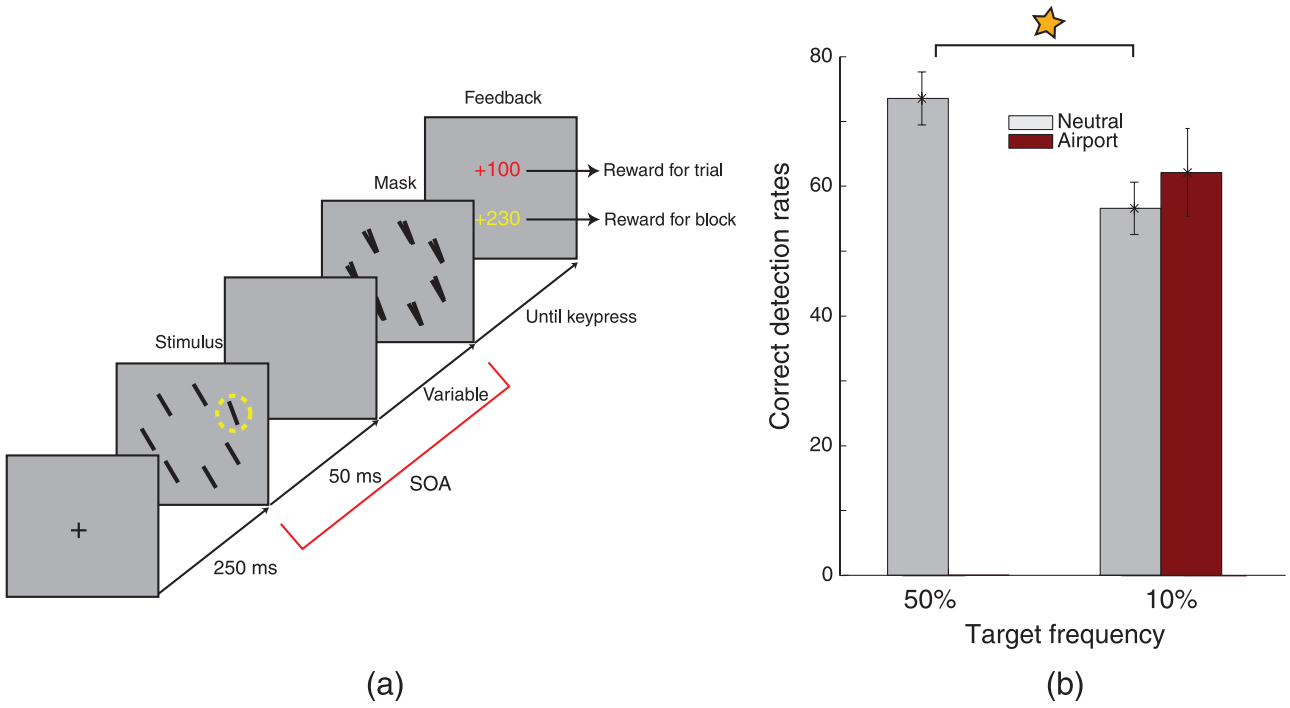

Figure 1. (a) Stimulus used in Experiments 1 and 2. Subjects were instructed to search for the target bar (oriented 70° clockwise from the horizontal) among eleven distractor bars (oriented 60°). Each display was presented briefly (Palmer et al., 2000; 100–200 ms SOA, to minimize eye movements) and was followed by a mask (Braun, 1994). After the mask appeared, subjects were asked to respond whether the target was present or not. Subjects received feedback on correct and incorrect responses. (b) Results from Experiment 1: In the Neutral reward scheme, target detection rates dropped as the target frequency decreased from 50% to 10%. Changing the reward scheme from Neutral to Airport did not significantly improve the detection rates in the 10% target frequency condition.

improved) until detection rates stabilized at 75–80%. The SOA was thereafter fixed at the stabilized value for the entire session. Task performance is known to improve with practice over days (Goldstone, 1998), hence to equate task difficulty in the 50% Neutral reward condition across all sessions, we repeated the SOA calibration procedure at the beginning of each day's session. The SOAs for subjects ranged from 100 to 200 ms.

Before the start of a new experimental condition, subjects received one block of training (100 trials) to become familiar with the new target frequency or reward scheme. During training (as well as the main experiment), subjects received feedback on reward earned per trial, so they could adjust their decision criterion τ to optimize reward.

Results

In Experiment 1A, we found that for a fixed reward structure (Neutral), as the target frequency decreases from 50% to 10% the detection rate (pooled across subjects) decreases significantly (p-value < 0.05, two-tailed t-test) from close to 75% down to 55% (Figure 1, light gray bars; mean ± standard error in detection rates for 50% and 10% target frequencies are 73 ± 4% and 56 ± 4%). This replicates the finding by Wolfe et al. (2005) that detection rates drop as the target frequency decreases. In Experiment 1B, we studied the effect of changing the reward structure from Neutral to Airport (Table 1b) when the target frequency decreases to 10% (infrequent). We found

	r_{00} (Correct rejection)	r_{01} (False alarm)	r_{10} (Miss)	r_{11} (Correct detection/hit)
(a) Neutral	+1	−50	−50	+1
(b) Airport	+1	−50	−900	+100
(c) Gain	+1	−50	−50	+950

Table 1. Reward schemes. The numbers inside the boxes indicate rewards (positive numbers) and penalties (negative numbers) in arbitrary units. r_{00} refers to the reward upon correctly rejecting the target when it is absent; r_{01} is the penalty upon falsely reporting a target; r_{10} is the penalty for missing the target when it is present; r_{11} is the reward for correctly detecting the target when it is present. **(a) Neutral reward scheme**—equal penalty on both false alarm and miss errors. **(b) Airport reward scheme**—severe penalty on miss errors compared to false alarms. **(c) Gain reward scheme**—generous reward for finding the target compared to correct rejection.

that the detection rate increases but not significantly (p-value > 0.05, two-tailed t-test; Figure 1, red bars; detection rates for 10% target frequency is 62 ± 7%).

Ideal observer model

To gain a quantitative understanding of the influence of target frequency and reward on detection performance, we turn to an ideal observer that maximizes the expected reward per trial. We use signal detection theory (Green & Swets, 1966) to model the ideal observer. Note that there exist alternate theories based on sequential analysis (for review, see Bogacz et al., 2006) to model speed-accuracy tradeoffs and reaction time measures. However, since ours is a brief fixed length display paradigm (100–150 ms SOA) without significant differences in reaction time, we follow previous literature in brief display paradigms

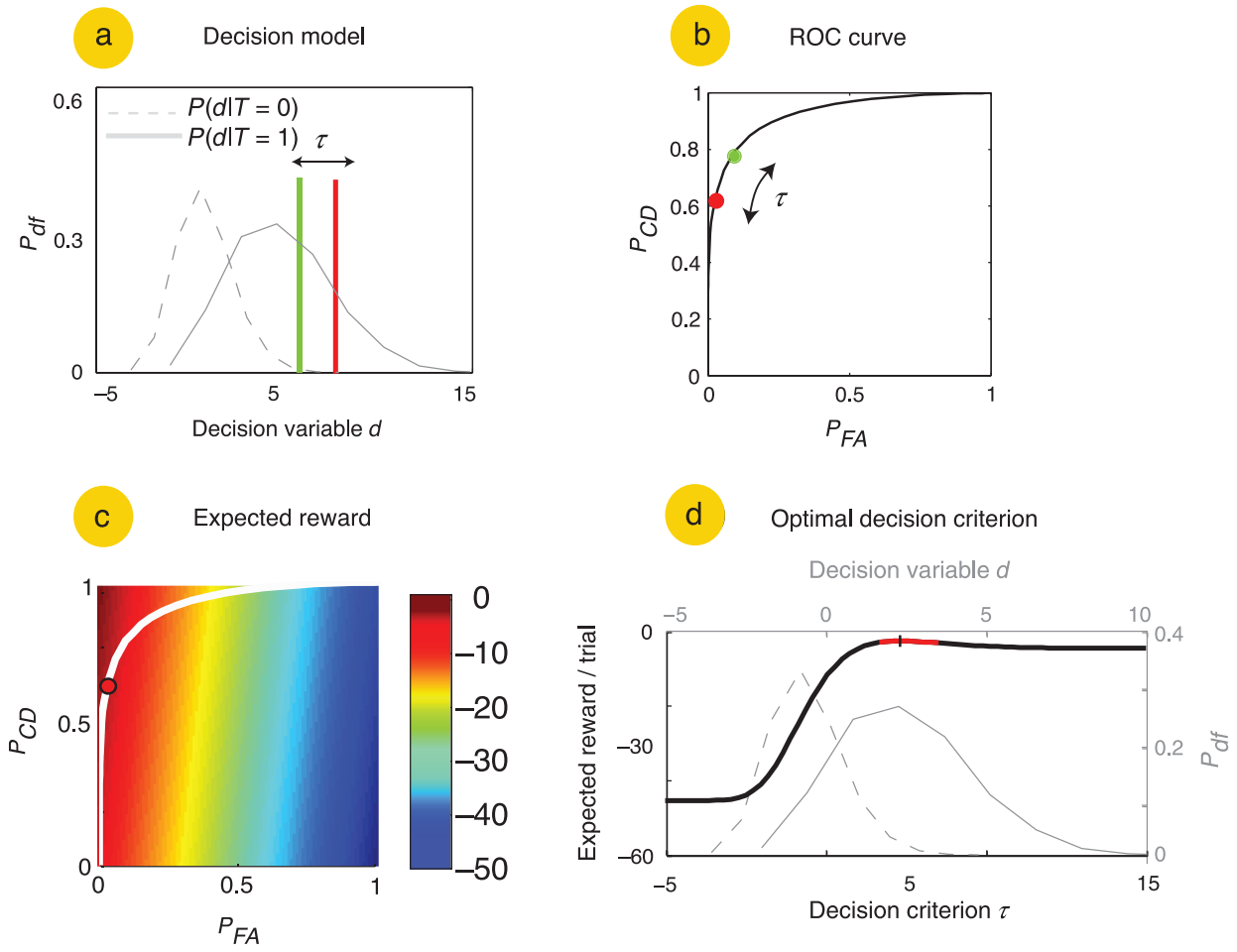

Figure 2. Ideal observer model. This figure illustrates our model and shows how an ideal observer would shift the decision criterion in the 10% target frequency (Neutral reward) condition to maximize expected reward per trial. (a) The two curves represent the distribution of decision variable values when the target is absent (dashed line, left) and present (solid line, right) in the display (d' = 1.52, see Comparing subjects and ideal observers section for details on how they are obtained from subject's data). The two vertical lines represent different decision criteria τ: the green criterion produces equal number of missed targets and false alarms, while the red criterion generates few false alarms at the expense of more missed targets. (b) The tradeoff between missed targets and false alarms for different criteria can be visualized using an ROC curve (the two dots are color coded to correspond to the two criteria in (a)). As the decision criterion shifts to the right in (a), the observer moves down the ROC curve in (b), resulting in more misses than false alarms. (c) The expected reward for each trial ($E[R]$), which depends on the number of missed targets and false alarms and their costs (see Equation 1), is shown in this color-coded map. The red color indicates high expected reward values, while blue indicates low values. The ideal observer (red symbol) operates at the point of maximum expected reward on the ROC curve (shown in white). (d) The gray curves show the distribution of decision variables d to target present (solid) and absent displays (dashed) as in (a). The expected reward for different choices of the decision criterion is superimposed (thick black line). The expected reward increases as the decision criterion shifts to the right and then decreases. Hence the ideal observer shifts his/her criterion slightly to the right (from the point of intersection of curves $P(d \mid T = 1)$ and $P(d \mid T = 0)$) to maximize reward. The error bar denotes standard error of the mean reward (*SEM*) and the red line indicates the range of decision criteria that lead to maximum $E[R]$ ±*SEM*.

(Palmer et al., 2000; Verghese & Stone, 1995) and use signal detection theory to model our ideal observer. We explain the theory below (illustrated in Figure 2a). Each display is represented internally by a single number, the decision variable d (see Comparing subjects and ideal observers section for details of the decision variable). The observer maintains a fixed internal representation of decision variables in target present and absent displays (Comparing subjects and ideal observers section explains how this internal representation may be obtained from subject's data; for now, let us assume these are known). The observer uses a linear discrimination threshold as a decision criterion (τ) to distinguish between likelihood of target presence ($P(d|T = 1)$) vs. absence ($P(d|T = 0)$) in the display (Palmer et al., 2000). Upon seeing a new display, the observer decides whether the target is present or absent by checking whether the observed decision variable d exceeds the decision criterion τ. Detection performance varies as the decision criterion changes and is visualized by the Receiver Operating Characteristic (ROC) curve that plots the probability of correct detection (hit) (P_{CD}) vs. false alarm (P_{FA}) for all possible values of the decision criterion τ (Figure 2b).

How does an ideal observer behave according to this model? The ideal observer will choose the decision criterion τ that maximizes the expectation of reward at each trial. The expected reward per trial $E[R]$ can be computed for different points on the ROC curve (P_{FA}, P_{CD}) explicitly by summing the four possible outcomes, each weighted by the probability of the corresponding event as described in Table 1. Refer to Table 1 for notation. P_1 is the probability of target presence ($P_0 = 1 - P_1$ is that of target absence):

$$E[R] = r_{11}P_1P_{CD} + r_{10}P_1(1-P_{CD}) + r_{01}P_0P_{FA} + r_{00}P_0(1-P_{FA}). \quad (1)$$

Note that both P_{CD} and P_{FA} are functions of τ. The ideal observer will choose a value of τ that produces P_{CD} and P_{FA} such that the previous expression is maximized.

Figures 2c and 2d illustrate this. As seen in Figure 2c, for low target frequency (10%) conditions in the Neutral reward scheme, the expected reward per trial ($E[R]$) is higher in the middle left region of the ROC curve. Hence, the optimal behavioral strategy is to operate at the middle left on the ROC curve. This is better visualized in Figure 2d, where $E[R]$ is shown as a function of the decision criterion τ. To maximize $E[R]$, the ideal observer would shift the decision criterion τ slightly to the right (from the point of intersection of curves $P(d|T = 1)$ and $P(d|T = 0)$), similar to the red line in Figure 2a. Consequently, the false alarms decrease and the detection rates are poor (red symbol in Figure 2b). Thus, the ideal observer theory predicts that detection performance for rare targets will be poor (as observed in previous studies (Wolfe et al., 2005) and replicated here). In the next section, we compare the behavior of the ideal observer and our subjects.

Comparing subjects and ideal observers

For each subject and the average subject, we computed the best fitting ROC curve across all experimental conditions (50%, 10% Neutral, and 10% Airport) as follows. We assumed that the sensory response to a stimulus at the ith location ($i \in \{1 \ldots 12\}$) is drawn from a Gaussian distribution, $x_i \sim G(\mu, \sigma)$ ($\mu = 1$ for the target, and $\mu = 0$ for the distractor). Next, we assumed that the display is represented internally as a single number, the decision variable d. Examples of decision variables for our yes/no visual search task include the maximum response at all locations in the display (Palmer, Ames, & Lindsey, 1993; Verghese, 2001), the likelihood ratio of target presence vs. absence in the display (see ideal observer, A.2.3, (Palmer et al., 2000); for use of the ideal rule in 2AFC tasks and cueing tasks, see Eckstein, Shimozaki, & Abbey, 2002; Schoonveld, Shimozaki, & Eckstein, 2007). We choose the latter as it is the ideal rule to integrate information across the display. According to this rule, the likelihood of target presence (or absence) in the display can be expressed in terms of the likelihood of target presence (or absence) at each location. Let $T = 1$ represent target presence in the display ($T = 0$ denotes target absence) and let $T_i = 1$ represent target presence at the ith location in the display ($T_i = 0$ denotes target absence):

$$P(x|T = 0) = \prod_{i=1}^{12} P(x_i|T_i = 0) \quad \text{(assuming responses at each location are drawn independently)}, \quad (2)$$

$$P(x|T = 1) = \frac{1}{12}\sum_{i=1}^{12} P(x_i|T_i = 1)\prod_{j\neq i} P(x_i|T_i = 0) \quad \text{(assuming target is equally likely to appear anywhere)}, \quad (3)$$

$$d = \frac{P(x|T = 1)}{P(x|T = 0)} = \frac{1}{12}\sum_{i=1}^{12}\frac{P(x_i|T_i = 1)}{P(x_i|T_i = 0)}. \quad (4)$$

Thus the decision variable (likelihood of target presence vs. absence in the display) equals the sum of likelihood ratios of the target presence vs. absence at each location in the display. In this model, we assume that the observer has a fixed internal representation of the values of the decision variable in target present and absent displays ($P(d|T = 1)$,

$P(d|T=0)$) and varies the decision criterion τ to operate at different points on the ROC curve. We find the best fitting ROC curve through a maximum likelihood estimation procedure that determines the value of σ that maximizes the likelihood of subject's data. The resulting ROC curves are asymmetric (Figure 2b), as reflected by the difference in shape of the distributions of decision variables to target present vs. absent displays (Figure 2a). Such asymmetric ROCs have also been observed in other studies (Wolfe et al., 2007).

Do subjects behave as ideal observers? To test this, we compared our subjects' reward to that of the ideal observer. We find that for different target frequencies (50%, 10%) in the Neutral reward scheme (i.e., the default, symmetric reward scheme), subjects operate in the region of maximum reward (Figures 3b and 3c)—there is no difference between subject reward and maximum expected reward (z-test, p-value > 0.05). This replicates previous findings (Green & Swets, 1966; Healy & Kubovy, 1981; Kubovy & Healy, 1977; Lee & Janke, 1964, 1965; Lee & Zentall, 1966; Maddox, 2002) that changing the sensory prior (target frequency) causes an optimal shift in the decision criterion. In contrast, we find that changing the reward scheme from Neutral to Airport for a fixed target frequency (10%) yields a suboptimal shift in the decision criterion (Figures 3c and 3d)—subjects fail to maximize expected reward in the 10% Airport condition (they have fewer target detections and fewer false alarms than the ideal observer). These results confirm previous findings (Green & Swets, 1966; Healy & Kubovy, 1981; Kubovy & Healy, 1977; Lee & Janke, 1964, 1965; Lee & Zentall, 1966; Maddox, 2002) that subjects adjust optimally to changes in target frequency, but they adjust suboptimally to changes in the reward scheme.

Experiment 2: Contest between subjects

It is somewhat surprising that changing the reward scheme in Experiment 1 did not lead to a significant improvement in target detection performance, especially since several economic theories assume humans to be reward or utility-maximizing agents (von Neumann & Morgenstern, 1953). One possible reason for the small effect of reward is that the reward values, r_{ij} ($i, j \in \{0, 1\}$) in Equation 1, are perceived differently by the subjects (not as the objective reward value set by the experimenter)

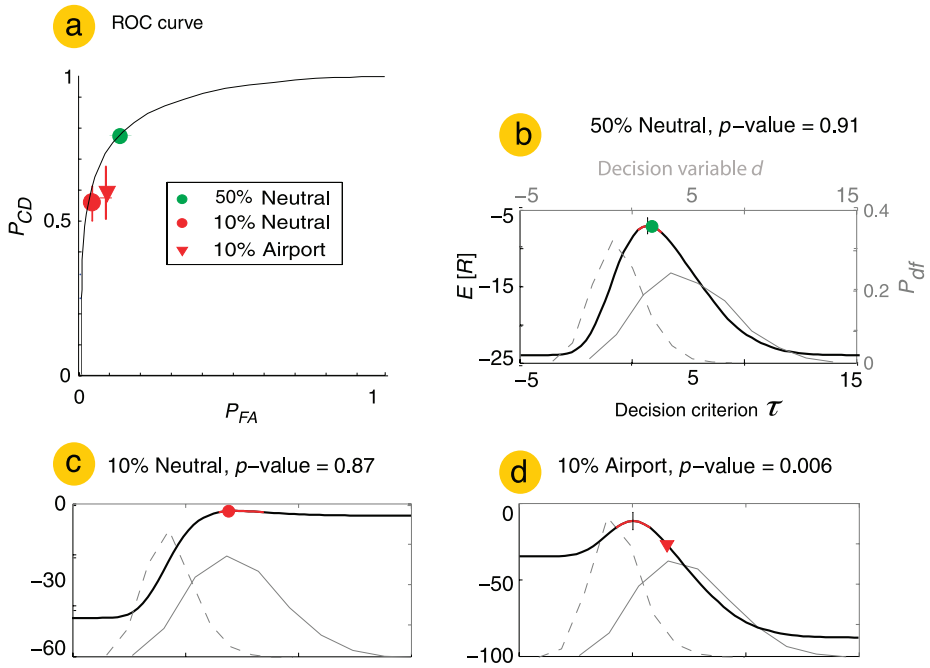

Figure 3. Experiment 1: Comparing subject behavior to the ideal observer. (a) ROC curve. Averaged subject behavior ($n = 4$) in the 50% and the 10% Neutral and in the 10% Airport conditions fits well to an ROC curve. (b–c) Changing the target frequency yields an optimal shift in the decision criterion. For each target frequency in the Neutral reward condition, we show the expected reward per trial for different points on the ROC curve. The operating point of the average subject (shown in big colored symbols) lies in the region of maximum reward (shown in red line)—the subject's reward is statistically indistinguishable (z-test, significance level 0.05, p-value indicated in the title of each panel) from the maximum reward (mean ±*SEM*, shown with thin black error bars). Similar results are obtained for the individual subjects. This shows that for different target frequencies in the Neutral reward scheme, subjects operate at the optimal decision criterion that maximizes expected reward per trial. (d) Changing the reward payoff scheme yields a suboptimal shift in the decision criterion. Subjects do not maximize expected reward when the reward scheme changes from Neutral to Airport for a fixed target frequency (10%).

due to filtering through a subjective utility function. Costs and benefits perceived by humans are not linearly related to monetary rewards; in particular, the utility of a dollar gained is typically lower than the negative utility of a dollar lost (Kahneman & Tversky, 1979), and the utility of a dollar diminishes with increasing gains (von Neumann & Morgenstern, 1953). Such non-linearity confounds the results from previous studies (replicated in Experiment 1) as the observed suboptimality (failure to maximize expected reward) may be due to the diminished utility of the reward values (e.g., a reward value of 100 points perceived as less than 100). A diminished value of reward and penalties is consistent with the conservative strategy (low detection rates and low false alarms) adopted by subjects in Experiment 1 (and previous studies).

To encourage subjects to follow our reward scheme faithfully, we redesigned each experimental condition as a contest among subjects: the sum of the gains and losses in each experimental condition was used as a score. The subject with the highest score in each condition won $50 (in addition to the regular hourly pay of $15.50). Subjects were unaware of each other's performance. Subjects may lose the contest even after earning 1000 points, hence in order to win, subjects must not diminish their effort or utility and must use the reward values literally in order to maximize their scores. In Experiment 2, we repeated Experiment 1 (on the same subjects) under such competitive settings. In addition to 50% and 10% target frequency, we also tested subjects at 2% target frequency. In addition to one block of training, each subject performed 3 blocks of 300 trials in total in the 50%, 10% target frequency conditions, 10 blocks of 1000 trials in total, containing an average of 20 target present trials in the 2% target frequency condition.

Results

In Experiment 2A, we found that for a symmetric reward structure (Neutral), as the target frequency decreases from 50% to 2% the detection rate (pooled across subjects) decreases significantly (p-value < 0.05, two-tailed t-test) from close to 80% down to 30% (Figure 4, light gray bars; mean ± standard error in detection rates for 50%, 10%, and 2% target frequencies are 78 ± 1%, 58 ± 4%, and 29 ± 4% respectively). This replicates the finding by Wolfe et al. (2005) that rare targets are often missed. In Experiment 2B, we studied the effect of changing the reward structure from Neutral to Airport (Table 1b) when the target frequency decreases to 10% (infrequent) and 2% (rare). We found that the detection rate increases significantly (more than double when the target is rare), restoring performance to levels statistically indistinguishable from the 50% target frequency condition (Figure 4, red bars; detection rates for 10% and 2% target frequencies are 77 ± 3% and 61 ± 11%). In fact, in two out of four subjects, the detection rate in the 10% Airport

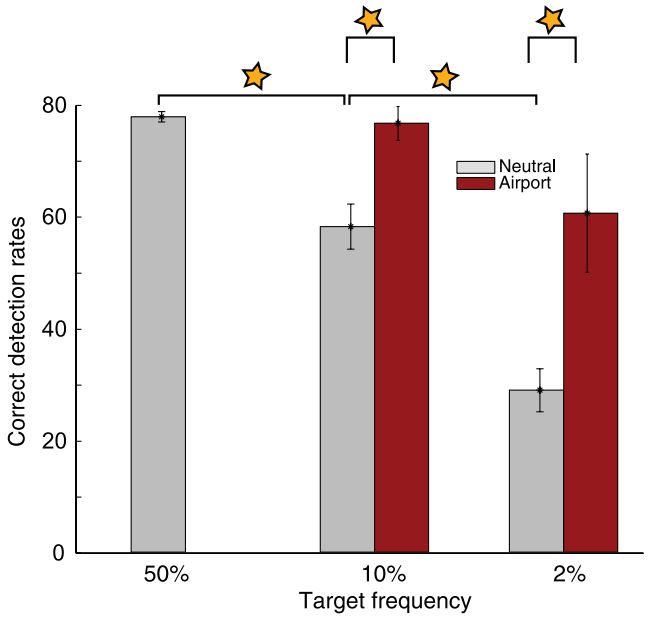

Figure 4. Results of Experiment 2. This figure shows that in the Neutral reward condition (gray bars), target detection rates drop significantly when the target becomes rare. However, changing the reward scheme from Neutral to Airport (red bars) yields a significant increase in detection rates, showing that reward has a strong influence on detection performance.

condition was significantly higher than in the 50% Neutral condition. This shows that reward can significantly influence target detection performance.

Next, we investigate to what extent the ideal observer model can account for subject's behavior. We find that a single ROC curve fits our subjects' behavior well in different reward and target frequency conditions (Figure 5a, $d' = 1.52$) and allows good quantitative predictions as shown below. When target frequency decreases, subjects move down the ROC curve toward lower detection and lower false alarm rates. In contrast, when reward changes from the Neutral to the Airport scheme, subjects move up along the ROC curve toward higher detection and false alarm rates. This suggests that the internal representation of the stimulus display ($P(d|T = 0)$, $P(d|T = 1)$) remains the same across all conditions—i.e., the subjects' attention/arousal levels remain the same—but the decision criterion changes.

Can the ideal observer model predict subject's behavior quantitatively? Our results show that subjects operate in the region of maximum reward (Figures 5b–5f)—there is no difference between subject's reward and maximum expected reward (z-test, p-value > 0.05). Not only does this show that our model predicts subject's data well, but also that the behavior of subjects is optimal, i.e., subjects maximize expected reward per trial.

Subjects' behavior in the Airport reward scheme (Experiment 2B) may be interpreted as avoiding losses, as the penalty on missing the target is much higher than the penalty on a false alarm error. To test whether subjects

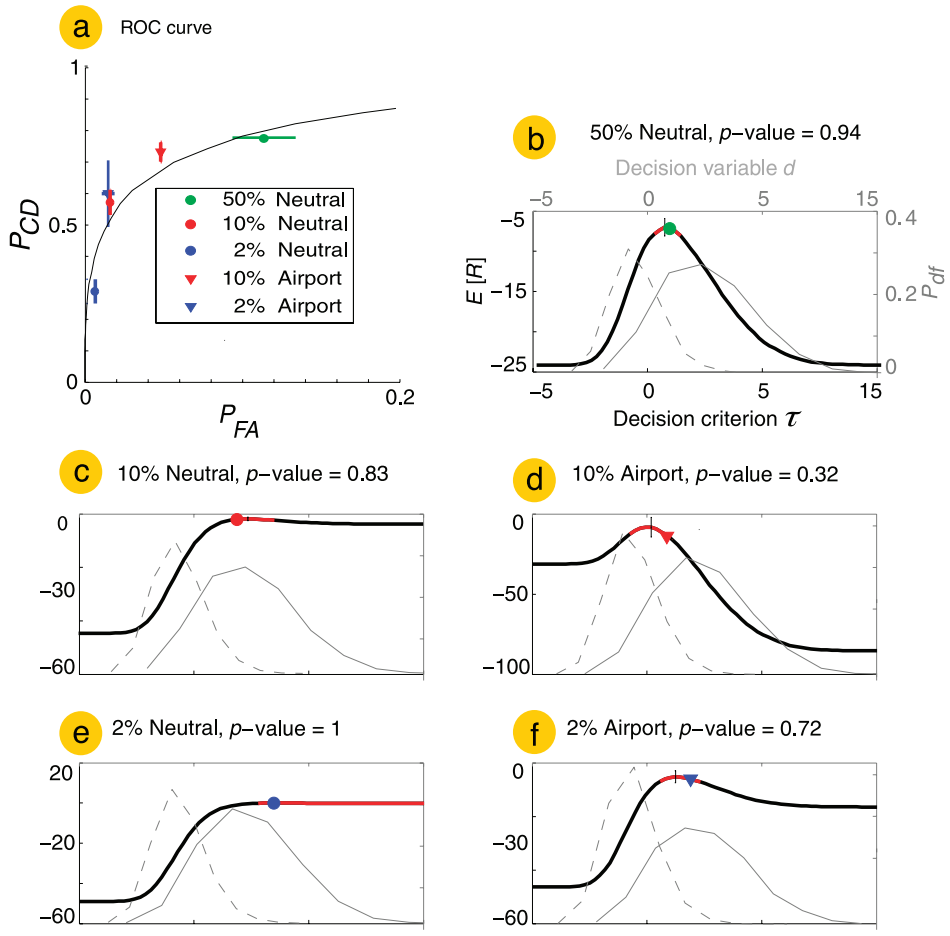

Figure 5. Experiment 2: Comparing subject's behavior to the ideal observer. (a) ROC curve. Subject's behavior (average as well as individual data) in different frequency and reward conditions fits well to an ROC curve. Only the averaged subject data are shown here due to space constraints. (b–f) Subjects maximize expected reward per trial. For each reward and target frequency condition, we show the expected reward per trial for different points on the ROC curve. The operating point of the average subject (shown in big colored symbols) lies in the region of maximum reward (shown in red line)—the subject's reward is statistically indistinguishable (*z*-test, significance level 0.05) from the maximum reward (mean ±*SEM*, shown in black error bars). Similar results are obtained for average as well as individual subject data.

are avoiding losses, or seeking gains, or both, we designed a 'Gain' reward scheme (Experiment 2C, Table 1) where the gain on correctly detecting the target is much higher than the gain on correctly rejecting it, and the penalty on miss and false alarm errors is equal. In this new reward scheme that emphasized gains rather than losses, we found again that subjects maximize their expected gain (Figure 6). Experiments 2B and 2C together show that subjects avoid losses, as well as seek gains; in other words, they maximize the overall expected reward.

Next, we tested whether subject's decisions could be explained purely based on sensory information (by ignoring rewards, i.e., setting $r_{ij} = k$ in Equation 1), or reward information (value-based choice that ignores sensory information, i.e., setting $P(d|T = 0) = P(d|T = 1)$ or $P_{CD} = P_{FA}$ in Equation 1). As seen in Figure 7, we find that only a model that combines sensory and reward information optimally can explain subject's data across all

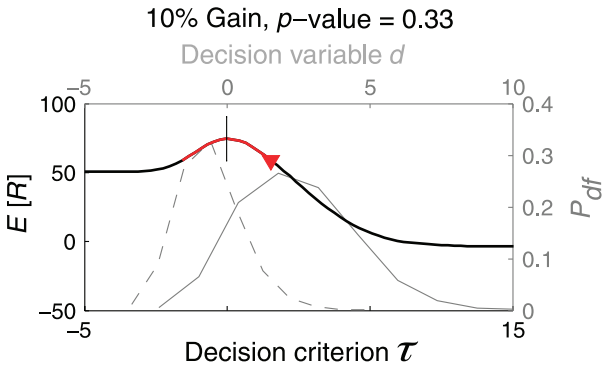

Figure 6. Subjects maximize expected gain. As in Figure 5d where subjects avoid losses in the Airport reward scheme, here (10% target frequency, Gain scheme) we find that the average subject ($n = 4$) maximizes expected gain per trial. Similar results are obtained for the individual subjects.

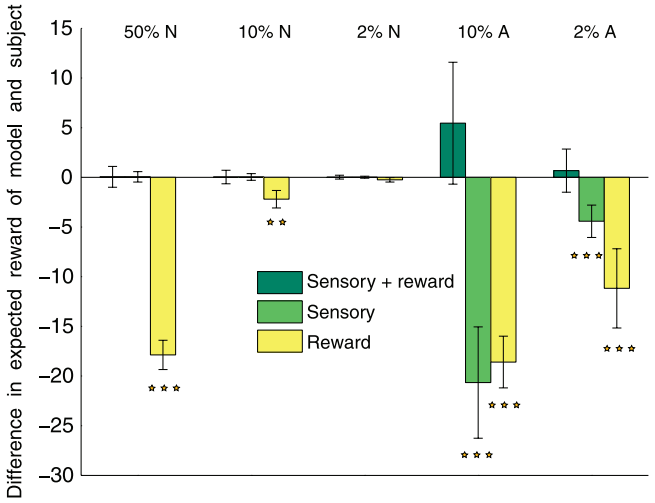

Figure 7. Subjects combine sensory evidence with reward optimally. Can subjects' decisions be explained by purely sensory or economic decision making models? To test this, we compare the predictions of three models of decision making based on 1) sensory evidence, 2) reward information, and 3) optimal combination of both. We test the ability of these models to predict subject's expected reward in different target frequency and reward schemes. The stars indicate significant difference between model's prediction and subject's reward (single star: p-value < 0.05, two stars: p-value < 0.01, three stars: p-value < 0.001). Subject data across all conditions are best explained by a model that combines sensory and reward information optimally.

target frequencies and reward schemes. Thus, subjects in our experiments are combining sensory priors and reward outcomes optimally, rather than purely sensory or economic decision making.

Learning the optimal decision criterion

Subject learning rates

In our experiments under competitive settings (Experiments 2A and 2B), we find that humans deploy the optimal decision criterion τ_{opt} that maximizes expected reward per trial. This suggests that humans learned τ_{opt} within the block of 100 trials of training that they received at the beginning of each experimental condition (i.e., before the start of a new combination of reward scheme and target frequency). We analyzed the training data as follows. For each experimental condition, we determined the correct detection/hit (P_{CD}) and false alarm rates (P_{FA}) as a function of the number of trials seen in the training sequence, used these to determine subject's reward using Equation 1, and asked when the subject's reward becomes statistically indistinguishable from the ideal observer's reward (z-test, significance level 0.05). Analysis of the training data reveals that on average, subjects learned τ_{opt} rapidly, within 14, 31, 7, 24, and 42 trials in the 50%, 10%, 2% Neutral, and 10%, 2% Airport conditions, respectively. What are the underlying computational and neural mechanisms of such rapid reward-based learning? How much information does a learner need to determine the optimal decision criterion (e.g., can the learner perform well without full knowledge of the statistics of responses in target present and absent displays)? How much memory does a learner need to determine and maintain the optimal decision criterion? In the next section, we address these issues and propose a neurally plausible model of learning.

Models of learning

We simulate five different models of learning: 1) fully informed, ideal observer who operates at the optimal decision criterion τ_{opt}, 2) optimal learner (with perfect and infinite memory), 3) bio learner (neurally plausible implementation of the optimal learner), 4) unit memory learner (who decides based on the previous trial only), 5) finite memory learners (who decide based on the previous 32, 64, and 128 trials). All model learners except the bio learner have perfect memory but with varying memory capacity (i.e., number of previous trials in memory). We explain each model below.

The ideal observer knows everything about the experiment except the ground truth of whether the target is present or absent on every trial ($T^i = 1$ if the target is present on the ith trial). Thus, the ideal observer knows the probability density of responses in the target present $P(d|T = 1) \sim \mathcal{G}(\mu_1, \sigma_1)$ and absent displays $P(d|T = 0) \sim \mathcal{G}(\mu_0, \sigma_0)$; the target frequency $P1$; and the reward structure (r_{00}, r_{01}, r_{10}, r_{11}). This observer can compute the optimal decision criterion τ_{opt} that maximizes expected reward $E[R]$ (Equation 1) using gradient ascent or other optimization techniques. This may be done ahead of seeing the stimulus, thus the ideal observer behaves optimally from the beginning.

Unlike the ideal observer with full knowledge, other models do not know the display statistics (distribution of decision variable values when the target is present or absent in the display), or target frequency. They only know the values of the decision variable, the ground truth (is the target present or not), and reward feedback in previously seen n trials (n = 1,32,64,128 for the finite memory learners, and infinite for the optimal learner). Rather than a slow process of explicitly learning the target frequency or display statistics, then determining the expected reward profile, and subsequently finding its peak through gradient ascent (or some optimization method), the models use a faster algorithm explained below.

The optimal learner (Figure 8) knows much less than an ideal observer—it knows only the ground truth T^i and

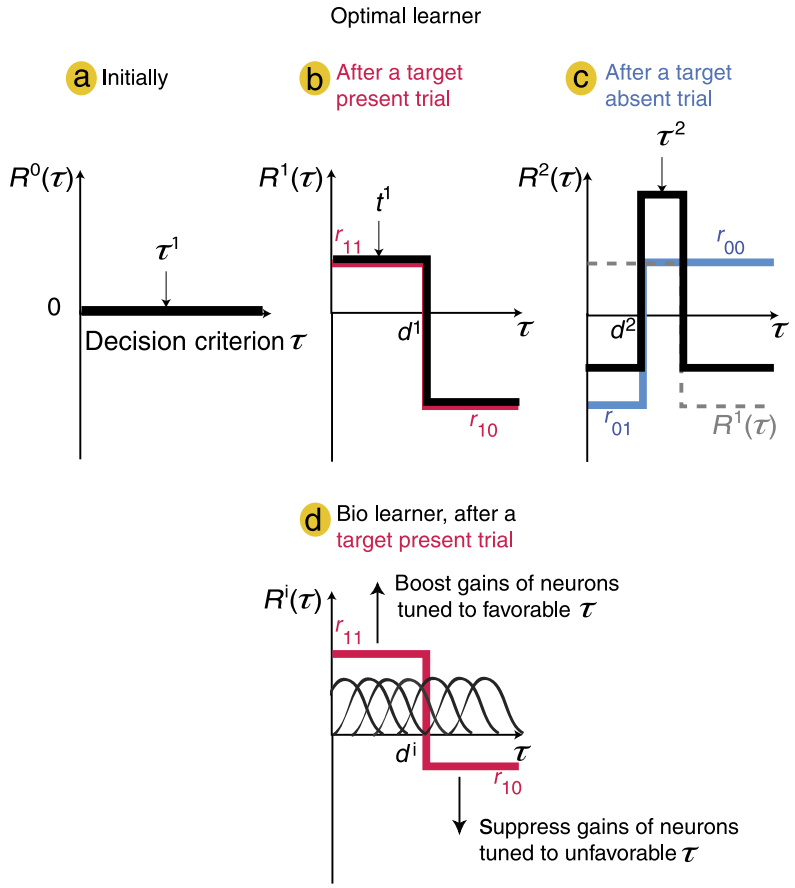

Figure 8. Optimal and bio learner. Panels (a)–(c) illustrate the functioning of the optimal learner. Initially, the reward earned by operating at all decision criteria τ is set to zero ($R^0(\tau) = 0, \forall\tau$). After seeing a target present trial whose value of decision variable is d^1, the reward score of favorable decision criteria ($\tau < d^1$) is incremented by the reward earned for correct detection or hit (r_{11}), and reward score of unfavorable decision criteria ($\tau \geq d^1$) is decremented by the penalty for missing the target (r_{10}). The resulting reward scores $R^1(\tau)$ is shown in black, and the model selects the decision criterion corresponding to maximum reward score as its criterion for the next trial (τ^1; note: if there are multiple decision criteria with maximum reward score, the model selects their mean). Upon seeing a target absent trial (decision variable value d^2), the model increments the reward score of favorable decision criteria ($\tau > d^2$) by r_{00} (otherwise, decrements reward score by r_{01}). The new decision criterion after these 2 trials (τ^2) is chosen as the one with maximum reward score. **Bio learner.** Panel (d) illustrates how the optimal model learner may be implemented neurally through gain modulation of a population of neurons encoding different decision criteria. After each trial, the gain of neurons encoding favorable decision criteria is enhanced, while other neurons (tuned to unfavorable decision criteria that will result in error responses) are suppressed. Further details are provided in the equations and main text.

reward feedback in n trials seen so far, and the corresponding values of the decision variable ($d^i|T^i$, $i \in \{1, n\}$). We call it the optimal learner as it learns τ_{opt} in a minimum number of trials (and using few parameters). For each decision criterion τ, the optimal learner computes the total reward $R^i(\tau)$ earned by operating at τ in the previous i trials and selects the τ with maximum reward as the decision criterion for the next trial (τ^{i+1}) (if there are multiple τ with maximum reward, the learner chooses the mean τ). The algorithm is explained below: Initially, all decision criteria (spanning the range of decision variable values $[d_{\min}, d_{\max}]$) have equal reward scores ($\forall\tau, R^0(\tau) = 0$) and hence are equally likely to be chosen. Upon observing a value of the decision variable d^i on the ith trial, the learner updates the score at each τ with the reward earned by operating at τ (which is positive for a correct response and negative for an incorrect response). The criterion for the next trial, τ^{i+1}, is chosen as the one that yields maximum reward $R^i(\tau)$. The exact algorithm is given below:

$$\begin{aligned} \text{Initialize} &: R^0(\tau) = 0, \forall\tau \in \{d_{\min}, d_{\max}\} \\ R^i(\tau) &= R^{i-1}(\tau) + r(\tau, d^i, T^i) \\ r(\tau, d^i, 1) &= r_{11}, \text{ if } \tau < d^i; r_{10}, \text{ else} \\ r(\tau, d^i, 0) &= r_{00}, \text{ if } \tau > d^i; r_{01}, \text{ else} \\ \tau^{i+1} &= \text{argmax}_\tau(R^i(\tau)). \end{aligned} \quad (5)$$

The finite memory learner is similar to the optimal learner except that it has limited memory of the previous k trials only ($k \in \{1,32,64,128\}$).

Bio learner (Figure 8) is a neurally plausible implementation of the optimal learner. We assume that the optimal decision criterion is encoded by a population of neurons (with Gaussian tuning curves) tuned to different decision criteria. Upon seeing the ith display with a value of decision variable d^i, all neurons whose preferred decision criteria favor the correct response undergo response gain enhancement proportional to the sum of reward gained by correct response and penalty avoided by incorrect response. The optimal decision criterion is set as the preferred criterion of the most active neuron in the population, read out through a winner-take-all competition. For simulation purposes, we implemented a population of neurons with equi-spaced Gaussian tuning curves, with standard deviation $\sigma = 0.25$ and inter-neuron spacing equal to σ. This population consists of 33 neurons representing values of τ in $[d_{\min}, d_{\max}]$:

$$\begin{aligned}
\text{Initialize}: g^0(\tau) &= 1, \forall \tau \in \{d_{\min}, d_{\max}\} \\
g^i(\tau) &= (1-\alpha) g^{i-1}(\tau) + \alpha r(\tau, d^i, T^i) \\
&\quad (\text{exponential smoothing factor } \alpha) \\
r(\tau, d^i, 1) &= (r_{11} - r_{10}), \text{ if } \tau < d^i; 0, \text{else} \\
r(\tau, d^i, 0) &= (r_{00} - r_{01}), \text{ if } \tau > d^i; 0, \text{else} \\
R^i(\tau) &= \textstyle\sum_{\tau'} g^i(\tau') \exp\left(-\frac{(\tau - \tau')^2}{2\sigma^2}\right) \\
\tau^{i+1} &= \text{argmax}_\tau (R^i(\tau)).
\end{aligned} \tag{6}$$

Simulation results

For each target frequency and reward scheme, we computed the expected reward per trial ($E[R]$, Equation 1) and error bar (*SEM,* standard error of the mean reward per trial) earned by the ideal observer (who operates at the optimal decision criterion). The error bar is computed as $\sigma[\bar{R}] = \sigma[\sum_i R_i/N] = \sigma[R]/\sqrt{N}$, where $\bar{R}$ is the mean reward per trial, R_i is the reward in the ith trial, R is the reward per trial, and N is the sample size. Rather than setting N to an arbitrarily large number of iterations, to ensure a fair comparison between the subjects and models, we set N to be the number of trials performed by the subject in each experimental condition, thus $N = \{300, 300, 1000\}$ in the 50, 10, and 2% target frequency conditions, respectively. We then simulated the models for 100 blocks of N trials (300 trials for 50, 10%; and 1000 trials for 2% frequency). On each trial, the decision variable (d) was drawn independently from a target present distribution ($P(d|T = 1)$) with probability P_1 (0.5, 0.1, 0.02 depending on target frequency) or from a target absent distribution ($P(d|T = 0)$) with probability $(1 - P_1)$. All learners were fed the same pseudo-random sequences as input. The distributions $P(d|T = 1)$ and $P(d|T = 0)$ were determined from the subjects' ROC curve (Figure 3, see Comparing subjects and ideal observers section for details on how to obtain these distributions from subject's data). The model's decision criteria evolved as a function of the values of the decision variable in previously seen displays (see equations in previous section). We determined the learning rate of each model learner as the number of trials (or displays) taken by the model to maximize expected reward (using a z-test to determine statistical difference between model's median reward and ideal observer's expected reward).

The learning curve of different models are shown in Figure 9 for different target frequency and reward schemes. Our simulations show that the optimal and bio learners perform similarly across all target frequency and reward conditions, consistently outperforming a unit memory learner. In particular, the optimal and bio models learn within 8, 16, and 2 trials (respectively) in the 50, 10 and 2% Neutral conditions; and within 16 and 64 trials in the 10 and 2% Airport conditions. As shown in Table 2, the trend in learning rates between humans and the models is consistent. The learning rate of the optimal and bio learners depends, intuitively, on the time taken to see a single target present and target absent trial—thus, a minimum of 2, 10, 50 trials are required on average for the 50, 10, 2% target frequency conditions. However, exceptions occur when the expected reward profile has a flat peak. For example, in the 2% Neutral condition, the expected reward profile has a flat peak (Figure 5e), so a large range of decision criteria (yielding minimal false alarm errors) are optimal as they maximize expected reward. An average of a single target absent trial suffices to shift the decision criterion near the optimal, so that false alarm errors are minimized. Accordingly, Figure 9c shows that the models are statistically indistinguishable from the ideal observer after the first trial (there is an increase in expected reward after 32 trials, but this is small and non-significant). Thus, although at a first glance, rapid learning within 2 trials in the 2% Neutral condition (Figure 9c) seems counterintuitive (it is much less than the average of 50 trials required to see a single target present and absent trial), rapid learning occurs due to the flat peak in the expected reward profile.

How much memory does a model need to learn as rapidly as the optimal model? Is larger memory better? As shown in Figure 9, we find that finite memory learners with equal or greater memory capacity than the learning rate of the optimal model perform equally well. The longest learning rate of the optimal model is 64 trials (in the 2% Airport condition). Thus, for the target frequencies considered here (50, 10, 2%), a minimum of 64 trials in memory is required to perform as well as the optimal learner with infinite memory. Although it may seem intuitive that model performance should improve with larger memory capacity, under our experimental conditions (specific payoff structure and number of trials per subject), we do not find any evidence to support this hypothesis (comparison of the mean reward of two

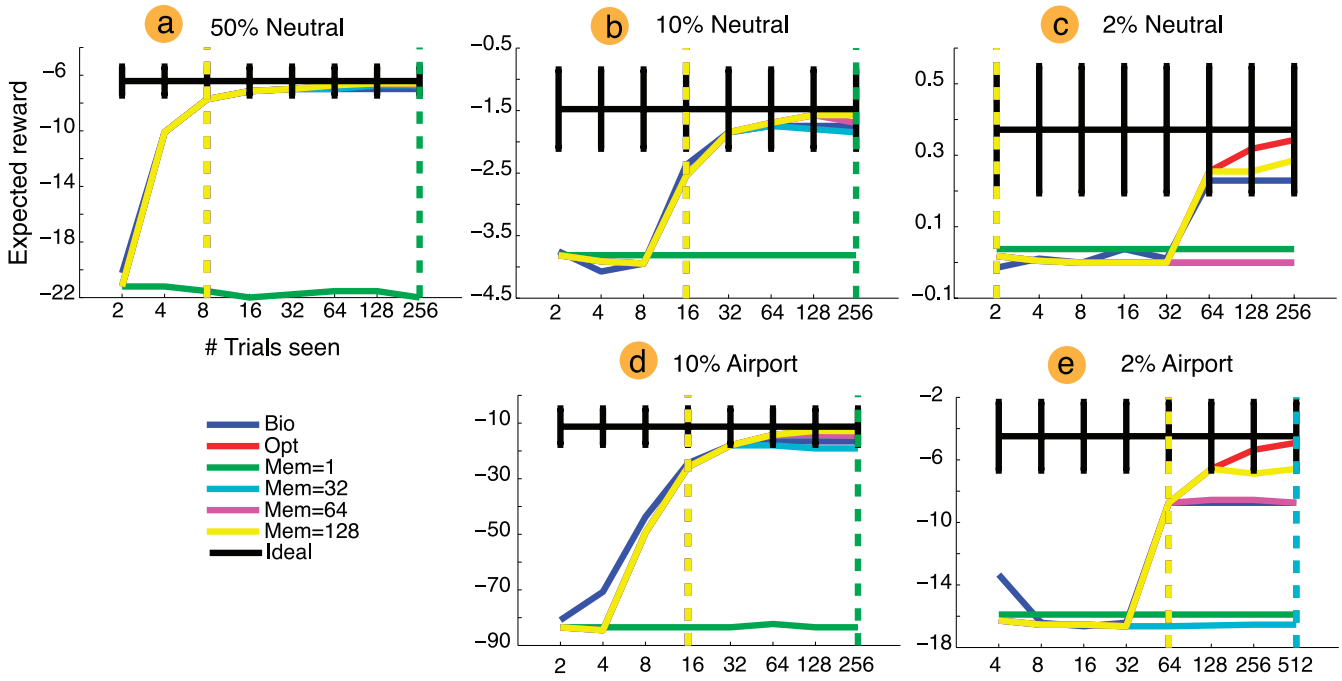

Figure 9. Learning rates of models. The distributions of expected reward per trial, *ER,* for various model learners (color coded) over 1000 simulated blocks are shown as a function of the number of trials seen. Consider the example in Panel (a): The expected reward of the ideal observer is shown in black (error bar represents the standard error of the mean). The different models are color coded (see legend). The learning curves of all models (except the unit memory learner) are similar (they overlap with the yellow curve). The dashed vertical line represents the number of trials *t* taken by the different models to learn to maximize *ER*—below *t* the model's median *ER* is significantly different from the ideal observer's *ER* (*z*-test, significance level 0.05). In (a), the dashed vertical line in green shows that the unit memory learner fails to learn the optimal decision criterion even after 256 trials, while the yellow dashed vertical line shows that all other models learn to maximize their expected reward within 8 trials. As seen in (a)–(e), the bio and optimal models learn to maximize *ER* within 8, 16, 2, 16, and 64 trials in the 50, 10, 2% Neutral; and 10, 2% Airport conditions, respectively. Across all conditions, finite memory learners with 64 or larger trials in memory perform as well as the optimal and bio models.

models with 128 or 64 trials in memory shows that the models are statistically indistinguishable, $p = 0.63$ from a *t*-test).

Experimental condition	Subject's learning rate	Optimal learner's rate
50% Neutral	14	8
10% Neutral	31	16
2% Neutral	7	2
10% Airport	24	16
2% Airport	42	64

Table 2. Learning the optimal decision criterion. The numbers inside the boxes indicate the average number of trials taken by four subjects and the optimal learner to maximize their expected reward. See the text for details on how the learning rates of subjects and the models are computed. Note that the subject's learning rate is computed from a single training sequence of 50 trials (in the 50, 10% target frequency condition) or 100 trials (in the 2% target frequency condition) and depends on the values of decision variable seen in that sequence (resulting in substantial variability across subjects). Across all experimental conditions, the trend in model's learning rates and subjects is consistent.

Discussion

To summarize, we investigated how reward outcomes combine with sensory prior (or target frequency). Most research has focused on medium and high priors like 10% and 50% (Palmer et al., 2000; Treisman & Gelade, 1980), ignoring low priors like 2% or less. However, such low priors are critical in several searches like detection of rare diseases in medical images (Gur et al., 2004) and detection of bombs in airline passenger bags (Rubenstein, 2001). We investigated how reward outcomes combine with a wide range of sensory priors (2, 10, 50%) to influence decision criterion and target detection rates in visual search. In our experiments under competitive settings, we find that humans engaged in visual search rapidly learn the optimal decision criterion that maximizes expected reward per trial.

Our experimental results suggests that poor detection rates in rare target searches (under the Neutral reward scheme) are not due to subject's fatigue (Parasuraman, 1986), carelessness, or lack of vigilance (Warm, 1993). They are not due to an attentional lapse (which would

impair stimulus representation), as we find that a model with fixed internal stimulus representation predicts subjects' data accurately by varying the decision criterion only. Another account suggests that rare targets are missed due to motor response errors (Fleck & Mitroff, 2007) arising from repeated "no" responses in rare target searches. However, this account fails to explain why such motor errors are not observed for rare target searches under the Airport reward scheme. A prominent account suggests that rare targets are missed due to premature abandoning of the search (Wolfe et al., 2005). While this account can explain reaction time data in experiments where the display remains until subjects terminate the search, it is not applicable to our study where we do not find any RT differences in response to a brief display that automatically terminates after a short duration (100–150 ms). We offer an alternate account of the poor detection performance in rare target searches (Neutral reward scheme) by suggesting that subjects are deploying the optimal strategy of maximizing their reward, which in rare target searches is dominated by the reward in the more frequent target absent trials. Hence, subjects shift their decision criterion to commit fewer false alarm errors.

Previous studies did not find a noticeable effect of reward on the detection rates (Green & Swets, 1966; Healy & Kubovy, 1981; Kubovy & Healy, 1977; Lee & Janke, 1964, 1965; Lee & Zentall, 1966; Maddox, 2002). They reported suboptimal effect of reward on the decision criterion. In particular, these studies noted that 1) the decision criterion was more conservative than the optimal decision criterion, 2) the decision criterion was closer to optimal when the target frequency changed, and suboptimal when the reward scheme changed. One study (Maddox, 2002) found that reward was less effective than target frequency in shifting the decision criterion, even at 25% target frequency. Another study conjectured that reward structures that would be effective for rare target searches (1–2% frequency) might exceed what is practical in the laboratory (Wolfe et al., 2007). Yet another concern is that the monetary reward in the laboratory setting may not relate to the reward values in the real-world scenario (e.g., getting fired from the job for missing a bomb in airport baggage screening cannot be quantified by an equivalent monetary value in the laboratory; Wolfe et al., 2005). In Experiment 1, we confirmed previous findings that reward payoff manipulation has a suboptimal influence on the decision criterion and does not yield significant improvement in target detection rates. However, a potential flaw in the design of previous studies is the lack of control for how subjects perceived the reward (Kahneman & Tversky, 1979). Note that the reward values r_{ij} $(i, j \in \{0, 1\})$ in Equation 1 may not be perceived by the subjects as the objective reward value (chosen by the experimenter). Non-linearity in subject's utility function such as diminished utility with increasing wealth (von Neumann & Morgenstern, 1953) may result in a gain of 10 points being perceived as less than 10, resulting in weak reward incentives. Thus, previously observed suboptimality (failure to maximize expected reward) may be due to the non-linearity of subject's utility function. We avoided this confound by rewarding subjects in a competitive setup—subjects may lose the contest even after earning 1000 points, hence to win the contest, subjects must try to maximize the number of points earned without slackening their effort or diminishing their utility. Under such competitive settings (Experiment 2), we find that subjects maximize expected reward. Thus, the drop in detection performance as the search targets become rare (Wolfe et al., 2005) can be compensated for by changing the reward scheme (Equation 1 shows that reward and target frequency are multiplied together and therefore should have equivalent effects on observer's behavior).

Our finding that subjects learn the optimal decision criterion is interesting and different from previous reports; however, it is of little practical value if subjects require hundreds of practice trials to learn. Analysis of the training data reveals a surprisingly rapid learning rate-subjects learn to maximize expected reward within an average of 14, 31, 7, 24, and 42 trials in the 50%, 10%, 2% Neutral, and 10%, 2% Airport conditions, respectively. What are the underlying computational and neural mechanisms of such rapid reward-based learning? One possibility is to first estimate the statistics of stimulus representation in the target present and absent displays, then compute the expected reward profile as a function of decision criteria (as in Equation 1) and use gradient ascent (or other rules) to learn the optimal decision criterion that maximizes expected reward. However, such a model would require many trials to learn. Instead, we propose a simple model that only requires knowledge of the sensory observation and ground truth on the previous trial (obtained via feedback on the reward) and uses an incremental learning rule to update the reward score earned so far by operating at different decision criteria. At any instant, it selects the decision criterion with maximum reward score. We show through simulations that this model and its neural implementation learn the decision criterion rapidly within a few trials (intuitively, the time to see at least one target present and absent trial). Such rapid learning raises interesting questions, such as whether subjects can quickly adapt to displays where the target frequency and reward payoffs change dynamically.

A potential application of our finding is the design of reward schemes to improve detection rates of rare targets. We demonstrate this in Figure 10. As the ideal observer model provides a quantitative account of subject's behavior, we can use it to predict subject's detection rates for any reward policy or target frequency, by simply knowing the subject's ROC curve. Thus, the subject's ROC curve may be easily determined from 100 to 200 trials (e.g., using confidence rating procedures (Green & Swets, 1966; Palmer et al., 2000) at 50% target frequency) and used to predict subject's detection rates in 2% target

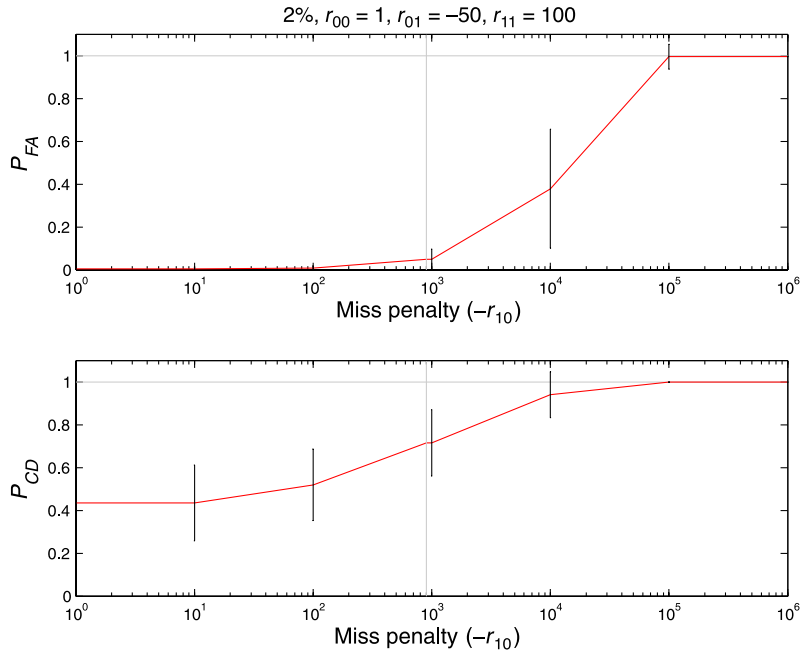

Figure 10. Designing reward schemes to obtain desired detection performance. This plot shows how the penalty r_{10} for missing a rare target (2% frequency) may be changed (while maintaining other reward values constant) to obtain the desired detection performance. We use the ROC curve obtained from Experiments A and B. For each reward policy in the 2% target frequency, we use the ideal observer model to predict the false alarm and correct detection (hit) rates that maximize expected reward per trial. Our Airport reward scheme shown by the vertical gray line ($r_{10} = -900$) yields detection rates of 71 ± 15%. Even higher detection rates may be obtained by increasing the miss penalties further ($r_{10} < -10000$) but at the expense of increasing false alarm rates.

frequency conditions, which would otherwise require 1000 trials (for 20 target present trials).

The current study focuses on a paradigm where the display automatically terminates after a brief exposure (100–200 ms SOA). Under these conditions, we find that subjects in a contest rapidly learn to maximize expected reward. However, several real-world tasks (e.g., visual screening of airline passenger bags for detecting dangerous weapons) involve displays that remain until the subject chooses to terminate the search. Such settings introduce speed-accuracy tradeoffs and interesting reaction time effects on target detection rates (e.g., a recent study (Wolfe et al., 2005) found that rare targets are missed when the search is abandoned faster than the mean time to find the target). It remains to be tested whether subjects maximize expected reward per trial in such reaction time settings.

Many cognitive decisions (e.g., whether to buy a stock or not) are influenced by economic incentives. Such tasks are complex, cognitive, deliberative, and often involve suboptimal decision strategies (Kahneman & Tversky, 2000; e.g., loss aversion, risk aversion). In contrast, our study focuses on simple, non-cognitive, early sensory decisions such as whether we see a salient target or not; under these conditions, we find that humans behave optimally. Our finding that humans maximize expected reward in sensory decision making complements recent findings of optimality in rapid motor planning tasks (Trommershäuser, Maloney, & Landy, 2003). We also find that humans learn the optimal decision criterion rapidly within 50 trials during training and propose neurally plausible models of such rapid reward-based learning. Potential real-life applications of this study include designing reward schemes to improve target detection rates in life-critical searches.

Acknowledgments

This work was supported by grants from the National Geospatial-Intelligence Agency (NGA), the Office of Naval Research (ONR), the National Science Foundation (NSF), and the National Institutes of Health (NIH). We would like to thank John O'Doherty, Shin Shimojo, Preeti Verghese, Jeremy Wolfe, two anonymous reviewers, and the editor for their valuable comments. The authors affirm that the views expressed herein are solely their own and do not represent the views of the United States government or any agency thereof.

Commercial relationships: none.
Corresponding author: Vidhya Navalpakkam.
Email: vidhya@caltech.edu.
Address: Division of Biology, 216-76, California Institute of Technology, Pasadena, CA 91125, USA.

References

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review, 113,* 700–765. [PubMed]

Braun, J. (1994). Visual search among items of different salience: Removal of visual attention mimics a lesion in extrastriate area V4. *Journal of Neuroscience, 14,* 554–567. [PubMed] [Article]

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

Eckstein, M. P., Shimozaki, S. S., & Abbey, C. K. (2002). The footprints of visual attention in the Posner cueing paradigm revealed by classification images. *Journal of Vision, 2*(1):3, 25–45, http://journalofvision.org/2/1/3/, doi:10.1167/2.1.3. [PubMed] [Article]

Fleck, M. S., & Mitroff, S. R. (2007). Rare targets are rarely missed in correctable search. *Psychological Science, 18,* 943–947. [PubMed]

Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annual Review of Neuroscience, 30,* 535–574. [PubMed]

Goldstone, R. L. (1998). Perceptual learning. *Annual Review of Psychology, 49,* 585–612. [PubMed]

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: John Wiley and Sons.

Gur, D., Sumkin, J. H., Rockette, H. E., Ganott, M., Hakim, C., Hardesty, L., et al. (2004). Changes in breast cancer detection and mammography recall rates after the introduction of a computer-aided detection system. *Journal of the National Cancer Institute, 96,* 185–190. [PubMed] [Article]

Healy, A. F., & Kubovy, M. (1981). Probability matching and the formation of conservative decision rules in a numerical analog of signal detection. *Journal of Experimental Psychology: Human Learning and Memory, 7,* 344–354.

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews, Neuroscience, 2,* 194–203. [PubMed]

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47,* 263–292.

Kahneman, D., & Tversky, A. (2000). *Choices, Values, and Frames*. New York: Cambridge University Press.

Kubovy, M., & Healy, A. F. (1977). The decision rule in probabilistic categorization: What it is and how it is learned. *Journal of Experimental Psychology: General, 106,* 427–466.

Lee, W., & Janke, M. (1964). Categorizing externally distributed stimulus samples for three continua. *Journal of Experimental Psychology, 68,* 376–382. [PubMed]

Lee, W., & Janke, M. (1965). Categorizing externally distributed stimulus samples for unequal molar probabilities. *Psychological Reports, 17,* 79–90. [PubMed]

Lee, W., & Zentall, T. R. (1966). Factorial effects in the categorization of externally distributed stimulus samples. *Perception & Psychophysics, 1,* 120–124.

Navalpakkam, V., & Itti, L. (2007). Search goal tunes visual features optimally. *Neuron, 53,* 605–617. [PubMed] [Article]

Nothdurft, H. C. (1992). Feature analysis and the role of similarity in preattentive vision. *Perception & Psychophysics, 52,* 355–375. [PubMed]

Palmer, J., Ames, C. T., & Lindsey, D. T. (1993). Measuring the effect of attention on simple visual search. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 108–130. [PubMed]

Palmer, J., Verghese, P., & Pavel, M. (2000). The psychophysics of visual search. *Vision Research, 40,* 1227–1268. [PubMed]

Parasuraman, R. (1986). *Vigilance, monitoring, and search* (vol. 2). New York: John Wiley and Sons.

Rosenholtz, R. (2001). Search asymmetries? What search asymmetries? *Perception & Psychophysics, 63,* 476–489. [PubMed]

Rubenstein, J. (2001). *Test and evaluation plan: X-ray image screener selection test*. Washington, DC: Office of Aviation Research (No. DOT/FAA/AR-01/47).

Schoonveld, W., Shimozaki, S. S., & Eckstein, M. P. (2007). Optimal observer model of single-fixation oddity search predicts a shallow set-size function. *Journal of Vision, 7*(10):1, 1–16, http://journalofvision.org/7/10/1/, doi:10.1167/7.10.1. [PubMed] [Article]

Sugrue, L. P., Corrado, G. S., & Newsome, W. T. (2005). Choosing the greater of two goods: Neural currencies for valuation and decision making. *Nature Reviews, Neuroscience, 6,* 363–375. [PubMed]

Maddox, W. T. (2002). Toward a unified theory of decision criterion learning in perceptual categorization. *Journal of the Experimental Analysis of Behavior, 78,* 567–595. [PubMed] [Article]

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12,* 97–136. [PubMed]

Trommershäuser, J., Maloney, L. T., & Landy, M. S. (2003). Statistical decision theory and the selection of

rapid, goal-directed movements. *Journal of the Optical Society of America A, Optics, Image Science, and Vision, 20,* 1419–1433. [PubMed]

Verghese, P. (2001). Visual search and attention: A signal detection theory approach. *Neuron, 31,* 523–535. [PubMed] [Article]

Verghese, P., & Stone, L. S. (1995). Combining speed information across space. *Vision Research, 35,* 2811–2823. [PubMed]

Vickery, T. J., King, L. W., & Jiang, Y. (2005). Setting up the target template in visual search. *Journal of Vision, 5*(1):8, 81–92, http://journalofvision.org/5/1/8/, doi:10.1167/5.1.8. [PubMed] [Article]

von Neumann, J., & Morgenstern, O. (1953). *Theory of games and economic behavior*. New York: Princeton University Press.

Warm, J. S. (1993). *Vigilance and target detection*. Washington, DC: National Academy Press.

Wolfe, J. M., Horowitz, T. S., & Kenner, N. M. (2005). Cognitive psychology: Rare items often missed in visual searches. *Nature, 435,* 439–440. [PubMed]

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. [PubMed]