Optimal Visual Search with Highly Heuristic Decision Rules

Anqi Zhang and Wilson S. Geisler

Abstract

Visual search is a fundamental natural task for humans and other animals. We investigated the decision processes humans use when searching briefly presented displays having well-separated potential target-object locations. Performance was compared with the Bayesian-optimal decision process under the assumption that the information from the different potential target locations is statistically independent. Surprisingly, humans performed slightly better than optimal, despite humans' substantial loss of sensitivity in the fovea ("foveal neglect"), and the implausibility of the human brain replicating the optimal computations. We show that three factors can quantitatively explain these seemingly paradoxical results. Most importantly, simple and fixed heuristic decision rules reach near optimal search performance. Secondly, foveal neglect primarily affects only the central potential target location. Finally, spatially correlated neural noise causes search performance to exceed that predicted for independent noise. These findings have far-reaching implications for understanding visual search tasks and other identification tasks in humans and other animals.

Introduction

A fundamental and ubiquitous visual task is searching the local environment for specific targets. Humans and other primates have foveated visual systems with high spatial resolution in the direction of gaze (in the fovea) and rapidly declining resolution away from the direction of gaze (in the periphery). Thus, visual search typically involves a series of fixations. During each fixation, the scene is covertly searched to detect the target and/or possible target locations. A sensible and common research strategy is to begin by studying the mechanisms of covert search and then generalize to search with eye movements (overt search). This study focuses on covert search.

Carefully controlled studies of covert search typically present the stimuli briefly to prevent eye movements and with the potential target locations arrayed at a fixed distance from the fixation location to hold the target visibility at the different locations approximately constant[1–5]. In some studies, the task is simply to indicate whether the target is present or absent. In other studies, the target is always present, and the task is to indicate the location of the target. The covert search task in our study is more complex and realistic: identify whether the target is present or absent, and if present, indicate its location.

Under natural search conditions, the number of potential target locations often varies from one situation to the next and hence varying the number of potential locations is a key experimental manipulation. When there is just a single potential target location, the search task reduces to a

simple identification, discrimination, or detection task. In general, as the number of potential target locations increases, search accuracy and speed decrease. The major scientific questions are what stimulus and neural factors are responsible for the decreases and whether models that incorporate the relevant neural factors can quantitatively predict search performance.

Bayesian statistical decision theory provides a principled computational framework for addressing these questions. Specifically, measurements of detection performance for each potential target location, when the location is known, can be used to predict quantitatively the best possible (ideal) performance in the search task when the location of the target is unknown. These predictions provide the normative benchmark for evaluating the various potential stimulus and neural factors affecting search performance. For example, hypothesized factors that cause the Bayesian observer's performance to fall below the measured human performance can be confidently rejected.

As mentioned above, in many well-controlled covert-search studies the targets are presented briefly, and the potential target locations are at fixed retinal eccentricity. Studies show that when the task is to report whether a single target is present or absent, there is a fairly wide range of conditions where human search accuracy is consistent with an optimal decision rule given statistical independence at the potential target locations[1,2,4,6]. However, fixed eccentricity displays are not representative of natural search, where potential target locations are typically more uniformly distributed across the visual field. Also, a target's visibility typically varies somewhat around a circle at fixed retinal eccentricity[7–10]. Finally, many previous studies have not used backgrounds with dense random noise, which is more typical in natural conditions.

Therefore, we designed covert search tasks in white noise backgrounds for target locations that cover the central 16 deg of the visual field. The stimuli were presented for 250 msec—the typical duration of fixations during overt search[11,12]. Search performance was measured in human observers. To increase interpretability, we also directly measured the detectability of the target at each potential target location when the target location was known to the observer. Furthermore, we carefully interleaved the detection and search sessions to eliminate differences in practice effects for the two kinds of sessions.

The results are surprising. First, all four observers performed the search task slightly better than the Bayesian ideal searcher, given the measured detectability when the target locations were cued and the assumption of statistical independence of the responses at the different locations. Second, the Bayesian ideal searcher takes into account the prior probability of the target being present at each potential location (the "prior map"), as well as the detectability of the target at each potential target location (the "d' map"). It seems implausible that during the experiment the observers could precisely learn the prior map and their own d' map, and then optimally apply this information in making responses. Third, in the search task, all four observers showed a substantial loss of sensitivity in the fovea ("foveal neglect"), a phenomenon we reported in a recent study of covert search in continuous noise backgrounds[13].

We show that three factors can explain our seemingly paradoxical results. First, we show that it is not necessary to know accurately the prior and d' maps. Extremely crude and fixed heuristic decision rules, in combination with local normalization (e.g., luminance and contrast gain control), are sufficient to obtain effectively optimal performance. This finding differs from the assumptions in much of the perceptual decision-making literature concerned with understanding the how the brain takes into account the stimulus reliability and prior probabilities[14–16]. Second, foveal neglect primarily affects only the central target location. Third, spatially correlated noise corresponding to about 45% of the total noise variance is sufficient to predict the slightly supraoptimal performance, even with the foveal neglect and the heuristic decision rules.

There are several important implications of the findings. The near optimality of highly heuristic decision rules promises to greatly simplify the development of a general theory of visual search. For example, even though under natural conditions the actual d' map changes with every fixation during visual search, the central decision mechanisms can assume a nearly fixed simple d' map and still achieve effectively optimal performance. These simple decision mechanisms are what evolution is likely to select. Furthermore, the wide range of heuristic decision rules that achieve near optimal overall performance allows the possibility of having a wide range of individual differences in decision rules. We argue in the discussion that this fact is something that could be capitalized upon by natural selection, especially in social species like humans and non-human primates. Another implication is that the highly controlled conditions of the current study provide even stronger evidence for the phenomenon of foveal neglect than our previous study[13]. Finally, our results reveal a new potential effect of correlated neural noise on human performance.

Results

Carefully controlled experimental conditions were used to compare performance in covert search with that in simple detection where the potential location of the target is known. The stimulus presentation sequence for both the search and detection experiments is shown in Figure 1a. On each trial, the observer fixated the central circle in a cueing display of faint circles showing the 19 possible target locations. After a brief period of blank screen, the stimulus display was presented for 250 msec, followed by a response period of up to 3000 msec, during which the cueing display was again presented. In the simple detection experiment, the potential location of the target was cued with a low contrast dark circle, with light circles at all other locations. Observers either indicated target absent with a right mouse click or indicated the estimated location of the target with a left mouse click.

All 19 potential target locations contained at 3.5° diameter randomly sampled patch of high contrast (20% rms contrast) white noise. The target, if present, was a small wavelet (6 cycles/deg), windowed to a diameter of 0.8 deg, and added to the center of the noise patch. For each observer, preliminary detection measurements were made at the central display location

(fovea) to determine the target contrast giving a bias-corrected accuracy of approximately 95% correct (d' = 4.5). This target contrast remained fixed for all search and detection trials. Alternating sessions of search and detection were carried out multiple times in a counterbalanced order across observers.
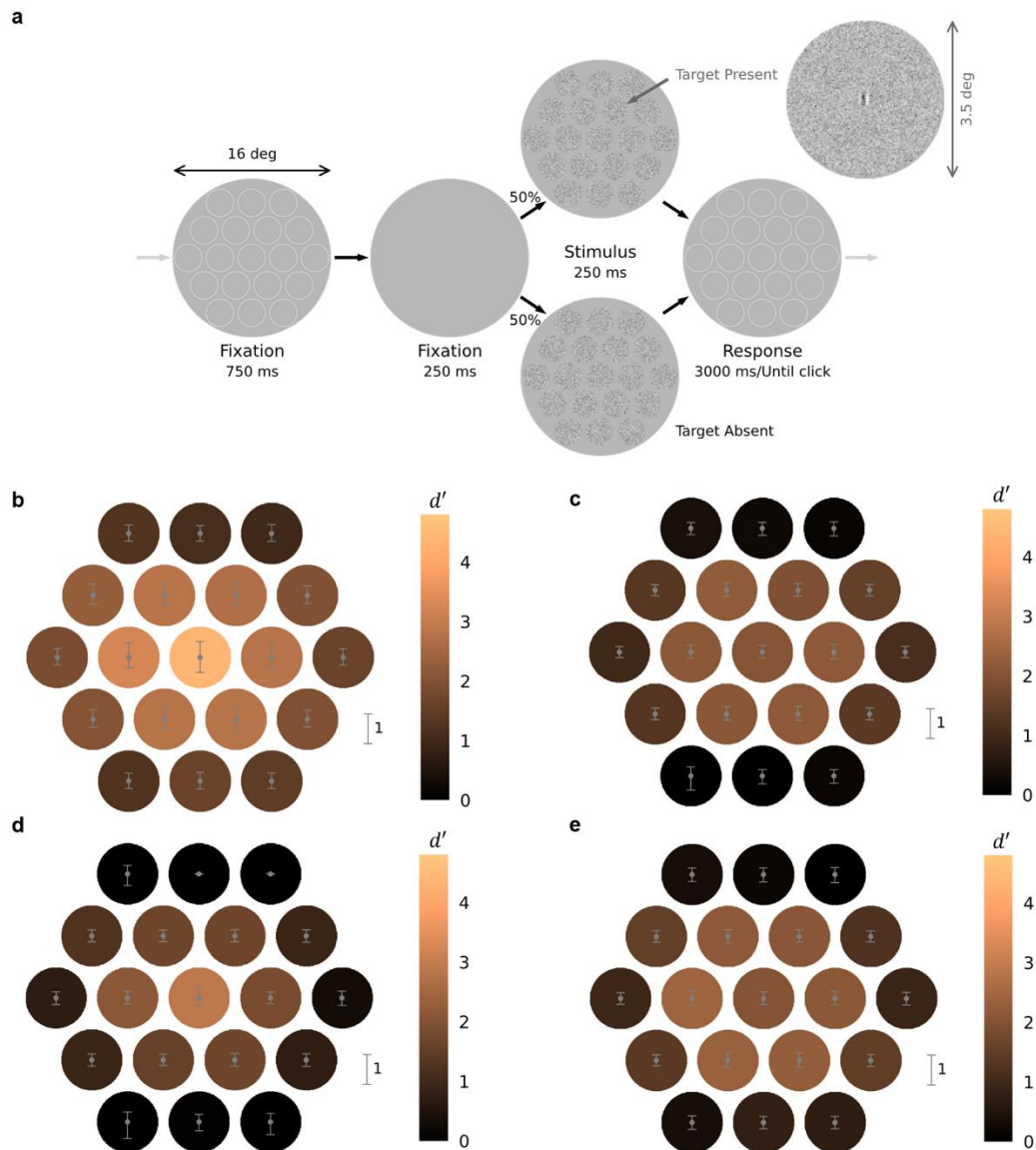

Figure 1 Tasks and performance. **a**. Detection and search stimuli and timeline in a trial. **b**. Human detectability (d') map for the detection task. **c**. Human d' map for the search task. **d**. Ideal searcher d' map. **e**. Heuristic searcher d' map given correlated noise and foveal neglect.

The average target detectability (d') at each of the 19 display locations (the d' map) is shown in Figure 1b. The average d' across all locations is 2.17, and the over proportion correct is 84.2%. While there are some individual differences in these maps across the four observers,

they show the same qualitative pattern: highest detectability in the fovea, intermediate at the 6 locations nearest the fovea, and poorest in the remaining 12 locations, with relatively lower detectability in the upper and lower visual fields (Figure S1a-d). This qualitative pattern is consistent with previous studies[7–10].

The average detectability in the covert search task is shown in Figure 1c. Here, the detectability was computed from the hit rate at each target location and the overall correct rejection rate. The overall detectability is 1.22, corresponding to an overall accuracy of 69.8%, which is considerably lower than that in the detection task. This pattern is seen in all four observers (Table S1). Although there is a falloff in d' with eccentricity, the d' within the central 7 locations is relatively constant. This pattern holds for individual observers (Figure S1e-h).

The gray bars in Figure 2 show the pattern of correct responses and errors across retinal location averaged over all four human observers. Note that the false hit rate (FH) is the proportion of trials where the observer reported a location different from the target location.
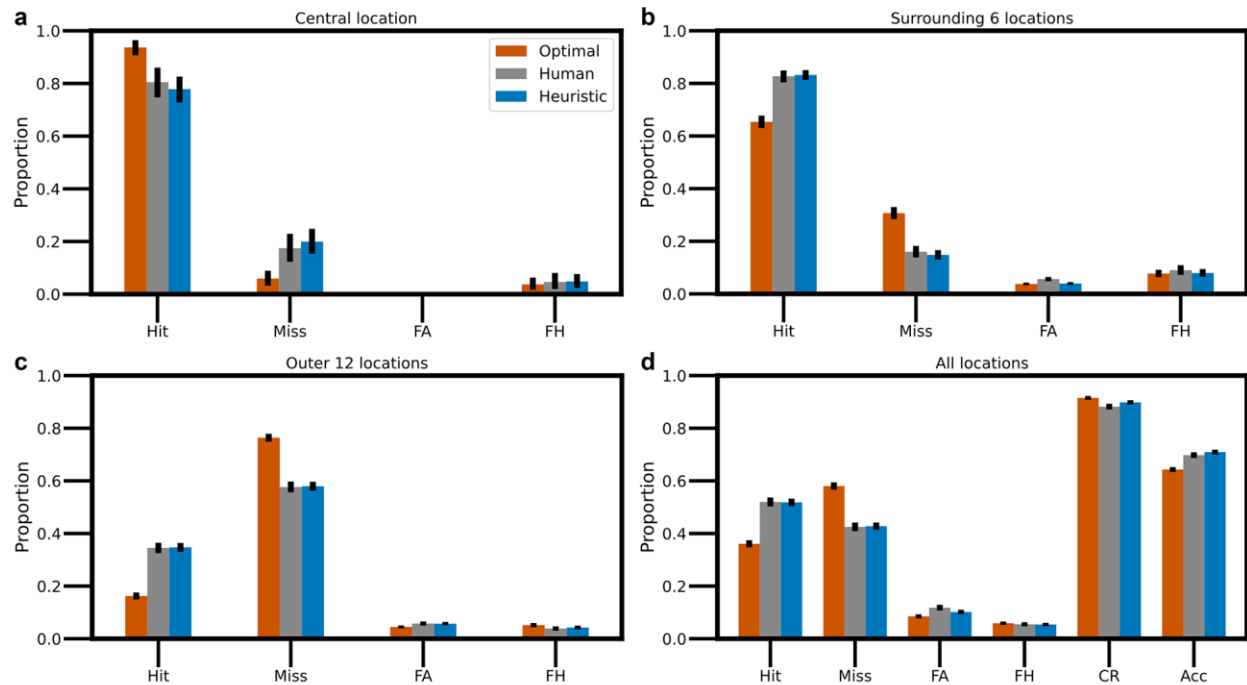

Figure 2 Correct and incorrect responses in the search task by retinal eccentricity. **a**. Histograms of hits, misses, false alarms (FA) and false hits (FH) in the central location, for the average observer (gray), the optimal observer assuming statistical independence of responses across the 19 locations (orange), and a heuristic model observer (blue). **b**. Histograms for the surrounding six locations. **c**. Histograms for the outer twelve locations. **d**. Histograms for all locations; correct rejections and overall accuracy are also included. (Error bars are bootstrapped 95% confidence intervals.) The fit of the heuristic model observer is much better than the ideal observer for the combined human observer and for all the individual human observers (see Supplementary information). The heuristic fit: log-likelihood = -11758, AIC = 23529. The ideal searcher fit: log-likelihood = -12039, AIC = 24078. The fall-off model is $e^{274.5}$ times as probable as the ideal searcher.

To understand the relationship between the performance in covert search and in the detection tasks, it is useful to consider what to expect from a covert searcher that uses the Bayes optimal decision rule. Given that we have directly measured the cued detectability of the target at each potential target location, this modeling analysis can be done within the standard signal detection theory framework[17]. We have shown[13,18,19] that when the receptive field response is normally distributed, the Bayes optimal decision rule is given by

$$\hat{k} = \arg\max_{k \in [0,n]} \left(\ln p_k + d'_k R'_k - 0.5 d'^2_k \right) \qquad (1)$$

where n is the number of potential target locations, $\hat{k}$ is the estimated target location, p_k is the prior probability that the target is at location k, d'_k is the detectability of the target at location k in the cued detection task, and R'_k is the normalized response on that trial at location k. In the standard signal detection framework, R'_k is a random sample from a Gaussian distribution with a standard deviation of 1.0, and a mean of d'_k when the target is present and a mean of 0.0 when the target is absent. In the signal detection framework, R'_k represents the normalized log likelihood ratio of target present versus absent and can be thought of as the observer's decision variable in the cued detection task. Note that $k = 0$ represents the event that the target is absent.

Figure 1d and the orange bars in Figure 2 show covert search performance using the optimal decision rule, given the average measured d' map shown in Figure 1b, the uniform prior probabilities used in the experiment, and the assumption of statistical independence of the responses from the potential target locations. Statistical independence is plausible because the targets were small, the potential target locations were well separated, and the random noise backgrounds were statistically independent. While there is some general qualitative agreement between the ideal searcher and the average human searcher, there are several puzzling facts.

First, it is implausible that human observers implement calculations equivalent to the Bayes optimal decision rule. The optimal decision rule requires weighting the response at each potential target location by the detectability of the target at that location and adding the log of the prior probability of the target appearing at that location (see Equation 1). Learning all 19 detectabilities and 19 priors during the experiment seems unlikely. Worse yet, under natural conditions the d' map is different on every fixation, even for the same target, because the masking properties of the background are different on every fixation. Also, the prior probability map varies depending on the scene context. For the optimal decision rule to be implemented under natural conditions, the human visual system would need extremely sophisticated neural mechanisms to estimate in parallel, during each fixation, the d' map over the visual field for any desired target. Similarly, sophisticated mechanisms would be required to estimate the prior map from the current scene context. Thus, it is highly likely that the observers are using

heuristic decision rules, making it highly unlikely that they can reach the performance predicted by the optimal decision rule.

Second, the overall accuracy of the average human observer is, in fact, slightly higher than that predicted by the optimal decision rule (see Figure 2d). This is also true for the individual human observers and for different numbers of potential target locations (Table S1).

Finally, although the overall accuracy of the human observers is higher than predicted by the optimal decision rule, their performance is suboptimal in the central location (compare Figures 1c and 1d and see Figure 2a). This result holds for the four observers (Figures S1 and S6). This phenomenon, which we call "foveal neglect", replicates an earlier study measuring covert search performance in continuous noise backgrounds[13]. In principle, foveal neglect should guarantee that human search performance falls below that predicted by the optimal decision rule.

What possible explanation is there for these seemingly paradoxical results? We argue that there are three factors that together could explain the results and that have broad implications for understanding and predicting search performance under natural conditions. First, and most important, a wide range of extremely simple heuristic decision rules can achieve optimal overall performance. Second, correlated neural noise would cause the measured d' values in the detection task to be an underestimate of the effective d' values in the search task. Third, foveal neglect primarily affects only the central location out of the 19 locations.

Consider first the effect of heuristic decision rules. A principled approach for evaluating heuristic decision rules is to quantitatively compare their performance with that of the Bayes optimal decision rule. To do this we simulated optimal and heuristic covert search performance for a wide range of possible d' maps. Roughly speaking, detectability falls off with retinal eccentricity according to a two parameter function, where one parameter is the peak value of d' in the center of the fovea, d'_{max}, and the other is the eccentricity at which d' falls by a factor of 2 from the maximum, e_2 (see Equation 6 in Methods). For example, the average d' map in Figure 1b is best fit with a d'_{max} of 4.69 and an e_2 of 4.68. The colored curves in Figure 3a show cross sections of d' maps with a d'_{max} of 6.0, for e_2 values ranging from 1° to 9°. This range represents, approximately, the falloffs for wavelet targets from about 1 cycles/deg to 16 cycles/deg[20,21], a range of spatial frequency that covers most of the area under the human contrast sensitivity function (CSF), and hence probably covers most search targets. Fine targets have small values of e_2, coarse targets have large values of e_2.

The solid black line in Figure 3b shows the overall accuracy in the covert search task of the Bayes ideal searcher for values of e_2 ranging from 1° to 9°, when the prior probability of target absent is 0.5 and the prior probability over all 19 target locations is uniform. The solid gray line shows the performance of the idea searcher when the prior probability of target absent is 0.0 (i.e., a target is always present somewhere), and the prior probability over all 19 target locations

is uniform. The ideal searcher always uses the exact d' map for the target (see Equation 1). As expected, the accuracy of the ideal searcher increases substantially as the value of e_2 increases.

One of the simplest possible heuristic decision rules is to assume a completely flat d' map: the horizontal dotted line in Figure 3a. The dotted curves in Figure 3b show the performance of a heuristic searcher that uses this single fixed flat d' map. As can be seen, performance is nearly identical to that of the optimal decision rules. A slightly more complex but still easy to implement heuristic is a d' map with a fixed peak and fixed falloff: the dashed curve in Figure 3a. The dashed curves in Figure 3b show the performance of this heuristic searcher, which is also near optimal.
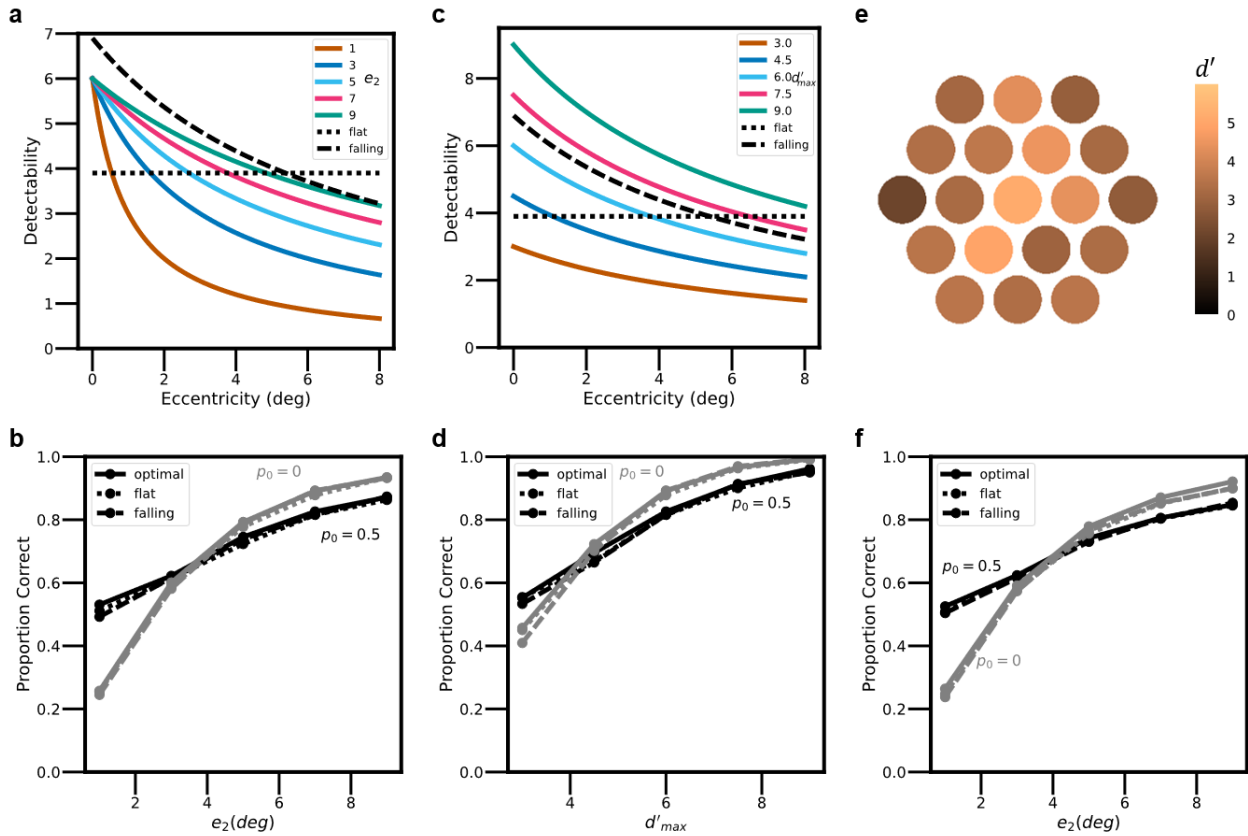

Figure 3 Optimal versus heuristic decision rules. **a**. Actual maps for a range of e_2 values with a fixed d'_{max} of 6.0 (colored curves). A flat heuristic map (dotted line). A heuristic map with the same fall-off as the best fit heuristic to the average human observer (dashed curve). **b**. Accuracy for ideal and heuristic searchers, for conditions in **a.** and target absence rate (p_0) of 0.5 and 0. **c**. Actual maps for a range of d'_{max} values with a fixed e_2 of 7.0 (colored curves). The same flat heuristic map as in **a** (dotted line). The same heuristic map with a fall-off as in **a** (dashed curve). **d**. Accuracy for ideal and heuristic searchers for conditions in **c** and target absence rate (p_0) of 0.5 and 0. **e**. An example of the d' map varying randomly per trial. The baseline d' map has a d'_{max} of 6.0 and a e_2 of 7.0. **f**. Accuracy for ideal and heuristic searchers for the baseline d' map with a d'_{max} of 6.0 and e_2 ranging from 1° to 9°.

These results strongly suggest that the human visual system uses a highly heuristic rule, because there would be little or no benefit in implementing the much more complex optimal rule. If the

visual system does use a heuristic rule, can that rule be estimated from behavioral data even though many heuristics give near optimal overall performance? The answer appears to be at least a partial yes. The dashed curve in Figure 3a is, in fact, the falloff parameter ($e_2 = 7°$) that best fits the pattern of corrects and errors. The blue histograms in Figure 2 show the predictions for this falloff parameter.

We also tested the same heuristic rules for d'_{max} varying from 3.0 to 9.0, while holding e_2 at 7.0 (see Figure 3c). Figures 3d shows that again both heuristic decision rules achieve near optimal performance in all those conditions.

Under natural conditions, the properties of the background scene vary substantial over space and hence the d' map is generally different with every new fixation. We simulated this situation by having a baseline d' map with various e_2 parameter values and then assumed that the random background properties cause the actual d' value at each location to be scaled up or down by a random percentage that is normally distributed, with a standard deviation of 20%. Figure 3e shows a single example of a random d' map, where the baseline e_2 is 7.0 and the baseline d'_{max} is 6.0. Figure 3f shows that the heuristic decision rules in Figures 3a achieve near optimal performance even when the d' map varies randomly on each trial.

Similar results to those in Figure 3 hold for all combinations of e_2 and d'_{max} (Figure S7), and for prior probability maps (Figure S9). Indeed, very simple heuristics are often near optimal under a very wide range of natural conditions. Nevertheless, heuristic rules with extreme deviations from the actual d' and prior-probability maps are detrimental to search performance. For example, if there are substantial search regions where the d' or prior probability go to zero, then no weight should be assigned to those locations in the decision. Similarly, performance can suffer substantially if regions where the d' and prior probability are non-zero are given no weight in the decision.

Next, consider the effects of correlated neural noise. Besides statistically independent sources of noise at each potential target location, we assume the existence of noise with common sources that are added to the responses at all potential target locations. These common sources cause the total noise at the different potential target locations to be partially correlated. For simplicity, we assume that the independent and common noise sources are Gaussian and statistically independent of each other, with standard deviations of σ and σ_0, respectively. Thus, the total noise variance at each target location is $\sigma^2 + \sigma_0^2$. In the detection task, the correlated noise component necessarily lowers the detectability. However, the common noise has little or no effect on the optimal decision rule in the search task. For example, for the flat heuristic rule, the effect of the correlated noise on d' is cancelled out by the max rule, because the correlated noise causes the same increase or decrease of responses equally at all potential target locations. Also, we note that the optimality of the max rule still holds even when the

response is correlated across locations[19]. We show in the Methods that the effective d' map for the search task compared to the detection task is scaled up by a factor of $\sqrt{\sigma^2 + \sigma_0^2}\big/\sigma$. A correlated noise that is approximately 45% of the total noise variance explains the better search performance of human observers compared to the optimal searcher using the d' map from the detection task (see below).

Finally, consider the phenomenon of foveal neglect. Even though humans perform slightly better than the optimal searcher, they have worse search performance in the foveal region (Figures 1c and 2a). In a previous study[13], we showed that this reduction in accuracy in the fovea is not due to bias in estimating the prior probability, but to a reduction in detectability in the fovea. We also showed that given a fixed total amount of attentional gain resources, foveal neglect is a principled phenomenon. The argument is that if there were limited attentional gain resources, they would be applied to some fixed number of cortical neurons, not to some fixed number of image pixels. Because of foveation, there are fewer neurons encoding the periphery and they have larger receptive fields. Thus, to maximize search performance over large regions of space it is better to assign less attentional gain to the dense, small-receptive-field neurons encoding the image near the line of sight.

To quantitatively evaluate this hypothesis, we assumed a fixed total amount of attention gain, distributed over a retinotopic map based on the measured anatomical density of the ganglion cells in the human retina[22,23]. This retinotopic map is represented by the grid of lines in Figure 4a and is similar to the retinotopic map of primary visual cortex[24]. The color shading in Figure 4a and the curve in Figure 4b show the smooth modulation of attentional gain with retinal eccentricity (see Equation 3 in Methods) that best explains the data in Figures 1c and 2. As can be seen, attention gain is lowest in the fovea and increases to a maximum in the periphery. Nominally, the attention gain cannot exceed 1.0, because under the optimal attention distribution hypothesis d' cannot exceed its value when the target location is cued. However, correlated noise allows the gain to exceed 1.0 in our search task. The amount above 1.0 provides an estimate of the fraction of the total noise variance that is due to correlated noise (see Equation 4 in Methods). In the present case, the correlated noise is estimated to be approximately 45% of the total noise variance. Similar estimates of foveal neglect and correlated noise were obtained for all four observers, although there are individual differences (see Figure S6).

In sum, the combination of the three factors described here provides a plausible, quantitatively accurate explanation of the seeming paradoxical detection and search results shown in Figures 1 and 2. The quantitative predictions of the best fitting heuristic searcher with these three factors are much more consistent with the human search behavior than the ideal searcher without these three factors.
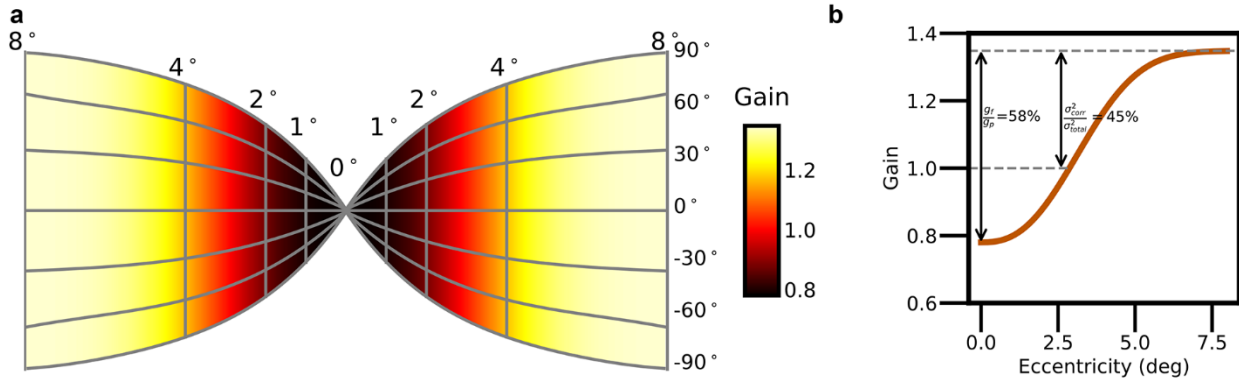
Figure 4. Foveal neglect and correlated noise. **a**. Retinotopic map. The flattened cortical sheet has a constant density of neurons. The grid of contours shows the retinal locations of the cortical neurons' receptive fields. **b**. The estimated variation in attentional gain with retinal eccentricity (also shown by coloring of map in **a**).

Discussion

Covert search and cued detection performance were measured for wavelet targets in Gaussian noise under carefully controlled conditions. The detectability (d') map measured in the cued detection task was used to predict covert search performance for the Bayes optimal decision rule, assuming statistically independent sensory responses from the potential target locations. We found that human performance slightly exceeded the predictions of the Bayes optimal decision rule, despite the complexity of the optimal rule, and despite the fact the humans showed a loss of sensitivity in the fovea (foveal neglect). We found that these seeming paradoxical results can be explained quantitatively by three facts: (1) very simple heuristic decision rules, together with local normalization, can achieve essentially optimal performance, (2) correlated neural noise causes the d' values measured in the detection task to underestimate the effective d' values in the search task, and (3) foveal neglect only reduces the effective d' value at the central target location.

An obvious question is how general these findings are. We have also made some measurements of covert search for 7, 61 and 91 potential target locations (Figure S4, Table S1). The results are generally consistent with those for 19 locations shown in Figures 1 and 2. However, we are less confident about the model predictions for 61 and 91 locations, because the d' maps needed to be interpolated and extrapolated from the 19 locations. Importantly, we find that the heuristic decision rules work equally well for 7, 19, 61, and 91 locations. Thus, it seems safe to conclude that the findings will hold over a wide range of potential target locations.

We chose a 6-cpd wavelet target because the detectability of this target varies substantially over the 16° diameter search region. It would be informative to repeat the measurements for other targets, but given what is known from the detection and search literature, and the fact the simple heuristic decision rules are near optimal for a wide range of d' maps, it is likely that the findings will generalize well across a wide range of spatially-localized targets.

We chose white-noise backgrounds because they have a dense naturalistic structure yet are statistically simple, and because they have been widely used in studies of visual performance. Natural backgrounds, on the other hand, are statistically complex and non-stationary, so that the masking properties of the background typically vary across potential target locations. Thus, the d' map on each trial (e.g., eye fixation) is due to a combination of foveation and variation in the masking properties of the background. The fact that simple fixed heuristic decision rules are effectively optimal even when the d' map changes randomly on each trial (Figures 3e–f, and Figure S7) suggests that our findings may hold across a wide range of stationary and non-stationary backgrounds. The difficulty in directly testing this hypothesis is that to calculate ideal performance one must know the d' map for the stimulus on each trial. Estimating the d' map is tractable for white noise backgrounds with luminance and contrast varying across spatial locations, because of the lawfulness of the effects of noise contrast and luminance [25–27], but is more problematic for natural backgrounds. However, progress has been made in modeling and predicting detectability in arbitrary natural backgrounds[18,26,28–30]. Given the current study, it is possible that such models will be able to predict human covert search performance in natural backgrounds by including simple heuristic decision rules, correlated noise, and foveal neglect.

Our theoretical analysis and modeling make the standard signal-detection-theory assumption that the observer's decision variable in the detection task is normally distributed. The psychophysical literature shows that this simplifying assumption is robust. Also, in previous work we have shown that linear receptive field responses to natural backgrounds are approximately normally distributed if the responses are normalized by the values of the background properties (e.g., luminance and contrast) that are known to have a substantial masking effect[26]; this finding is also consistent with earlier related work[31].

Where humans tend to fall below the predictions of the optimal decision rule is when the task is to identify the locations of multiple targets[32], identify the target location when the target may be drawn many different categories of target or, more generally, when there are substantial demands on memory or on high-level cognitive computation (e.g., which location contains a number divisible by 13). Theories of covert search for such conditions need to include limitations in memory and cognitive computation. Nonetheless, there are many real-world situations where observers are covertly searching for targets that require low cognitive effort.

Our findings for the covert search task are likely to generalize to many other identification tasks. Most identification tasks can be described as choices between mutually exclusive events. For such tasks the max rule remains optimal or near optimal. Thus, a Bayesian heuristic decision analysis like the one described here may provide useful insights and testable predictions.

Normalization is a fundamental and ubiquitous property of cortical processing[33–35]. The present results suggest that it may play a more important role in perceptual decision making than previously appreciated. For detection tasks in natural backgrounds, it has been shown that normalization by local luminance, contrast, and similarity allows near-optimal decisions with a

single fixed decision criterion[26]. The present results greatly expand this conclusion by showing that such normalization should also allow extremely simple heuristic decision rules to achieve near optimal performance for a wide range of natural identification tasks. Without normalization, the heuristics described here do not work nearly as well on natural and other non-stationary backgrounds. For example, if the detectability at different locations varies because of differences in the contrast of the white noise, then simple fixed heuristics without normalization give substantially poorer performance (see Figure S8).

The existence of simple fixed decision rules that are near optimal increases the likelihood of developing accurate quantitative theories of covert search under natural conditions. It is relatively easier for evolution to find such decision rules, so there is higher likelihood that near-optimal rules are implemented in real visual systems. Also, the simplicity of these near-optimal rules should make it easier to uncover the neural circuits that implement them.

An important implication of our heuristic decision analysis is that there can be substantial individual differences in covert-search decision rules with no effect on overall search accuracy. In other words, seeing individual differences in the patterns of correct responses and errors, or in the trial-by-trial responses to the same stimuli, does not imply that one of the observers has a better search strategy.

Another implication is that the neural circuitry for optimal covert-search decisions is likely to be relatively simple. Thus, animals with relatively few neural resources should be able to implement the same optimal decision rules as animals with many more neural resources.

Finally, the fact that many simple fixed heuristics perform optimally may have important implications for the evolution of decision processes. Specifically, natural selection may cause an increase in individual differences in covert-search decision rules. This may be especially true in social species, because bigger individual differences may reduce competition in foraging and increase the probability that the group detects predators, prey, and other targets of value to the group. At the same time, this increase in individual differences would not be detrimental to the individual when away from the group.

The optimal Bayesian decision rule for covert search also takes into account the prior probability map (see Equation 1). In the current experiments, the prior probability of target absent was 0.5 and prior probability across the potential target locations was uniform. Given the simplicity of this prior map, we did not focus on the effect of assuming simple heuristics. However, in the supplementary material (Figure S9) we report some initial simulations of simple heuristics (uniform priors) when the actual prior maps are not uniform. We find that it is important to know the target-absent prior probability reasonably well in order to approach optimal performance. However, as with the d' maps, a uniform target-present prior map achieves near optimal overall performance for a wide range of actual prior maps.

The classic view in the Bayesian decision making literature is that taking into account reliabilities and prior probabilities is critical for making decisions that optimize accuracy[14–16]. Our simulations and experimental results show that this is not true for covert visual search tasks and probably many other identification tasks[19]. While it is important to take into account the regions of stimulus space where reliabilities and priors go essentially to zero, it is not important to take into account the specific pattern of values within the non-zero regions. This finding depends on the assumption of parallel processing, which holds when applying the max rule. Parallel processing also holds for most deep neural networks (DNNs) and hence DNNs also benefit from the fact that they can achieve near optimal overall performance without taking into account detailed variations in prior probability and reliability. It follows that DNNs can differ in their responses to specific stimuli even when their overall performance is similar and near optimal. If processing is more limited and serial, then effective decision rules need to take into account detailed priors and reliabilities. For example, in a covert search task with a 50% prior of target absent, it would be a mistake to spend a limited amount of processing time on a few locations with low prior probability and/or low reliability.

The max rule is the optimal decision rule not just for covert search, but for many identification tasks. For these tasks, correlated noise (across all the category responses) has a negligible or minor effect on the accuracy of the decision. We showed that correlated noise can create a mismatch between the d' values estimated in the cued detection task and the effective d' values in the covert search task. Such mismatches might occur in other identification tasks, and thus would be an important factor to consider.

Our results do not prove that correlated noise is the source of the supraoptimal accuracy of the observers in our experiments, but it is a plausible hypothesis because there are many likely sources of correlated noise, such as slow modulations in membrane potential and in more indirect neural-response measures such as the bold response [36–40].

An interesting possibility is that the nervous system evolved to inject correlated variations into the pathways transmitting information to the circuits or brain areas that perform identification tasks. Because these correlated variations do not hurt identification performance, they could provide an independent (low bitrate) channel for communicating other kinds of information such as reward signals, arousal signals, task-relevant global context information, etc. Such correlated variations could be used to broadcast these other kinds of information to any of the decision-making circuits that use the max rule. Piggybacking on existing major pathways could also reduce the need for separate specialized minor pathways. The benefits of this low bit rate communication channel may outweigh the cost of reduced sensitivity in cued detection tasks. A recent study provides evidence for the related hypothesis that stochastic co-modulation of specific neural populations serves as a label for task relevance in subsequent stages of processing[41].

Visual search under natural conditions is a mixture of covert and overt search. During the first part of each fixation there is a covert search event to identify the location of the target or potential target locations. During the second part of each fixation the next fixation location is computed. The optimal decision rule for picking the next fixation location also takes into account the d' map and the prior probability map[10]. An important next step will be to perform a Bayesian heuristic decision analysis for fixation selection. However, what is immediately clear is that a flat heuristic d' map will be inadequate. Using a flat map would always result in keeping fixation the same, rather than wasting the saccade time during which no useful covert-search information is obtained. If the human visual system uses the same d' map for covert search and fixation selection it could explain why our estimate of the average human heuristic for covert search is not flat, even though the flat heuristic is as good as the foveated heuristic (see Figures 3 and S2).

Methods

The study included four male observers, aged 19–26. They all had normal or corrected-to-normal acuity. All procedures in the experiments were approved by the University of Texas Institutional Review Board.

The stimuli in the experiments were generated with MATLAB 2023a and the Psychophysics Toolbox[42,43]. They had a resolution of 30 pixels per visual degree, with each pixel occupying a 2 x 2 screen pixel region on a calibrated Sony GDM-FW900 cathode-ray-tube (CRT) monitor, with a size of 1920 x 1200 pixels, a refresh rate of 85 Hz, and a bit depth of 8. The stimuli were gamma-compressed prior to display on the screen.

Circular cues were used to indicate possible target locations. The luminance outside the circular cues was set to the mean luminance of stimuli within the cues (60 cd/m^2). Light circular cues had a luminance of 66 cd/m^2 . Dark circular cues had a luminance of 51 cd/m^2 . Backgrounds centered on each potential target location were statistically independent samples of high-contrast Gaussian noise. Each patch had a root-mean-squared (RMS) contrast of 20%. For the 19-location configuration, the target locations were separated by 4 deg, and the backgrounds had a diameter of 3.5 deg. The target was a 6-cycles/deg wavelet windowed to a diameter of 0.8 deg that, when present, was added to the center of the noise patch.

For the detection task, the observers were asked to indicate target absence with a right mouse click, and target presence with a left mouse click. For search tasks, observers were asked to right click if the target was absent and to left click the judged location if present. We ran preliminary search trials with highly visible targets and found that the human observers made no errors in clicking on the target location, indicating that memory and motor-control limitations were not important factors in the study. Furthermore, we used a reverse counterbalancing design, where the observers completed the detection and search experiments in two opposite orders.

The allocation of attentional sensitivity gain was modeled as in the efficiency-limited-foveated (ELF) searcher and the retinotopic map definitions in [13]. The attentional efficiency gain $g(\mathbf{x})$ was applied to obtain a modified d' map:

$$d'_g(\mathbf{x}) = g(\mathbf{x})d'(\mathbf{x}) \quad (2)$$

A retinotopic map is shown in Figure 4 and supplementary Figure S6a. Figure S6a illustrates the cortical area corresponding to each of the 19 background regions. The attentional sensitivity gain map was modeled as a Weibull function

$$g(\mathbf{i}) = g_f + (g_p - g_f)\left[1 - e^{-(i/a)^b}\right] \quad (3)$$

where $\mathbf{i} = (i, j)$ is a coordinate in V1, i represents the coordinate along the horizontal meridian, g_f, g_p are the gains in fovea and periphery, respectively, and a, b are the steepness and shape parameters of the gain modulation, respectively (for more details see [13]).

Spatially correlated Internal noise is another component of our model. We assume that the independent noise variance at each location is σ^2, the correlated noise added all the locations is σ_0^2, that the independent and correlated noise sources are independent of each other. Thus, in the detection experiment the detectability is $d'_d = a\big/\sqrt{\sigma_0^2 + \sigma^2}$ and the detectability in the search task is effectively $d'_s = a/\sigma$, because the correlated noise is cancelled by applying the max rule.

Simulations show that, for moderate levels of correlated noise, if one applies the max rule in the cued detection task by comparing the cued location to another location having equal detectability there is no improvement in performance. This occurs because the comparison turns the task into a same-different discrimination task, which has square-root of two lower detectability than a yes-no task[14].

Because of the correlated noise, the peripheral gain parameter can exceed 1.0. The estimated value of the periphery gain provides an estimate of the proportion of the total variance due to correlated noise:

$$\frac{\sigma_0^2}{\sigma^2 + \sigma_0^2} = 1 - \hat{g}_p^{-2} \quad (4)$$

A heuristic covert searcher is any model that does not make decisions exactly as the ideal searcher. We focus on a family of models that uses simplified (heuristic) priors and detectabilities in max rule:

$$\hat{\mathbf{x}} = \arg\max_{\mathbf{x}} \left\{ \ln \hat{p}(\mathbf{x}) + \hat{d}'(\mathbf{x})\left[R'(\mathbf{x}) - 0.5\hat{d}'(\mathbf{x})\right] \right\} \quad (5)$$

where $\hat{p}(\mathbf{x})$ and $\hat{d}'(\mathbf{x})$ are the heuristic prior and detectability maps.

The simulated d' maps considered here were of the form

$$d' = \frac{d'_{\max} e_2}{e + e_2} \quad (6)$$

Where $d'_{\max}$ is the peak value of d' in the center of the fovea, e is eccentricity ($e = \|\mathbf{x}\|$), and e_2 is the eccentricity at which d' reaches half the value in the center of the fovea.

Data availability

All the data that support the findings of this study are available via GitHub at https://github.com/anqi-j/supraoptimal-covert-search.

Code availability

All codes that analyze the findings of this study are available via GitHub at https://github.com/anqi-j/supraoptimal-covert-search.

References

1. Palmer, J., Ames, C. T. & Lindsey, D. T. Measuring the effect of attention on simple visual search. *J. Exp. Psychol. Hum. Percept. Perform.* **19**, 108–130 (1993).

2. Eckstein, M. P., Thomas, J. P., Palmer, J. & Shimozaki, S. S. A signal detection model predicts the effects of set size on visual search accuracy for feature, conjunction, triple conjunction, and disjunction displays. *Percept. Psychophys.* **62**, 425–451 (2000).

3. Palmer, J. & Davis, E. Visual search and attention: An overview. *Spat. Vis.* **17**, 249–255 (2004).

4. Eckstein, M. P. Visual search: A retrospective. *J. Vis.* **11**, 14–14 (2011).

5. Carrasco, M. Visual attention: The past 25 years. *Vision Res.* **51**, 1484–1525 (2011).

6. Eckstein, M. P. The Lower Visual Search Efficiency for Conjunctions Is Due to Noise and not Serial Attentional Processing. *Psychol. Sci.* **9**, 111–118 (1998).