Computational imperfections in human visual search

Elina Stengård[1] & Ronald van den Berg[1*]

[1] Department of Psychology, University of Uppsala, Uppsala, Sweden

[*] Corresponding author

E-mail: ronald.vandenberg@psyk.uu.se (RVDB)

Abstract

Numerous studies have reported that human behavior on perceptual inference tasks – such as cue combination and visual search – is well accounted for by optimal models. However, others have argued that optimal models are often overly flexible and, therefore, lack explanatory power. In addition, it has been suggested that inference performed by neural systems is inherently noisy, which would preclude optimality in many perception tasks. Here, we reconsider human performance on visual search by devising an approach that strongly reduces model flexibility and tests for suboptimalities due to imprecisions in neural inference. Subjects performed a target detection task in which targets and distractors were ellipses with orientations drawn from Gaussian distributions with different means. We controlled the level of sensory uncertainty through stimulus presentation time (short *vs.* unlimited) and the elongation of the ellipses (low *vs.* high). Moreover, we created four levels of *external* uncertainty by varying the amount of overlap between the target and distractor distributions. Since sensory noise was negligible in the conditions with unlimited display time, we were able to estimate deviations from optimality without having to fit free parameters. In conditions with short display time, we limited the flexibility of the optimal model by using a separate task to estimate sensory noise levels. We found clear evidence for suboptimalities in all tested conditions. Moreover, we estimate that the performance loss due to computational imperfections was of comparable magnitude to the loss due to sensory noise. Our results provide support for the proposal that neural inference is inherently imprecise and challenge previous claims of optimality in perception.

Author summary

The main task of perceptual systems is to create truthful representations of the world. They do so by using sensory information that is often astonishingly imprecise due to measurement errors in our senses. Consequently, it is often impossible to be 100% correct all the time on tasks that involve perception, such as judging whether a visual target is present in a cluttered scene. Observers are typically defined as optimal if they perform as well as theoretically possible given the sensory imprecisions. Numerous studies have reported that humans are optimal observers in perception-based tasks, but the validity of these findings has recently been questioned for two different reasons. First, it has been argued that a lot of the evidence is based on studies that used overly flexible models. Second, there are indications that inference performed by brains is inherently imprecise, due to limitations in the neural systems performing the inference. In this study, we reconsider optimality in perception by devising a research method that makes several improvements over previous studies. We apply this method to a visual search task and find clear indications of suboptimalities. Our findings imply that the perceptual systems may indeed not be as perfect as previously thought.

Introduction

Visual perception is the brain's ability to make inferences about the external world from visual information. It is often reported that human performance on this type of inference is optimal or "Bayesian" [1–5], which would mean that they perform as well as is possible given their sensory noise levels. Evidence for this has mainly come from tasks in which subjects integrate two sensory cues to estimate a common source. The optimal strategy in these tasks is to compute a weighted average of the two cues [6], where each weight is proportional to the cue's reliability[1]. This type of weighting is a hallmark of Bayesian observers and predicts that a subject's

[1] Defined as the inverse variance of the sensory noise distribution.

estimates are biased towards the more reliable cue, which has been confirmed in a large range of studies with both humans and other primates. Examples include integration of a visual and haptic cue to estimate the height of an object [7], a visual and proprioceptive [8] or auditory [9] cue to estimate object location, and two visual cues to estimate object depth [10,11] or object slant [12]. Moreover, it has been reported that optimality in perception extends to tasks with up to at least eight cues and in which the optimal strategy involves non-linear computations, including visual search [13–17], categorization [18], change detection [19], change localization [20], and sameness discrimination [21] tasks.

While these studies have provided valuable insights into basic mechanisms of perception – such as that humans take cue reliability into account when integrating sensory cues – they have also been criticized. One criticism is that the emphasis on optimality has led to an underreporting and underemphasizing of studies that have found violations of optimality [22,23]. Another, more fundamental criticism is that optimal models often lack explanatory power due to being overly flexible [23–26]. The risk of too much flexibility is that it may allow an optimal model to account for data from suboptimal observers. For example, when sensory noise levels are fitted as free parameters – as in most studies – an optimal model may account for suboptimalities in inference by overestimating these noise levels. In fact, several recent studies have argued that computational imprecisions are inherent to any kind of inference performed by a brain, due to factors such as noise in the underlying neural mechanisms, imprecise knowledge of the task statistics, and the use of deterministic approximations to complex computations [27–31]. If this is true, then it would mean that the suboptimalities caused by such imprecisions must somehow have gone unnoticed in previous studies, possibly due to using overly flexible models.

These concerns call for a reconsideration of optimality claims in perception. Here, we contribute to this enterprise by revisiting human performance on visual search, which is one of the most commonly employed tasks in visual psychophysics and a task that has been reported to be performed optimally by humans [13–15]. Importantly, however, we use an approach that strongly limits model flexibility, in two different ways. First, we include experimental conditions in which sensory noise is negligible, such that deviations from optimality can be assessed without the need to fit any free parameters. In these conditions, we present stimuli at high contrast and with unlimited viewing time. Moreover, to avoid the task from becoming trivial, we add external noise to the stimuli. Hence, in these conditions, we essentially replace a latent parameter (sensory noise) with one that is fully under the experimenter's control (external noise). An additional advantage of this approach is that it makes the task more consistent with naturalistic conditions, where inference often involves dealing with both internal and external uncertainty [32]. Second, to reduce model flexibility in conditions *with* sensory noise, we measure subjects' sensory noise levels in a separate task and use these estimates to constrain the parameters of the optimal model. Finally, instead of testing the optimal model against a specific set of suboptimal decision rules – as previous studies have typically done – we test it against a more general model that is able to account for a broad variety of inferential imperfections. To preview our main result, in all tested conditions we find clear evidence for inferential suboptimalities.

Experimental methods

Data and code sharing

The experimental data are available at https://osf.io/dkavj/. Matlab code that reproduces the main results will be made available upon publication.

Subjects

Thirty subjects were recruited via advertisements at the psychology department of Uppsala University in Sweden and received payment in the form of cinema tickets or gift vouchers. All subjects had self-reported normal or corrected-to-normal vision and gave informed consent before the start of the experiment. No subjects were excluded from any of the analyses.

Stimuli

Stimuli were black ellipses (0.35 cd/m^2) with an area of 0.60 deg^2 presented on a gray background (71 cd/m^2; Fig. 1A). The task-relevant feature in all experiments was ellipse orientation, with 0° defined as vertical. The eccentricity (i.e., elongation) of an ellipse determined the level of sensory noise with which observers observed its orientation: the higher its eccentricity, the lower the level of sensory noise (Fig. 1B). Stimuli were generated using the Psychophysics Toolbox [33] for Matlab and presented at fixed locations along an invisible circle at the center of the screen and with a radius of 7 degrees of visual angle.

General procedure

Each subject completed multiple experimental sessions that lasted about one hour each. At the start of the first session, they received general information about the experiment and then performed a discrimination task (Fig. 1A) followed by one condition of the visual search task. In the remaining sessions, they only performed the visual search task (Fig 1C). We created eight conditions for the visual search task by using a 2×4 factorial design (Table 1). The factors specify the stimulus presentation time (short *vs.* unlimited) and the level of external uncertainty (none, 5%, 10%, and 15%; explained below). Different groups of subjects performed different subsets of these conditions.

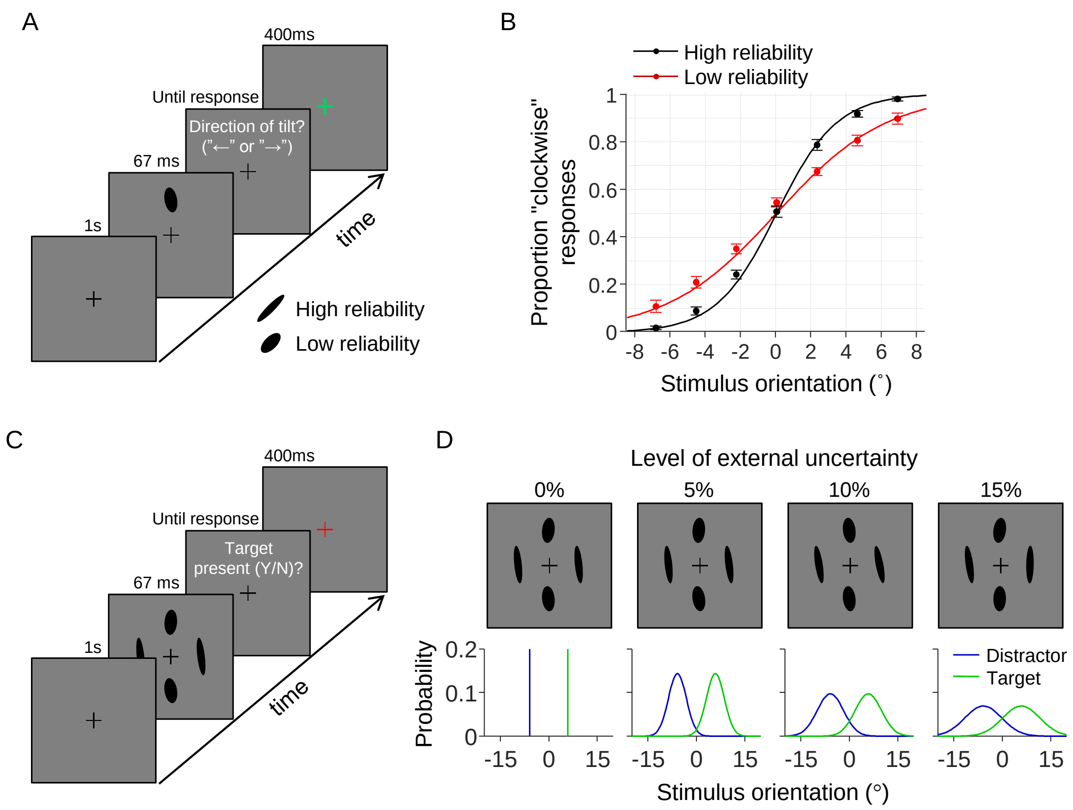

Fig. 1. Experimental design. (A) Illustration of a trial in the discrimination task. A single oriented ellipse was presented and subjects reported its direction of tilt relative to vertical. The elongation of the stimulus could take two values; we refer to the most elongated type of ellipse as a "high reliability" stimulus and the less elongated type as a "low reliability" stimulus. Feedback was provided by briefly turning the fixation cross red (error) or green (correct) after the response was given. (B) The subject-averaged data (filled circles) and model fits (curves) reveal that sensitivity was higher for stimuli with high reliability (black) compared to those with low reliability (red). Error bars represent 1 s.e.m. (C) Illustration of a trial in the visual search task with brief stimulus presentation time. (D) Top: examples of target-present displays under the four different levels of external uncertainty. Bottom: examples of distributions used in the experiment and from which the stimuli in the example displays were drawn. In all four examples, the top stimulus is a target and the other three are distractors.

Table 1. Overview of visual search task conditions and experimental subject groups. Each group consisted of 10 subjects. The condition with unlimited stimulus time and no external uncertainty was excluded from the experiment, because subjects are expected to perform 100% correct on it.

		Level of external uncertainty			
		None	**5%**	**10%**	**15%**
Stimulus display time	**Short (67 ms)**	A Group 1	B Group 2	C Group 1	D Group 3
	Unlimited	-	E Group 2	F Group 1	G Group 3

Discrimination task

On each trial, the subject was presented with a single ellipse (67 ms) and reported whether it was tilted clockwise or counterclockwise with respect to vertical (Fig. 1A). Trial-to-trial feedback was provided by briefly turning the fixation cross in the inter-trial screen green (correct) or red (incorrect). The eccentricity of the stimulus was 0.80 on half of the trials ("low reliability") and 0.94 on the other half ("high reliability"), randomly intermixed. On the first 20 trials, the orientation of the stimulus was drawn from a uniform distribution on the range -5° to $+5^\circ$. In the remaining trials, a cumulative Gaussian was fitted to the data collected thus far and the orientation for the next trial was then randomly drawn from the domain corresponding to the 55-95% correct range. This adaptive procedure increased the information obtained from each trial by reducing the number of extremely easy and difficult trials. Subjects completed 500 trials of this task.

Visual search without external uncertainty (condition A)

In this condition, subjects were on each trial presented with four oriented ellipses. On half of the trials, all ellipses were distractors. On the other half, three ellipses were distractors and one was a target. The task was to report whether a target was present. Targets were tilted μ degrees in clockwise direction from vertical and distractors were tilted μ degrees in counterclockwise direction. The value of μ was customized for each subject (Table 2) such that an optimal observer with sensory-noise levels equal to the ones estimated from the subject's discrimination-task data had a predicted accuracy of 85% correct. Stimulus display time was 67 ms and each stimulus was presented with an ellipse eccentricity of either 0.80 ("low reliability") or 0.94 ("high reliability"). On each trial, the number of high-reliability stimuli was drawn from a uniform distribution on integers 0 to 4 and reliability values were then randomly distributed across the four stimuli. Feedback was provided in the same way as in the discrimination task. The task consisted of 1500 trials divided equally over 12 blocks with short forced breaks between blocks.

Table 2. Overview of estimated sensory noise levels in the discrimination task ($\tilde{\sigma}_{\text{low}}$, $\tilde{\sigma}_{\text{high}}$) and the customized experimental parameters (μ, σ_{external}) in the visual search task.

Level of external uncertainty (%)	Subj ID	$\tilde{\sigma}_{\text{low}}$ (°)	$\tilde{\sigma}_{\text{high}}$ (°)	μ (°)	$\sigma_{external}$ (°)
0	1	7.1	4.6	8.0	0
0	2	5.5	2.0	5.0	0
0	3	6.4	2.3	5.8	0
0	4	3.8	1.8	3.8	0
0	5	4.6	2.2	4.5	0
0	6	4.0	3.3	5.0	0
0	7	3.1	1.3	2.9	0
0	8	6.8	2.8	6.4	0
0	9	3.3	2.4	3.9	0
0	10	4.5	3.0	5.1	0
5	11	4.4	3.4	5.3	2.4
5	12	3.7	3.0	4.5	2.1
5	13	4.2	2.6	4.6	2.1
5	14	3.9	2.2	4.1	1.9
5	15	4.3	3.1	4.9	2.3
5	16	6.5	2.4	5.9	2.8
5	17	6.2	3.3	6.4	2.9
5	18	3.8	2.0	3.9	1.8
5	19	5.1	3.0	5.4	2.5
5	20	4.6	2.6	4.8	2.2
10	1	7.1	4.6	8.0	5.6
10	2	5.5	2.0	5.0	3.4
10	3	6.4	2.3	5.8	4.1
10	4	3.8	1.8	3.8	2.6
10	5	4.6	2.2	4.5	3.1
10	6	4.0	3.3	5.0	3.4
10	7	3.1	1.3	2.9	2.1
10	8	6.8	2.8	6.4	4.3
10	9	3.3	2.4	3.9	2.7
10	10	4.5	3.0	5.1	3.4
15	21	6.2	2.6	5.9	5.8
15	22	7.7	2.4	6.8	6.7
15	23	6.9	4.2	7.5	7.1
15	24	6.7	4.8	7.9	7.5

15	25	7.5	4.5	8.2	7.8
15	26	5.1	3.6	5.8	5.7
15	27	4.1	2.3	4.3	4.3
15	28	8.1	3.3	7.7	7.4
15	29	7.2	5.2	8.5	8.1
15	30	7.1	3.3	6.9	6.8

Visual search with external uncertainty and short display time (conditions B-D)

The three visual search conditions with external uncertainty and short display time were identical to the condition just described, except that the orientations of the target and distractors were no longer fixed, but instead drawn from partly overlapping Gaussian distributions (Fig. 1D). These distributions had means μ and $-\mu$, respectively (see above), and a standard deviation σ_{external}. The value of σ_{external} was customized for each subject (Table 2) such that the accuracy of an optimal observer would drop by 5, 10, or 15% compared to the same condition without external uncertainty. We refer to each of these percentages as a *level of external uncertainty*. Subjects completed 1500 trials divided equally over 12 blocks with short forced breaks between blocks.

Visual search with external uncertainty and unlimited display time (conditions E-G)

These three conditions were identical to conditions B-D, except for the following two differences. First, stimuli were presented with an ellipse eccentricity of 0.97 and stayed on the screen until a response was provided, such that the sensory noise levels were reduced to a presumably negligible level. Second, this condition contained 500 instead of 1500 trials. Each subject completed this condition before the equivalent condition with short display times.

Statistical analyses

All Bayesian statistical tests were performed using the JASP software package [34] with default settings. The output of these tests is a Bayes factor, denoted BF_{10}, which specifies the ratio between how likely the data are under the alternative hypothesis compared to how likely they are under the null hypothesis. Hence, values smaller than 1 indicate evidence for the null hypothesis and values larger than 1 indicate evidence for the alternative hypothesis. When a test supports the null hypothesis we usually report $BF_{01}=1/BF_{10}$.

Modeling methods

Optimal decision variable

We denote target presence by a binary variable T (−1=absent, 1=present), set size by N, the stimulus values by $\mathbf{s}=\{s_1, s_2, \ldots, s_N\}$, and the observer's noisy observations of the stimulus values by $\mathbf{x}=\{s_1, s_2, \ldots, s_N\}$. For convenience, Table 3 presents an overview of all symbols we use in our mathematical descriptions of the models and experiments. The Bayesian optimal observer reports "target present" if the posterior probability of target presence exceeds that of target absence, $p(T=1|\mathbf{x})>p(T=-1|\mathbf{x})$. This strategy is equivalent to reporting "target present" if the log posterior ratio exceeds 0,

$$d(\mathbf{x}) \equiv \log \frac{p(T=1|\mathbf{x})}{p(T=-1|\mathbf{x})} > 0.$$

where $d(\mathbf{x})$ is referred to as the global decision variable. Taking into account the statistical structure of the task (S1 Fig), this evaluates to (S1 Appendix):

$$d(\mathbf{x}) = \log\left(\frac{1}{N}\sum_{i=1}^{N} d_{\text{local}}(x_i)\right), \tag{1}$$

where

$$d_{\text{local}}(x_i) = \exp\left[\frac{(x_i+\mu)^2-(x_i-\mu)^2}{2\left(\sigma_i^2+\sigma_{\text{external}}^2\right)}\right] \tag{2}$$

is referred to as the local decision variable. Hence, the optimal decision variable is the log of an average of local decision variables, each of which represents the evidence (posterior ratio) for target presence: $d_{\text{local}}(x_i)<1$ is evidence for a distractor at location i and $d_{\text{local}}(x_i)>1$ is evidence for a target. In the four conditions with short displays times (A-D, Table 1), the sensory noise level associated with a stimulus, σ_i, differed for stimuli with low and high reliability. In the three conditions with unlimited display time (conditions E-G), we assume σ_i to be 0.

We mentioned earlier that optimal observers – just like good detectives – weight each cue by its reliability. In the visual search task, each stimulus observation, x_i, is a cue. Equation (2) demonstrates that the optimal observer indeed weights these cues by their reliability: the larger the sensory noise, σ_i, the smaller the magnitude of the local evidence for target presence, $d_{\text{local}}(x_i)$.

Table 3. Overview of the symbols used in our mathematical descriptions of the experiments and models.

Symbol	Description
$\tilde{\sigma}_{\text{low}}, \tilde{\sigma}_{\text{high}}$	Estimated sensory noise levels in the discrimination task
$-\mu,\ +\mu$	Means of the distractor and target distributions, respectively
σ_{external}	Standard deviation of the distractor and target distributions in the visual search task
N	Number of stimuli in the visual search task
L	Location of target in the visual search task (when present)
T	Target presence in the visual search task (−1=absent, +1=present)
s_i	Orientation of the i-th stimulus in the visual search task
x_i	Noisy observation of s_i
σ_i	Estimated sensory noise levels associated with stimulus s_i in the visual search task; computed as $\sigma_i = \alpha\tilde{\sigma}_{\text{low}}$ when s_i has low reliability and as $\sigma_i = \alpha\tilde{\sigma}_{\text{high}}$ otherwise.
$\hat{\sigma}_i$	Sensory noise level assumed by the observer in the i-th stimulus when computing the local decision variable
η	Computational noise, a Gaussian random variable with mean zero
$\sigma_{computational}$	Standard deviation of the distribution of the computational noise

Models

To test for deviations from optimality, we extend the optimal model with two types of suboptimality.

Suboptimality 1: imperfect cue weighting. As explained above, a hallmark of optimal observers is that they weight sensory cues by their reliability, which requires subjects to have perfect knowledge of their own sensory noise levels. Although it has been suggested that such knowledge may be implicitly represented in the neural representation of the stimulus [35], we allow for possible imperfections in this knowledge by introducing a variable $\hat{\sigma}_i$ that represents the observer's "assumed" level of sensory noise in the *i*-th stimulus when computing the local decision variable related to this stimulus,

$$d_{\text{local}}(x_i) = \exp\left[\frac{(x_i+\mu)^2-(x_i-\mu)^2}{2\left(\hat{\sigma}_i^2+\sigma_{\text{external}}^2\right)}\right]. \tag{3}$$

The optimal local decision variable is the special case in which the assumed level of noise is identical to the true level of noise with which stimuli are encoded, $\hat{\sigma}_i = \sigma_i$.

Suboptimality 2: computational noise. The second kind of suboptimality that we incorporate in the model is "computational noise" on the global decision variable,

$$d(\mathbf{x}) = \log\left(\frac{1}{N}\sum_{i=1}^{N} d_{\text{local}}(x_i)\right)+\eta, \tag{4}$$

where η is a zero-mean Gaussian random variable with a standard deviation of $\sigma_{\text{computational}}$. One way to interpret this parameter is by thinking of it as capturing effects caused by random variability in the activity of the neurons that encode the global decision variable. However, simulations (Fig. 2) show that this parameter may also capture effects caused by other types of suboptimalities, including noise in the computation or representation of the local decision variables, incorrect knowledge of experimental parameters μ and σ_{external}, and incorrect cue weighting[2].

Factorial model design. Combining Eqs. (3) and (4) yields

$$d(\mathbf{x}) = \log\left(\frac{1}{N}\sum_{i=1}^{N}\exp\left[\frac{(x_i+\mu)^2-(x_i-\mu)^2}{2\left(\hat{\sigma}_i^2+\sigma_{\text{external}}^2\right)}\right]\right) + \eta. \tag{5}$$

Based on this equation, we implement a 3×2 factorial set of models (Table 4). The first factor determines how the model weights sensory cues:

i) *Optimal weighting*. The observer has perfect knowledge of the amount of sensory noise at each location and uses this knowledge perfectly when computing local decision variables, $\hat{\sigma}_i = \sigma_i$.

ii) *Suboptimal weighting*. The observer weights evidence differently for low-reliability and high-reliability stimuli, but possibly using weights that deviate from the optimal ones. We model this by setting $\hat{\sigma}_i = \hat{\sigma}_{\text{low}}$ for low-reliability stimuli and $\hat{\sigma}_i = \hat{\sigma}_{\text{high}}$ for high-reliability stimuli, where $\hat{\sigma}_{\text{low}}$ and $\hat{\sigma}_{high}$ are free parameters. This option allows for under-weighting or over-weighting of stimuli (e.g., $\hat{\sigma}_{\text{low}} < \sigma_{\text{low}}$ would result in overweighting of low-reliability stimuli).

iii) *No weighting*. The observer does not differentiate between low-reliability and high-reliability stimuli when computing local evidence, $\hat{\sigma}_i = \hat{\sigma}_{\text{both}}$, where $\hat{\sigma}_{\text{both}}$ is a free parameter.

The second factor determines the presence of computational imperfections:

[2] Even though incorrect cue weighting appears as computational noise on the decision variable, the models can still distinguish between the effects that both types of suboptimality have on the response data (see section *Verification of analysis methods* in Results).

i) *No computational imperfections*. There is no computational noise, $\sigma_{\text{computational}}=0$.

ii) *Imperfections in the computation of the global decision variable*. The global decision variable, $d(\mathbf{x})$, is corrupted by noise drawn from a zero-mean Gaussian distribution with a standard deviation $\sigma_{\text{computational}}\geq 0$.

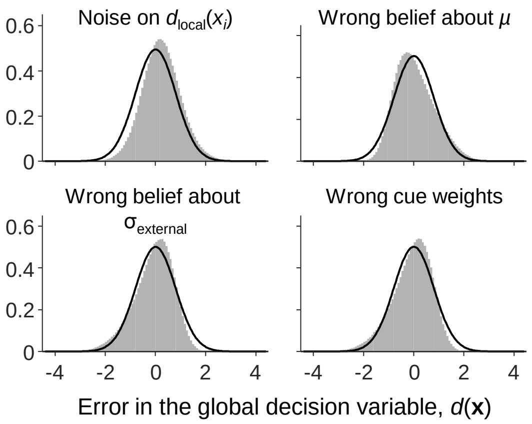

Fig. 2. Four types of suboptimality that produce near-normally distributed errors in the global decision variable. We simulated 1 million trials of the visual search task and computed for each trial the global decision variable in four suboptimal variants of the optimal model. The histograms (gray areas) show the distributions of the error in these suboptimal decision variables relative to the optimal one. In the first variant, local decision variables were corrupted by Gaussian noise (top left). In the second and third variants, local decision variables were computed using incorrect values for the mean (top right) or standard deviation (bottom left) of the stimulus distributions. In the last variant, local decision variables were computed using incorrect sensory weights (bottom right). All four distributions are reasonably well approximated by a Gaussian distribution (black curves). This suggests that the behavioral effects of these suboptimalities can be captured by a model in which the global optimal decision variable is corrupted by Gaussian noise.

Table 4. Overview of models for the visual search task with sensory uncertainty. Model M5 is the optimal model.

Model ID	Factor 1: Sensory weighting	Factor 2: Computational noise	Free parameters
M1	None	No	$\alpha,\ \hat{\sigma}_{\text{both}}$
M2	None	Yes	$\alpha,\ \hat{\sigma}_{\text{both}},\ \sigma_{computational}$
M3	Suboptimal	No	$\alpha,\ \hat{\sigma}_{low},\ \hat{\sigma}_{high}$
M4	Suboptimal	Yes	$\alpha,\ \hat{\sigma}_{low},\ \hat{\sigma}_{high},\ \sigma_{computational}$
M5	Optimal	No	α
M6	Optimal	Yes	$\alpha,\ \sigma_{computational}$

Simplified model for conditions without sensory uncertainty

In three of the experimental conditions (E-G, Table 1), stimuli are presented at high contrast and can be viewed for as long as the subject wants to. Assuming that sensory noise is negligible in these conditions, $\sigma_i \approx 0$, the stimulus observations, **x**, are equal to the presented stimulus orientations, **s**. The optimal decision variable can then be written directly as a function of **s**, and simplifies to

$$d(\mathbf{s}) = \log\left(\frac{1}{N}\sum_{i=1}^{N}\exp\left[\frac{(s_i+\mu)^2-(s_i-\mu)^2}{2\sigma_{\text{external}}^2}\right]\right). \quad (6)$$

Without sensory noise, there can be no suboptimality in sensory cue weights. Therefore, computational noise is the only suboptimality that we model in these conditions, in the same way as above. The optimal decision variable in the model for conditions without sensory noise is thus computed as

$$d(\mathbf{s}) = \log\left(\frac{1}{N}\sum_{i=1}^{N}\exp\left[\frac{(s_i+\mu)^2-(s_i-\mu)^2}{2\sigma_{\text{external}}^2}\right]\right) + \eta. \quad (7)$$

Model fitting and model comparison

We use an adaptive Bayesian optimization method [37] to find maximum-likelihood estimates of model parameters, at the level of individual subjects. Model evidence is measured as the Akaike Information Criterion [38] and interpreted using the rules of thumb provided by Burnham & Anderson [39].

Constraining of estimated sensory noise levels

We use the estimates of $\tilde{\sigma}_{\text{low}}$ and $\tilde{\sigma}_{\text{high}}$ from the discrimination task (Table 2) to constrain the sensory noise levels in the models for visual search. However, set size differed between the two experiments (1 *vs.* 4) and previous work has shown that sensory noise levels may increase with set size [15,36]. In particular, those studies found that the standard deviation of the noise distribution can increase up to a factor 2 between set sizes 1 and 4, depending on the heterogeneity of the distractors (Fig. 7B in [15] and Fig. 9 in [36]). Therefore, we expect that sensory noise levels in the visual search task were between a factor 1 and 2 larger as in the discrimination task. We denote this factor by α. Instead of fitting σ_{low} and σ_{high} in the visual search task as two entirely free parameters, we compute them as $\sigma_{\text{low}} = \alpha\tilde{\sigma}_{\text{low}}$ and $\sigma_{\text{high}} = \alpha\tilde{\sigma}_{\text{high}}$, and we instead fit α as a single free parameter. Moreover, we constrain α to take values between 1 and 2, by imposing a Gaussian prior on this parameter with a mean of 1.50 and a standard deviation of 0.20.

Results

Discrimination task

To estimate the effect of ellipse elongation on the sensory precision with which subjects encoded the stimulus orientations, we fit two models to each subject's data in the discrimination task. In both models, stimulus observations are assumed to be corrupted by Gaussian noise. Under this assumption, the predicted proportion of "clockwise" responses is a cumulative Gaussian as a function of stimulus orientation. We refer to the standard deviation of this Gaussian as the sensory noise level. In the first model, the noise level is independent of ellipse elongation and fitted as a single free parameter. In the second model, the sensory noise levels are fitted as separate parameters for the low- and high-reliability stimuli, which we denote by $\tilde{\sigma}_{low}$ and $\tilde{\sigma}_{high}$, respectively. The second model accounts well for the data (Fig. 1B) and model comparison favors this model for every subject (ΔAIC range: 0.50 to 22.3; mean±sem:

8.6±1.3[3]). Moreover, for all subjects the estimated noise level is higher for the low-reliability stimulus than for the high-reliability stimulus (Table 2). Hence, the stimulus-reliability manipulation works as intended. As described in Methods, we use the estimates of $\tilde{\sigma}_{\text{low}}$ and $\tilde{\sigma}_{\text{high}}$ to customize the target and distractor distributions in the visual search experiment (Table 2) and to constrain the models fitted to the data from that experiment.

Visual search with unlimited display time

Evidence for suboptimalities. Since stimuli in these conditions were high in contrast and could be inspected for as long as a subject wanted, we assume for the moment that sensory noise was negligible[4]. Under this assumption, Eq. (6) specifies the optimal decision variable, $d(\mathbf{s})$, which we can compute for each trial without the need to fit any free parameters. The optimal observer responds "target present" if $d(\mathbf{s})>0$ and "target absent" otherwise. In other words, if subjects are optimal, then their proportion of "target present" responses should be a step function of $d(\mathbf{s})$, transitioning from 0 to 1 at $d(\mathbf{s})=0$ (Fig. 3A, red lines). In all three conditions, subjects clearly deviate from this prediction (Fig. 3A, black circles), which suggests that they performed suboptimally.

We hypothesize that the suboptimality is caused by the kind of inferential imperfections that we intend to capture with the computational noise parameter. Consistent with this hypothesis, we find that the data are well accounted for by the model that uses Eq. (7) to compute the decision variable (Fig. 3A, black curve).

[3] Throughout the paper X±Y refers to mean±sem across subjects, unless stated otherwise.
[4] Later in this section, we provide an analysis where we do not make this assumption.

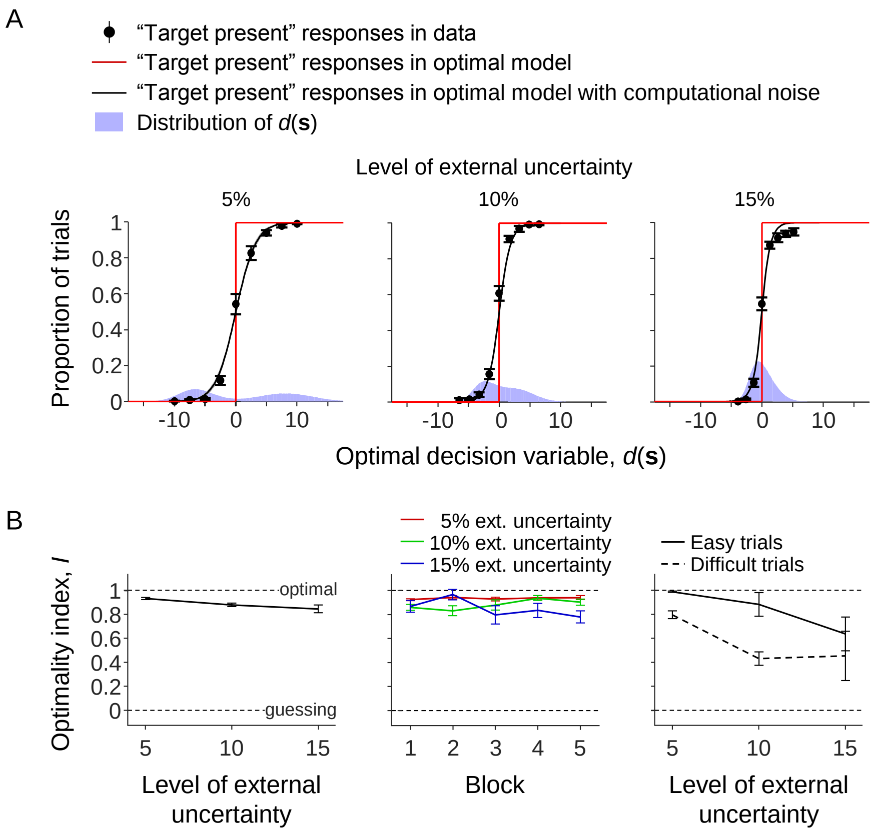

Fig. 3. Results from the visual search conditions with unlimited display. (A) Assuming that sensory noise was negligible in these conditions, the optimal model (red) predicts that the proportion of "target present" responses is a step function of the optimal decision variable, $d(\mathbf{s})$. The subject data (black markers) clearly deviate from this prediction. The data are well accounted for by a model in which the optimal decision variable is corrupted by Gaussian noise (black curves). (B) The optimality index decreases with the level of external uncertainty (left), did not systematically change across blocks (center), and was lower for difficult trials than for easy ones (right). Note that the distribution of $d(\mathbf{s})$ (purple areas in panel A) becomes more concentrated around zero as the level of external uncertainty increases. This explains why performance decreases with the level of external uncertainty (panel B), despite the response data becoming more step-like.

Quantification of optimality loss. To quantify how much subject performance deviated from optimal performance, we introduce an index I that measures performance on a scale that extends from random guessing to optimal performance,

$$I \equiv \frac{p_{\text{subject}} - p_{\text{guess}}}{p_{\text{optimal}} - p_{\text{guess}}}, \tag{8}$$

where p_{subject} is the subject's proportion of correct responses, p_{guess} is chance-level accuracy (0.50 in all our tasks), and p_{optimal} is the accuracy expected from an optimal observer. When there is no sensory noise, p_{optimal} can be computed in a parameter-free way, namely as the proportion of trials in which the sign of the optimal decision variable, Eq. (6), is equal to the sign of the binary variable that specifies target presence, T. The index takes values between 0 and 1, with 0 corresponding to random guessing and 1 to optimal behavior.

The subject-averaged optimality index across all three conditions is 0.877±0.015 (Fig. 3B, left). A Bayesian one-sample t-test reveals extremely strong evidence for the hypothesis $I<1$ ($BF_{-0}=5.51\cdot10^6$). This hypothesis is also supported when analyzing the index separately in each of the three conditions, $BF_{-0}>701$. Moreover, a Bayesian one-way ANOVA provides evidence for the hypothesis that the optimality index depends on the level of external uncertainty where the optimality index decreases with increasing levels of external uncertainty ($BF_{10}=15.7$). At first sight, the latter finding may seem to contradict the pattern observed in Fig. 3A: if the optimality index decreases with external uncertainty, how is it then possible that the response curve increasingly resembles the step function predicted by the optimal observer? This can be explained by considering differences in the distribution of $d(\mathbf{s})$ across conditions: the larger the external uncertainty, the more difficult the task, i.e., the more narrowly $d(\mathbf{s})$ is concentrated around 0. Therefore, when the level of external uncertainty is larger, the deviations from the step function matter more for task accuracy than when there is a wider spread of $d(\mathbf{s})$.

Ruling out alternative explanations of the optimality loss. A one-way Bayesian ANOVA reveals strong evidence against an effect of block on the optimality index ($BF_{01}=21.2$; Fig. 3B, center), which suggests that the suboptimality is not simply caused by a lack of learning at the beginning of the session. Moreover, we find a difference between the suboptimality index for easy and difficult trials (Fig. 3B, right; Bayesian one-way ANOVA $BF_{10}=134$)[5]. This suggests that the suboptimality is also not caused by attentional lapses, which would be expected to affect optimality similarly in easy and difficult trials. In addition, if subjects had been guessing on a

[5] Difficulty was measured as the absolute value of decision variable $d(\mathbf{s})$. Trials with $|d(\mathbf{s})|$ smaller than the first quartile were defined as difficult and trials with $|d(\mathbf{s})|$ larger than the third quartile were defined as easy.

significant proportion of trials, then the asymptotes in their response curves (Fig. 3A) would have deviated from 0 and 1, which does not seem to be the case. Hence, neither a lack of learning nor guessing due to attentional lapses seems to be a plausible explanation of the suboptimality.

Validating the assumption that sensory noise was negligible. So far, we have assumed that there is no sensory noise in these conditions, σ_i=0. However, despite the unlimited display time, it is unlikely that subjects encoded stimulus orientations without any error at all. To obtain an estimate of σ_i in these conditions, we conduct a control experiment that is identical to the discrimination experiment (Fig. 1A), except that the stimulus has an ellipse eccentricity of 0.97 and stays on the screen until a response is given. By fitting a cumulative Gaussian to the data from this experiment, we find an estimate of σ_i=0.875±0.097. Recomputing the optimality index under this value of sensory noise gives I=0.898±0.016, which is barely higher than the 0.877±0.015 we found under the assumption of σ_i=0. Indeed, a Bayesian t-test does not provide any evidence for a difference (BF_{01}=2.59). Hence, even though sensory noise may not entirely have been eliminated in the conditions with unlimited stimulus presentation time, it was so low that the effect on the optimality index was negligible.

Conclusions. We draw three conclusions from the visual search tasks with unlimited stimulus time. First, the data present evidence for a suboptimality. Second, the degree of suboptimality increases with the level of external uncertainty. Third, the suboptimality cannot be explained as the result of guessing or a lack of learning, but is well accounted for by a model in which the optimal decision variable is corrupted by Gaussian noise.

Visual search with short display times

Model comparison. We hypothesize that there may be two kinds of suboptimality beyond sensory noise: suboptimal cue weighting (model factor 1) and computational imperfections (model factor 2). To quantify evidence for these hypotheses, we fit a factorial set of six models and compute the relative support for each of them (Table 5). The best-fitting model – and the only one that is supported by the data – is the model that includes both types of suboptimality (M4). We find no support for models in which subjects ignore differences in stimulus reliability (M1 and M2), which is consistent with findings in previous studies [13,19,21]. However, in contrast to those studies, we also find no support for the models in which subjects weight stimuli optimally (M5 and M6)[6]. Therefore, the data suggest that subjects take stimulus reliability into account, but may do so suboptimally. Finally, there is no support for any model that does not include computational noise.

Table 5. Overview of strength of support for each model, measured as the subject-averaged AIC value relative to the best-fitting model (M4). The interpretation of the evidence is based on the rules of thumb provided in [39]**.**

Model	Sensory weighting	Computational noise	ΔAIC	Interpretation
M1	None	No	61.8 ± 8.8	No support
M2	None	Yes	34.3 ± 5.0	No support
M3	Suboptimal	No	12.1 ± 4.2	No support
M4	Suboptimal	Yes	0	Best model
M5	Optimal	No	60.1 ± 18.2	No support
M6	Optimal	Yes	37.7 ± 12.1	No support

Model fits. Although the data provide no support for the optimal model (ΔAIC=60.1±18.2), its account of the subject-averaged hit and false alarm rates is visually only slightly worse than that of the best-fitting model, except in the condition with the highest level of uncertainty (Fig. 4A). However, suboptimal behavior at the level of individual subjects may appear optimal when viewed at the level of the group [23]. Indeed, when analyzed at the level

[6] Note, however, that this finding is not necessarily inconsistent with previous studies. Those studies compared the optimal model only to the no-weighting model and did not consider the option of suboptimal weights. If we had done the same here, we also would have concluded that subjects weighted evidence optimally.

of individual subjects[7], the combined root mean squared error of the optimal model is almost twice as large as that of the best-fitting suboptimal model (0.0979 *vs.* 0.0571). The model fits in Fig. 4A also show that there are aspects of the data that neither of the models explain well. In particular, both models consistently underestimate the hit rate when the target is the only high-reliability item. This suggests that there may be additional suboptimalities in the data that are not captured by the model.

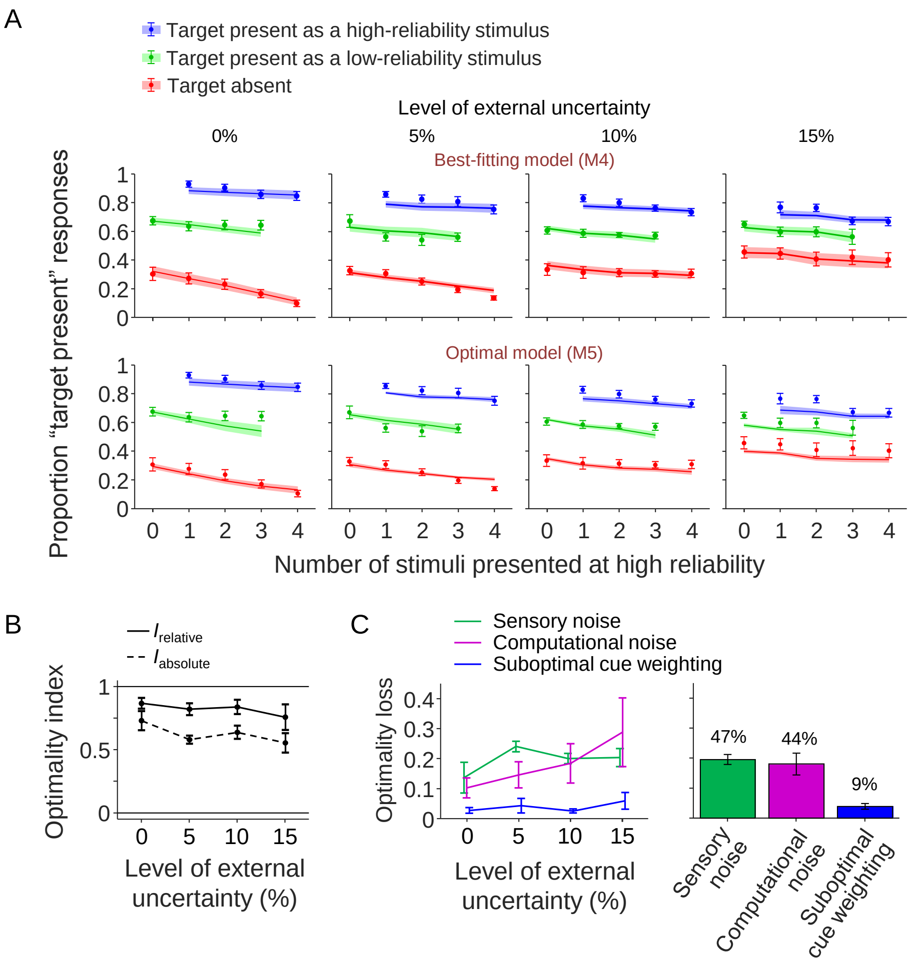

Fig. 4. Results from the visual search conditions with short display time. (A) False-alarm rates (red) and hit rates conditioned on whether the target had high reliability (blue) or low reliability (green). The data (markers) are well accounted for by the model with suboptimal weights and computational noise (curves). However, note that there seem to be some systematic deviations between the model behavior and the data, such as its consistent underestimation of the hit rate when only the target has high reliability. This suggests that there may be suboptimlaties in the data that are not captured in the model. (B) Relative and absolut optimality index. The difference between the two indices represents the amount of optimality loss caused by sensory noise and the difference between the relative optimality index and 1 measures the amount of optimality loss due to other sources. (C) Optimality loss decomposed into three sources.

[7] Individual plots can be generated using the code and data provided at https://osf.io/dkavj/.

Optimality index. We next estimate how much subject performance deviates from optimal performance in these conditions. Because of the presence of sensory noise, we now make a distinction between absolute and relative optimality [3]. An observer is defined as optimal in the *absolute* sense when their accuracy equals that of a noiseless optimal observer. In the condition without external uncertainty, this corresponds to an accuracy level of 100% correct; in the other conditions, the maximum accuracy levels are dictated by experimental parameters μ and σ_{external}. A subject is defined as optimal in the *relative* sense when accuracy is as high as possible given the subject's estimated levels of sensory noise. Both optimality indices – which we denote by I_{absolute} and I_{relative} – are computed using Eq. (8), with p_{optimal} being the only variable differing between the two measures. When computing I_{absolute}, we estimate p_{optimal} in the same way as we did in the analysis of conditions with unlimited display time (see above). Computing p_{optimal} for the relative optimality index, however, requires an estimate of the subject's sensory noise levels. To get the best possible estimate available, we average the estimated sensory noise levels ($\alpha\tilde{\sigma}_{\text{low}}$ and $\alpha\tilde{\sigma}_{\text{high}}$) across all six models, where we weight each value by the model's AIC weight [40], such that better models contribute more strongly.

Averaged across subjects and conditions, we find I_{absolute}= 0.626±0.032 and I_{relative}= 0.822±0.032 (Fig. 4B). The finding that $I_{\text{absolute}}<I_{\text{relative}}$ (BF_{-0}=910) indicates that sensory noise affected task accuracy, as expected. Moreover, the finding that $I_{\text{relative}}<1$ ($BF_{-0}=15.3\cdot10^3$) indicates that there are also suboptimalities beyond the effects of sensory noise.

Decomposing sources of optimality loss. We assess the relative impact separately for each of the three hypothesized sources of optimality loss: sensory noise, suboptimal weighting of sensory cues, and computational noise. We estimate the optimality loss due to sensory noise as the difference between I_{absolute} (the expected accuracy for a noiseless, optimal observer) and I_{relative} (the expected accuracy for an optimal observer with sensory noise). Averaged across subjects and levels of external uncertainty, we find that the optimality loss due to sensory noise is 0.195±0.016. To compute the optimality loss from the other two sources, we take for each subject the maximum-likelihood fit of model M4 and use simulations to compute how much the optimality index increases when "turning off" either type of suboptimality. We find that replacing the suboptimal weights with the optimal ones increases the optimality index with 0.039±0.009 and setting the computational noise to zero increases it with 0.179±0.035 (Fig. 4C). Expressed in percentages, sensory noise accounts for an estimated 47% of optimality loss, computational noise for 44%, and suboptimal weighting for the remaining 9%.

Comparison of optimality loss between conditions with and without sensory noise. In the conditions with unlimited display time we found an optimality index of 0.877±0.015, which is slightly higher than the relative suboptimality index $I_{relative}$=0.822±0.032 found in the conditions with short display times. However, a Bayesian independent samples t-test provides no evidence for the hypothesis that the average suboptimality index differs between conditions with and without sensory noise (BF_{10}=0.53)[8]. Hence, adding sensory uncertainty to the visual search task does not seem to have a major impact on suboptimalities from other sources. In contrast to the conditions with unlimited display time, however, we now find no evidence for an effect of the level of external uncertainty on the relative optimality index (BF_{01}=5.06).

Conclusions. The four main findings from the analysis of the visual search tasks with short display time are as follows. First, there is evidence for a deviation from optimality that goes beyond the effects of sensory noise, which is at odds with previous reports of (relative) optimality in human visual search. Second, although there is strong evidence for a suboptimality in cue weighting, the effects are relatively small. Third, addition of sensory noise to this task does not substantially affect the degree of suboptimality in downstream computations. Finally,

[8] This test was performed only on the data from the conditions with 5, 10, and 15% external uncertainty, because the experiment with unlimited display time did not include a condition with 0% external uncertainty. When including the condition with 0% external uncertainty in the experiment with short display time, the outcome is very similar (BF_{10}=0.58).

there is no evidence that the degree of suboptimality depends on the level of external uncertainty.

Verification of analysis methods

A possible concern about our analyses is that – despite the measures we took to reduce model flexibility – sensory noise, suboptimal cue weighting, and computational noise may have had partly interchangeable effects on the model predictions. If so, then there is a risk that data from optimal observers – whose performance is only affected by sensory noise – may give rise to spurious evidence for suboptimal cue weighting and computational noise. Therefore, we verify the reliability of our findings by applying the same methods to three synthetic data sets.

Synthetic dataset 1: optimal observers. In the first of these analyses, we use the optimal model (M5) to generate ten datasets at each level of external uncertainty, with the same number of trials as in the subject data. In each simulation, the value of the only free model parameter, α, is drawn from a uniform distribution between 1 and 2. If our methods are reliable, then we should find no evidence for suboptimalities in these synthetic data. Consistent with this, we find that an AIC-based model comparison correctly selects the optimal model as the preferred model. Moreover, we find a relative optimality index of 0.986±0.012 and evidence for the null hypothesis that the average of the population from which these values originate is 1 (BF_{01}=3.2). Hence, our methods are unlikely to provide evidence for suboptimalities on data from optimal observers.

Synthetic dataset 2: observers with suboptimal cue weights. The second analysis is analogous to the first one, except that we use the model variant with suboptimal weights (M1) to generate the synthetic data. We find that the model comparison correctly selects M1 as the preferred model. Moreover, we find a relative optimality index equal to 0.857±0.017, which indicates clear evidence for a suboptimality beyond sensory noise. Decomposition of the optimality loss suggests that 59% is caused by sensory noise, 38% by suboptimal cue weighting, and 3% by computational noise. Hence, as expected, our methods identify sensory noise and suboptimal cue weighting as factors that strongly affected accuracy in this dataset, but not computational noise. This means that even though suboptimal weighting appears as computational noise on the optimal decision variable (Fig. 2, bottom right), these two factors are not confounded in the models. The reason is that suboptimal weighting causes *systematic* errors in the decision variable, which model M1 can explain on a trial-by-trial basis, while models with computational noise only explain the effect that these errors have on the hit- and false-alarm rate averaged across trials.

Synthetic dataset 3: observers with computational noise. Finally, we perform a similar analysis on data generated from the model with computational noise (M6). As expected, model comparison selects M6 as the preferred model on these data. Moreover, the relative optimality index is 0.665±0.033 and the decomposition analysis estimates that 33% of the optimality loss is due to sensory noise, 1% due to suboptimal weighting, and 66% due to computational noise.

Conclusions. These simulation results show that our methods do not produce evidence for suboptimalities when applied to data generated from the optimal model. Hence, the suboptimalities found in the subject data are unlikely to be the result of a flaw in analysis methods. Moreover, the results show that the methods reliably distinguish between effects of suboptimal cue weighting and effects of computational noise, which validates our decomposition analysis.

General discussion

Summary of results

In this study, we re-examined human performance on visual-search tasks by making several methodological improvements over previous studies. First, we used an experimental design that included conditions without sensory noise, which allowed us to estimate deviations from optimality without fitting any free parameters. Second, in conditions with sensory noise, we constrained model parameters by using prior knowledge obtained from a separate task. Third, we tested the optimal model against a more general type of suboptimal model than previous studies. Fourth, we decomposed loss of optimality into three sources and quantified the contribution of each of them. Our results show evidence for suboptimalities in human visual search, at all tested levels of internal and external noise. In the conditions with sensory noise, we estimated that about 47% of the accuracy loss was due to sensory noise, 9% due to suboptimal weighting of sensory cues, and the remaining 44% due to other computational imperfections. Our results are consistent with previous evidence that humans take stimulus reliability near-optimally into account during perceptual inference (e.g., [13,18,19,21]). However, they do not support previous suggestions that human visual-search behavior is optimal (e.g., [13–16]). Instead, they support recent evidence suggesting that inference in neural systems is inherently imprecise [27,28,30,31].

Related work

Our approach and our findings bear similarities to recent work by Drugowitsch and colleagues [27]. They used a visual categorization task in which subjects were presented on each trial with a sample of sixteen stimuli, whose orientations were drawn from one of two or three distributions. The subject's task was to indicate from which of the distributions the orientations were drawn. Just as in our task, the optimal decision variable for this task is based on a sum of local decision variables, each of which is itself a non-linear function of a stimulus observation. Drugowitsch et al. reported evidence for suboptimalities in behavior due to computational imprecisions, which is consistent with our own findings. However, whereas we found that such imprecisions accounted for around 53% of the total optimality loss in the conditions with sensory noise, Drugowitsch et al. reported an estimate of over 90%. One possible explanation for this difference is that computational imprecisions may be smaller in visual search compared to categorization. However, an alternative explanation – which we believe is more plausible – is that sensory noise levels were simply larger in our experiment, due to difference in stimulus presentation time (67 ms to encode four stimuli in our study; 333 ms per stimulus in the study by Drugowitsch et al.).

Although we are not aware of any other studies that have decomposed sources of suboptimality in perceptual inference tasks, several other studies have reported evidence for the general presence of suboptimalities in such tasks. For example, numerous sensory cue combination studies have reported overweighting of one of the sensory cues [41–48], Bhardwaj et al. [49] found that visual search performance is suboptimal when stimuli are correlated, and Qamar et al. [50] found that some of their human and monkey subjects used a suboptimal decision rule in a visual categorization task.

Our study is also not the first to test models in which the optimal decision variable is corrupted by noise. An example of our own previous work – in which we referred to it as "decision noise" – concerns the change detection study by Keshvari et al. [19], where we found that inclusion of this noise parameter did not substantially improve the model fits. However, sensory noise levels in that study were fitted in an entirely unconstrained way, while it is conceivable that there was a trade-off between effects of noise on the decision variable and effects of sensory noise on model predictions. Moreover, in that study we assumed random variability in encoding precision, which – as it turns out – is easily confounded with decision noise [51]. Therefore, it is possible that imprecisions in inference went unnoticed due to confounding them with sensory noise and variability in precision. Another example of work that has considered noise on the decision variable is a collection of studies by Summerfield and colleagues, who refer to it as "late noise" (e.g., [52,53]). They have shown that in the presence of such noise, subjects can and often do obtain performance benefits by using "robust averaging", i.e., down-weighting outlier cues (compared to the optimal observer) when computing the global decision variable. We performed simulations to examine whether this strategy may also give performance benefits in our visual search task, but we did not find any clear evidence for it. One difference between our task and the tasks typically used by Summerfield et al. is that in their case each cue is a stimulus observation, while in our case they are non-linear functions of stimulus observations (Eq. (2)). Further work is required to examine whether this difference explains why robust averaging does not seem to be a beneficial strategy in visual search.

Finally, the within-display manipulation of stimulus reliability that we used here has been applied in earlier studies using visual search [13], categorization [18], change detection [19], and same/different discrimination [21] experiments. Consistent with those studies, we found strong evidence against models that give equal weight to stimuli with low and high reliability. Although we found evidence for a discrepancy between the estimated weights used by the subjects and the optimal weights, the optimality loss caused by this discrepancy was small (8%). Therefore, our findings suggest that – despite our evidence for inferential suboptimalities – subjects weighted sensory cues near-optimally by their reliability.

External uncertainty

Unlike most previous studies on visual search, we added external noise to the stimuli (however, see [54] for a similar manipulation). We believe that this approach has two advantages over using deterministic stimuli that could make it valuable in other perception studies too. First, it allows the experimenter to include task conditions without sensory noise. As demonstrated here, such conditions allow the experimenter to assess deviations from optimality without the need to fit free parameters. However, a second advantage is that tasks with external uncertainty may be more consistent with naturalistic conditions, where errors in judgment often arise not only due to sensory uncertainty but also due to external ambiguities, for example caused by imperfect correlations between features in the environment. Due to such ambiguities, naturalistic stimuli are often probabilistic rather than deterministic [32], which prevents even a noiseless optimal observer from reaching 100% correct performance.

Our results regarding the effect of external uncertainty on optimality are inconclusive: an effect was found in the conditions without sensory noise, but not in the conditions with sensory noise. One possibility is that the identified effect was a statistical fluke. However, an alternative possibility is that an effect is simply harder to establish in the presence of sensory noise. As explained above, computing the optimality index then requires estimating the subject's level of sensory noise. Imprecisions in these estimates will increase the variance of the optimality-index estimates which, in turn, will reduce the likelihood of finding statistically significant effects. Consistent with this reasoning, we found that the variance in the optimality index estimates was indeed more than 7 times larger in the conditions with sensory noise.

Further decomposition of sources of suboptimality

In the conditions with sensory noise, we decomposed suboptimalities into three sources: sensory noise, suboptimal cue weighting, and computational noise. In our experiment, these sources accounted for about 47%, 9%, and 44%, respectively, of the optimality loss. The first two sources have quite a specific interpretation, but effects captured by the computational noise parameter could stem from a number of different sources, such as random variability in the neurons that represent the local and global decision variables, imprecise knowledge of experimental parameters – such as the stimulus distributions – and the use of a suboptimal cue integration rule. We tried to further decompose suboptimalities into these more specific sources, but were unable to do so. The problem is that different types of suboptimalities have near-identical effects on the response data, due to which we were unable to reliably distinguish models that implemented more specific kinds of suboptimalities. Future studies may try to solve this model-identifiability problem by using experimental paradigms that provide a richer kind of behavior data to further constrain the models (e.g., by collecting confidence ratings [55]). Our experimental design also did not allow us to distinguish between systematic suboptimalities and random ones. Future studies could address this by using a double-pass paradigm [27,55]. The feasibility of this approach is highlighted in the study by Drugowitsch et al. [27], who estimated that in their visual categorization task about two-thirds of the inferential imprecisions were random and the other one third systematic.

References

1. Knill DC, Pouget A. The Bayesian brain: The role of uncertainty in neural coding and computation. Trends in Neurosciences. 2004. pp. 712–719. doi:10.1016/j.tins.2004.10.007

2. Knill DC, Richards W, editors. Perception as Bayesian Inference. Cambridge University Press; 1996.

3. Ma WJ. Organizing probabilistic models of perception. Trends Cogn Sci. 2012;16:

511–8. doi:10.1016/j.tics.2012.08.010

4. Geisler WS. Contributions of ideal observer theory to vision research. Vision Research. 2011. pp. 771–781. doi:10.1016/j.visres.2010.09.027

5. Trommershäuser J, Körding KP, Landy MS. Sensory Cue Integration. Sensory Cue Integration. 2012. doi:10.1093/acprof:oso/9780195387247.001.0001

6. Landy MS, Banks MS, Knill DC. Ideal-Observer Models of Cue Integration. Sensory Cue Integration. 2012. doi:10.1093/acprof:oso/9780195387247.003.0001

7. Ernst MO, Banks MS. Humans integrate visual and haptic information in a statistically optimal fashion. Nature. 2002;415: 429–33. doi:10.1038/415429a

8. van Beers RJ, Sittig AC, Gon JJ. Integration of proprioceptive and visual position-information: An experimentally supported model. J Neurophysiol. 1999;81: 1355–1364.

9. Alais D, Burr D. The ventriloquist effect results from near-optimal bimodal integration. Curr Biol. Elsevier; 2004;14: 257–62. doi:10.1016/j.cub.2004.01.029

10. Jacobs RA. Optimal integration of texture and motion cues to depth. Vision Res. 1999;39: 3621–3629. doi:10.1016/S0042-6989(99)00088-7

11. Landy MS, Maloney LT, Johnston EB, Young M. Measurement and modeling of depth cue combination: in defense of weak fusion. Vision Res. 1995; doi:10.1016/0042-6989(94)00176-M

12. Hillis JM, Watt SJ, Landy MS, Banks MS. Slant from texture and disparity cues: Optimal cue combination. J Vis. 2004;4: 1. doi:10.1167/4.12.1

13. Ma WJ, Navalpakkam V, Beck JM, Berg R van den, Pouget A. Behavior and neural

basis of near-optimal visual search. Nat Neurosci. 2011;14: 783–790. doi:10.1038/nn.2814

14. Vincent BT, Vincent BT, Baddeley RJ, Baddeley RJ, Troscianko T, Troscianko T, et al. Optimal feature integration in visual search. J Vis. 2009;9: 1–11. doi:10.1167/9.5.15.Introduction

15. Mazyar H, van den Berg R, Ma WJ. Does precision decrease with set size? J Vis. 2012;12: 10. doi:10.1167/12.6.10

16. Palmer J, Verghese P, Pavel M. The psychophysics of visual search. Vision Research. 2000. pp. 1227–1268. doi:10.1016/S0042-6989(99)00244-8

17. Navalpakkam V, Koch C, Perona P. Homo economicus in visual search. J Vis. 2009;9(1): 1–16. doi:10.1167/9.1.31.Introduction

18. Shen S, Ma WJ. A Detailed Comparison of Optimality and Simplicity in Perceptual Decision Making. Psychol Rev. 2016;123: 452.

19. Keshvari S, van den Berg R, Ma WJ. Probabilistic computation in human perception under variability in encoding precision. PLoS One. 2012;7.

20. Van Den Berg R, Shin H, Chou W-C, George R, Ma WJ. Variability in encoding precision accounts for visual short-term memory limitations. Proc Natl Acad Sci. 2012;109: 8780–8785. doi:10.1073/pnas.1117465109

21. van den Berg R, Vogel M, Josic K, Ma WJ. Optimal inference of sameness. Proc Natl Acad Sci U S A. 2012;109: 3178–83. doi:10.1073/pnas.1108790109

22. Rosas P, Wichmann FA. Cue Combination: Beyond Optimality. Sensory Cue Integration. 2012. doi:10.1093/acprof:oso/9780195387247.003.0008

23. Rahnev D, Denison RN. Suboptimality in Perceptual Decision Making. Behavioral and Brain Sciences. 2018: 1–107. doi:10.1017/S0140525X18000936

24. Jones M, Love BC. Bayesian Fundamentalism or Enlightenment? On the explanatory status and theoretical contributions of Bayesian models of cognition. Behav Brain Sci. Cambridge University Press; 2011;34: 169-88; disuccsion 188-231. doi:10.1017/S0140525X10003134

25. Bowers JS, Davis CJ. Bayesian just-so stories in psychology and neuroscience. Psychol Bull. 2012;138: 389–414. doi:10.1037/a0026450

26. Marcus GF, Davis E. How Robust Are Probabilistic Models of Higher-Level Cognition? Psychol Sci. 2013; doi:10.1177/0956797613495418

27. Drugowitsch J, Wyart V, Devauchelle A-DD, Koechlin E, Acerbi L, Vijayakumar S, et al. Computational Precision of Mental Inference as Critical Source of Human Choice Suboptimality. Neuron. 2016;92: 1398–1411. doi:10.1016/j.neuron.2016.11.005

28. Beck JM, Ma WJ, Pitkow X, Latham PE, Pouget A. Not Noisy, Just Wrong: The Role of Suboptimal Inference in Behavioral Variability. Neuron. 2012. pp. 30–39. doi:10.1016/j.neuron.2012.03.016

29. Faisal AA, Selen LPJ, Wolpert DM. Noise in the nervous system. Nat Rev Neurosci. 2008;9: 292–303. doi:10.1038/nrn2258

30. Renart A, Machens CK. Variability in neural activity and behavior. Current Opinion in Neurobiology. 2014. pp. 211–220. doi:10.1016/j.conb.2014.02.013

31. Wyart V, Koechlin E. Choice variability and suboptimality in uncertain environments. Current Opinion in Behavioral Sciences. 2016. pp. 109–115.

doi:10.1016/j.cobeha.2016.07.003

32. Juslin P, Olsson H. Thurstonian and Brunswikian origins of uncertainty in judgment: a sampling model of confidence in sensory discrimination. Psychol Rev. 1997;104: 344–366. doi:10.1037/0033-295X.104.2.344

33. Brainard DH. The Psychophysics Toolbox. Spat Vis. 1997;10: 433–436. doi:10.1163/156856897X00357

34. JASP Team. JASP (Version 0.8.4.0) [Computer program]. 2018.

35. Ma WJ, Jazayeri M. Neural Coding of Uncertainty and Probability. Annu Rev Neurosci. 2014;37: 205–220. doi:10.1146/annurev-neuro-071013-014017

36. Mazyar H, Van den Berg R, Seilheimer RL, Ma WJ. Independence is elusive : Set size effects on encoding precision in visual search. J Vis. 2013;13: 1–14. doi:10.1167/13.5.8.doi

37. Acerbi L, Ma WJ. Practical Bayesian Optimization for Model Fitting with Bayesian Adaptive Direct Search. Advances in Neural Information Processing Systems 30. 2017. pp. 1836–1846. doi:https://doi.org/10.1101/150052

38. Akaike H. A New Look at the Statistical Model Identification. IEEE Trans Automat Contr. 1974;19: 716–723. doi:10.1109/TAC.1974.1100705

39. Burnham KP, Anderson DR. Multimodel inference: Understanding AIC and BIC in model selection. Sociological Methods and Research. 2004. pp. 261–304. doi:10.1177/0049124104268644

40. Wagenmakers E-J, Farrell S. AIC model selection using Akaike weights. Psychon Bull Rev. 2004;11: 192–196. doi:10.3758/BF03206482

41. Battaglia PW, Jacobs RA, Aslin RN. Bayesian integration of visual and auditory signals for spatial localization. J Opt Soc Am A. 2003; doi:10.1364/JOSAA.20.001391

42. Burr D, Banks MS, Morrone MC. Auditory dominance over vision in the perception of interval duration. Exp Brain Res. 2009; doi:10.1007/s00221-009-1933-z

43. Maiworm M, Röder B. Suboptimal auditory dominance in audiovisual integration of temporal cues. Tsinghua Sci Technol. 2011; doi:10.1016/S1007-0214(11)70019-0

44. Fetsch CR, Pouget A, Deangelis GC, Angelaki DE. Neural correlates of reliability-based cue weighting during multisensory integration. Nat Neurosci. 2012; doi:10.1038/nn.2983

45. Prsa M, Gale S, Blanke O. Self-motion leads to mandatory cue fusion across sensory modalities. J Neurophysiol. 2012; doi:10.1152/jn.00439.2012

46. Battaglia PW, Kersten D, Schrater PR. How haptic size sensations improve distance perception. PLoS Comput Biol. 2011; doi:10.1371/journal.pcbi.1002080

47. Knill DC, Saunders JA. Do humans optimally integrate stereo and texture information for judgments of surface slant? Vision Res. 2003; doi:10.1016/S0042-6989(03)00458-9

48. Rosas P, Wichmann FA, Wagemans J. Texture and object motion in slant discrimination: failure of reliability-based weighting of cues may be evidence for strong fusion. J Vis. 2007; doi:10.1167/7.6.3

49. Bhardwaj M, Van Den Berg R, Ma WJ, Josic K. Do people take stimulus correlations into account in visual search? PLoS One. 2016;11. doi:10.1371/journal.pone.0149402

50. Qamar AT, Cotton RJ, George R, Beck JM, Prezhdo E, Laudano A, et al. Trial-to-trial, uncertainty-based adjustment of decision boundaries in visual categorization. Proc Natl

Acad Sci U S A. 2013;110: 20332–7. doi:10.1073/pnas.1219756110

51. Shen S, Ma WJ. Variable precision in visual perception. bioRxiv. 2018; 153650. doi:10.1101/153650

52. Gardelle V De, Summer C, de Gardelle V, Summerfield C. Robust averaging during perceptual judgment. Proc Natl Acad Sci. 2011;108: 13341–13346. doi:10.1073/pnas.1104517108

53. Li V, Herce Castañón S, Solomon JA, Vandormael H, Summerfield C. Robust averaging protects decisions from noise in neural computations. PLoS Comput Biol. 2017;13: e1005723. doi:10.1371/journal.pcbi.1005723

54. Baldassi S, Verghese P. Comparing integration rules in visual search. J Vis. 2002;2: 3–3. doi:10.1167/2.8.3

55. Cabrera CA, Lu ZL, Dosher BA. Separating decision and encoding noise in signal detection tasks. Psychol Rev. 2015;122: 429–460. doi:10.1037/a0039348