

Gambling in the Visual Periphery: A Conjoint-Measurement Analysis of Human Ability to Judge Visual Uncertainty

Hang Zhang^{1,2*}, Camille Morvan^{1,2,3}, Laurence T. Maloney^{1,2}

¹ Department of Psychology, New York University, New York, New York, United States of America, ² Center for Neural Science, New York University, New York, New York, United States of America, ³ Department of Psychology, Harvard University, Cambridge, Massachusetts, United States of America

Abstract

Recent work in motor control demonstrates that humans take their own motor uncertainty into account, adjusting the timing and goals of movement so as to maximize expected gain. Visual sensitivity varies dramatically with retinal location and target, and models of optimal visual search typically assume that the visual system takes retinal inhomogeneity into account in planning eye movements. Such models can then use the entire retina rather than just the fovea to speed search. Using a simple decision task, we evaluated human ability to compensate for retinal inhomogeneity. We first measured observers' sensitivity for targets, varying contrast and eccentricity. Observers then repeatedly chose between targets differing in eccentricity and contrast, selecting the one they would prefer to attempt: e.g., a low contrast target at 2°

Author Summary

Human ability to discriminate drops dramatically with increasing distance from the center of vision. If you fixate a word on a page, you likely can not read words a short distance away. Because of this retinal inhomogeneity, we need to move our eyes to search a scene. The efficiency of search depends on how well the visual system compensates for inhomogeneity in planning eye movements. We used a simple decision task to find out what the observer "knows" about his or her own retina. We first measured observers' sensitivity for targets, varying contrast and eccentricity. Observers then repeatedly chose between targets differing in eccentricity and contrast, selecting the one they would prefer to attempt: e.g., a low contrast target at 2° versus a high contrast target at 10°. Could observers correctly trade off contrast and eccentricity and pick the more discriminable of the two targets? We found that observers exhibited large, patterned errors in their choices, making choices that were not even self-consistent.

of probability commonly reported in the decision under risk literature [11–13].

The decision task is an example of a conjoint measurement task [14–15]. We vary contrast and eccentricity and see how these variations affect the observer's ordering of eccentricity-contrast pairs $(c, e) > (c', e')$ by discriminability. If the observer's judgments satisfy certain conditions that, in effect, assess their coherence or self-consistency, then the experimenter can potentially recover estimates of the observer's "subjective" retinal sensitivity curves for each contrast [14–15] and compare them to observers' actual performance. If one or more of the conditions fail, then we further conclude that the observer's choices are not based on a coherent model of their own retinal sensitivity.

We test the observer's knowledge of his own ability to discriminate such targets in three ways, illustrated in Figure 2A. The first is a test of *equivalence*: Can observers correctly judge which pairs (c, e) and (c', e') are equally discriminable? We can represent these pairs by indifference curves as shown in Figure 2A.

The second test is *transitivity*: for any choice of eccentricities e, e', e'' and contrasts c, c', c'' , if $(c, e) > (c', e')$ and $(c', e') > (c'', e'')$, then $(c, e) > (c'', e'')$. Transitivity is a test of the self-consistency or coherence of observers' judgments.

The third is a test of *dominance*: if $c > c'$, does the observer correctly judge that $(c, e) > (c', e')$ for any choice of eccentricity? And if $e < e'$, does the observer correctly judge that $(c, e) > (c', e')$ for any choice of contrast c ? (Of course, we must verify experimentally that the two dominance claims are in fact true for our experimental conditions). Dominance is evidently the weakest of the three tests.

The three tests are distinct: an observer who fails equivalence may still satisfy transitivity and dominance. This outcome would imply that, while his or her estimates of discriminability are in error, the estimates he or she has do, at least, cohere. An observer who fails transitivity cannot trade off contrast and eccentricity in any consistent way, but he or she may still know that more contrast improves performance and that performance near the fovea is better.

Methods

Ethic statement

The experiment had been approved by the University Committee on Activities Involving Human Subjects (UCAIHS)

of New York University and informed consent was given by the observer prior to the experiment.

Apparatus

Stimuli were displayed on a 19-in. Sony Trinitron Multiscan G500 monitor controlled by a Dell Pentium D Optiplex 745 computer. The monitor was run at a frame rate of 100 Hz with 1280×1024 resolution in pixels. A forehead bar and chinrest were used to help the observer maintain a viewing distance of 57 cm. At that distance, the full display subtended 40.4°×30.3°. The observer viewed the display binocularly.

Monitoring fixation

Observers were required to fixate a fixation cross and all stimuli were presented relative to this fixation cross. We used an Eyelink II eye tracker to verify that observers did not make eye movements away from the fixation cross. At the beginning of each trial drift correction was made at the fixation cross. The criterion of eye movement was set to be a speed over 10 deg/s or an offset over 1 deg from the fixation cross. A trial would be cancelled if the fixation constraint were violated during the trial. The eye tracker was calibrated initially, drift corrected for each trial and recalibrated after every 100 trials or when drift exceeded 5 deg.

Stimuli

Stimuli were presented against a uniform gray (39.1 cd/m²) background. The fixation cross was black, spanning 0.6°×0.6° at the center of the screen. The target was a 1°×1° lighter gray (67.1 cd/m²) square with an even lighter gray dot of 0.16° diameter at its top or bottom. The luminance of the dot could be 74.4, 80.7, or 91.4 cd/m², i.e., a contrast of 1.11, 1.20 or 1.36 relative to the square. We refer these three levels of contrast as low, medium, and high contrast. The contrast c of a stimulus and the eccentricity at which it was presented, formed an *eccentricity-contrast pair* (c, e) .

Color codes and location cues. Each contrast was associated with a colored cue, which was a filled circle of 0.6° diameter behind the fixation cross. The colors for the low, medium, and high contrast were red, blue and white, respectively. The location of target was cued by a 1°×1° black frame square at the location of the would-be target. Targets or location cues were located at 18 possible locations uniformly distributed in the range of 0.8° to 17.5° to the right of the fixation cross.

Procedure and design

The experiment consisted of two three-hour sessions completed on two successive days. Observers were advised to take a break every about 350 trials and allowed to take breaks whenever necessary. Each observer went through the two tasks in sequence: calibration, then decision. The time courses of both tasks are illustrated in Figure 2B.

Calibration task

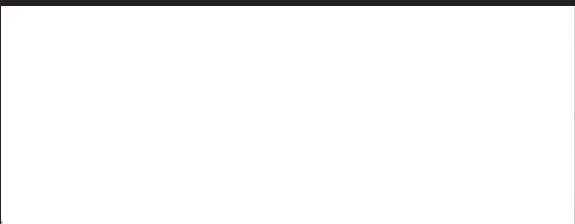
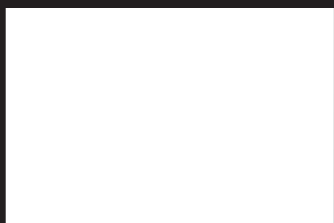
The calibration task allowed us to map probability correct as a function of eccentricity for each of the three contrasts. The observer's task was to decide whether the dot was at the top or at the bottom (Figure 2B). Fixation was monitored. No feedback was given.

For each of the three contrasts, the target could appear at any of 18 possible locations, i.e., 18 possible eccentricities, evenly spaced from 2° to 12.2° on the right of the fixation cross. There were five blocks, in each of which each location of each contrast repeated for six times, half top and half bottom, randomly mixed together. Each observer completed 3 contrasts×540 trials = 1620 calibration trials.

Before the experimental trials, there were 108 practice trials for the first session and 12 practice trials for the second session. To keep observers motivated, we rewarded observers with a bonus up to \$10 based on their overall probability correct in the calibration task of each session.

The probability correct of the calibration task was fitted against eccentricity with a Quick-Weibull psychometric function [16–17]:

$$\psi(e) = 0.5 + 0.5 \exp(-e/\alpha)$$



Observers learned the association between targets and cues at the beginning of the experiment during the calibration task and we verified that they had learned these associations by a short “quiz” before the decision task.

Observers knew that, at the end of the experiment, four of their choices would be chosen at random and that they would attempt to identify targets in the conditions corresponding to each choice. A correct response would lead to a \$5 reward. Correct response for all of the four trials would result in a \$20 bonus.

To measure the point of subjective indifference (equivalence) between targets that differed in contrast, we used one-up, one-down adaptive staircase procedures. In a staircase, one target of one contrast was fixed in eccentricity and the target of the other contrast varied in eccentricity. The fixed contrast in each staircase had an eccentricity corresponding to a probability correct of 0.6, 0.7, 0.8, or 0.9 separately estimated for each observer based on their calibration data. We estimated the eccentricity that the observer considered to be equally discriminable for the variable contrast. Each staircase consisted of 70 trials. There were 12 staircases (3 contrasts \times 4 probabilities), randomly interleaved. That is, 12 staircases \times 70 trials = 840 staircase trials. Based on these staircase trials, we tested equivalence and transitivity.

To test dominance, we included trials in which the two targets had different eccentricities but the same contrast (*equi-contrast trials*), or different contrasts but the same eccentricity (*equi-eccentricity trials*). The possible contrasts were low, medium, and high. The possible eccentricities were the eccentricities corresponding to a probability correct of 0.75 for each of the three contrasts, computed with the functions estimated in the calibration task for the particular observer. The number of equi-contrast trials was 3 contrasts \times 3 eccentricity combinations \times 10 repetitions = 90. The number of equi-eccentricity trials was 3 eccentricities \times 3 contrast combinations \times 10 repetitions = 90 as well.

The 840 staircase trials and 180 dominance trials were mixed in a random order, divided into three blocks and completed in the second session after the calibration task. There were 24 practice trials before the formal experimental trials.

Observers

Eight observers, four female and four male, participated. None of them was aware of the purpose of the experiment. All observers had normal or corrected-to-normal vision. The observers each received US \$12/hour for their time and a performance-related bonus. Total payment ranged from US \$87 to US \$112 across observers.

Results

Visual sensitivity curves

For each observer, we fit the data of the calibration task to Equation 1 separately for each contrast using the maximum likelihood method. Figure 3 shows both the data and fit for each observer.

Equivalence test

From the 12 staircases of the decision task, we acquired 12 pairs of eccentricity-contrast pairs judged to be equally discriminable ($(e, c) \sim (e', c')$) by the observer. Four of them had the target of low contrast in fixed eccentricity and the target of medium contrast in variable eccentricity, which we call a low-to-medium mapping. Another four staircases were medium-to-high mappings and a third set of four high-to-low mappings.

0° 6° 12° 18° 0° 6° 12°

correct for a particular contrast, like the true probability correct, is a function of eccentricity in the form of Equation 1. An equivalence transformation from one contrast to another contrast would then be linear on a log scale (see Text S1 for proof).

$$\log e' = a_{1 \rightarrow 2} \log e + b_{1 \rightarrow 2} \quad (3)$$

where e is the eccentricity for contrast 1, e' is the eccentricity for contrast 2, and $a_{1 \rightarrow 2}$ and $b_{1 \rightarrow 2}$ are parameters to be estimated.

If Equation 2 is satisfied, we should have

$$\begin{aligned} a_{L \rightarrow M} a_{M \rightarrow H} a_{H \rightarrow L} &= 1 \\ b_{L \rightarrow M} a_{M \rightarrow H} a_{H \rightarrow L} + b_{M \rightarrow H} a_{H \rightarrow L} + b_{H \rightarrow L} &= 0 \end{aligned} \quad (4)$$

Define $A = a_{L \rightarrow M} a_{M \rightarrow H} a_{H \rightarrow L} - 1$, $B = b_{L \rightarrow M} a_{M \rightarrow H} a_{H \rightarrow L} + b_{M \rightarrow H} a_{H \rightarrow L} + b_{H \rightarrow L}$. Testing for failure of transitivity requires only that we test whether either of A and B is significantly different from zero.

For each observer, we fitted Equation 3 separately for the low-to-medium, medium-to-high, and high-to-low transformations. With the estimated $a_{(\cdot)}$'s and $b_{(\cdot)}$'s we computed A and B . We obtained the 95% confidence intervals (Bonferroni corrected for two conditions) of A and B using a bootstrap method [18] by resampling the staircase data for 10,000 times.

Only one observer (S5) passed the transitivity test. The remaining seven observers' mean A and B values were both significantly different from zero. Interestingly, all the seven observers' deviations had the same direction. That is, all the A 's were less than zero (median across observers = -0.60). All the B 's were greater than zero (median across observers = 1.06). If, for any observer, A and B errors were independent and equally often positive or negative, the probability for all the seven observers to

have a less-than-zero A and a greater-than-zero B would be $(1/4)^7 = 6 \times 10^{-5}$. Therefore, the observed common pattern of failure of transitivity is unlikely to be the result of measurement error.

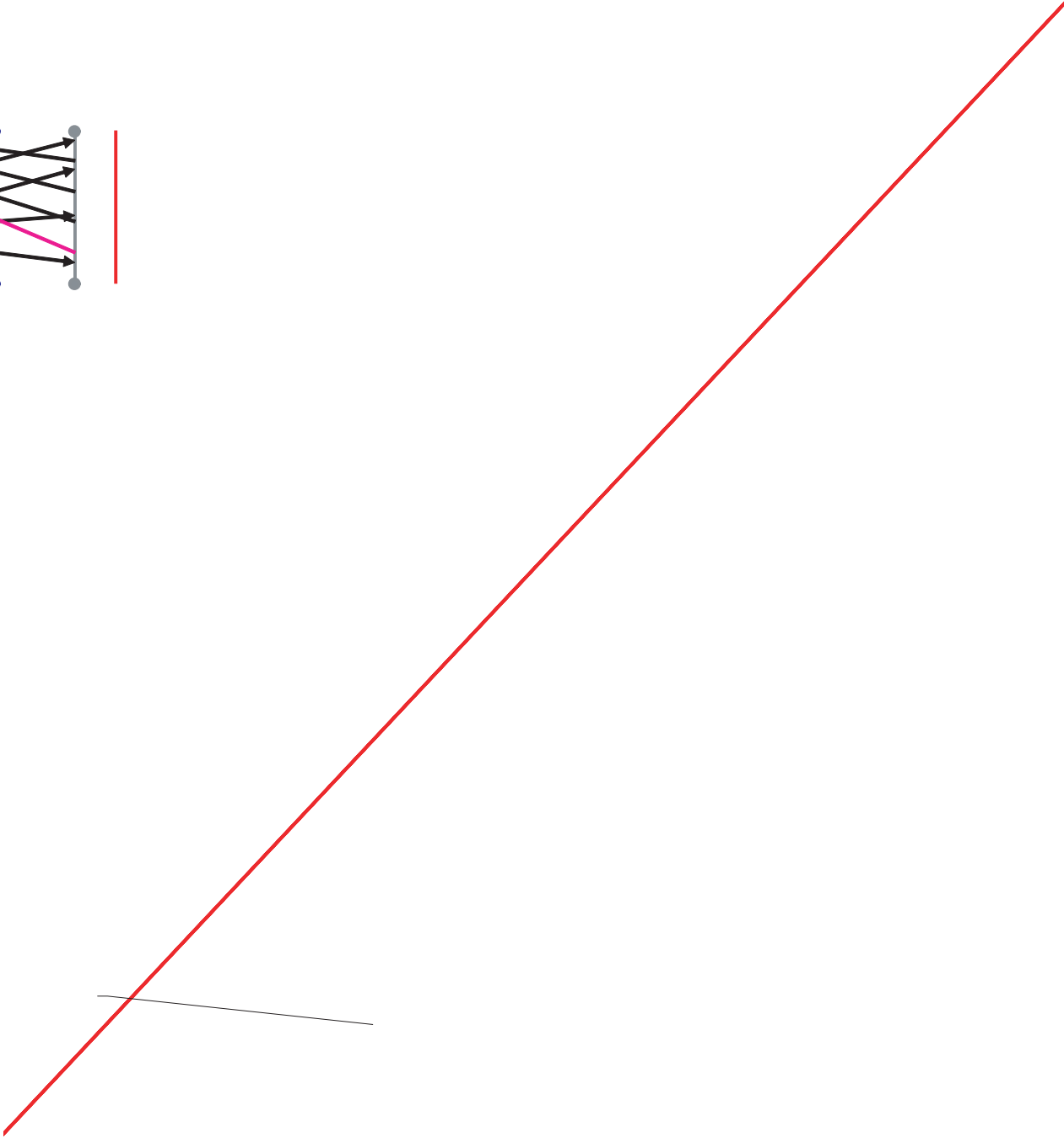
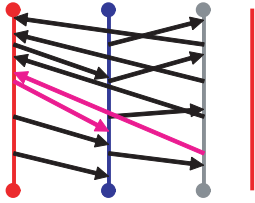
Figure 4B shows a sequence of transformations. The three axes in each subplot represent the eccentricities of the low, medium, and high contrasts in the log scale. For each observer, we start from a specific eccentricity at the low contrast find the equivalent eccentricity at the medium contrast, then we pass from medium to high and then high to low. If the transformations satisfy transitivity, we should return to the same eccentricity at the low contrast axis after going through the three transformations, low-to-medium, medium-to-high and high-to-low. If transitivity holds, we stop after one set of three transformations (low \rightarrow medium \rightarrow high \rightarrow low). If it does not we continue with a second set of three transformations to make the pattern of intransitivity easier to visualize.

Figure 4B illustrates the transitivity failure of seven out of eight observers and their common pattern of failure. We move from one axis to another axis in an arbitrary counter-clockwise way. Note that all the observers that failed the transitivity test had plots that tended to "corkscrew" outward. That is, when eccentricity is transformed from low contrast to medium contrast and then to high contrast, the resulting eccentricity difference between the low and high contrasts tended to be larger than when they mapped from low to high directly.

Dominance test

Observers failed the equivalence test and, with one exception, the transitivity test. The dominance test is, in conjoint measurement terms, a test that observer's preferences form a weak order satisfying single cancellation [15]. We are asking whether observers, given two targets of equal contrast at different

Figure



eccentricities, judge the target with smaller eccentricity to be more discriminable (equi-contrast dominance) and that, given two targets at the same eccentricity, judge the target with higher contrast to be more discriminable (equi-eccentricity dominance).

Figure 4C show the percentage of dominance errors for each observer. For each observer and condition, we computed the 95% confidence intervals for the percentage of errors by treating the true proportion of errors as a random variable with a beta distribution whose parameters are determined by the observed numbers of errors and non-errors. Although the percentage of errors was significantly larger than zero for most of the observers in either condition, the values were small. The medians across observers were 8.3% and 11%, respectively for the equi-contrast trials and equi-eccentricity trials. The upper limits of all the confidence intervals were far below 50%, the chance level.

Discussion

We employed a simple decision task with a conjoint measurement design to investigate what people know about their own visual uncertainty across the retina. In this task, observers were asked to judge which of two eccentricity-contrast pairs $(.c)$ or $(',c')$ was more discriminable. We measured the observer's ability to discriminate targets varying in contrast and eccentricity separately in a calibration task. Consequently, we could determine whether the observer correctly judged which of the two eccentricity-contrast pairs was more discriminable. We found that observers' judgments exhibited large, patterned errors.

Observers may err in judging equally-discriminable pairs, but be self-consistent in their erroneous judgments. We tested whether observers' judgments were transitive. An observer's judgments are transitive if and only if, for all choices of eccentricities $, ', ''$ and contrasts c, c', c'' : $(.c) > (',c')$ and $(',c') > ('',c'')$, then $(.c) > ('',c'')$. Seven out of eight observers failed to be transitive, exhibiting large and patterned errors.

The last test, dominance, assessed whether the observer would choose the eccentricity-contrast pair with smaller eccentricity if contrasts were the same or the eccentricity-contrast pair with larger contrast if eccentricities were equated. An observer need only have a crude sense that higher contrast leads to better performance and that performance is better at smaller eccentricities, at least for our choice of stimuli. In particular, an observer can "pass" dominance without any ability to trade off the

We consider a very similar task but in a different domain: judgments of retinal sensitivity as a function of contrast and eccentricity. In contrast to performance in these visual and motor tasks, however, our observers not only had markedly distorted representations of their retinal scaling functions for targets differing in contrast but also made choices that violated transitivity.

The observed failure to correctly judge visual sensitivity across retinal positions agrees well with reports in other areas. Patients suffering from scotomas (a retinal area with reduced visual acuity) are typically unaware of the scotoma even when testing reveals near total loss of visual sensitivity outside the fovea [27]. Most patients with a central scotoma prefer to use their left visual field to read, although the right visual field is found to be more efficient in reading than the left visual field [28–29]. Galvin & Williams [30] noted that, while objective visual performance in many tasks plummets with distance from the fovea, human observers seem to experience a retinal field that is unblurred and more or less uniform. People were found to have underconfidence and overconfidence at the same time for visual discrimination performance of stimuli of different size [31].

Our results suggest that people might have difficulty in integrating the uncertain visual information from across different retinal eccentricities to speed search. In fact, people are reported to be suboptimal at choosing where to saccade [32–35].

One possibility is that observers have inaccurate estimates of retinal eccentricity [36–37], which make precise mapping between eccentricity and probability of correct impossible. But the observers' transitivity failures suggest the failures are more profound: people likely do not have consistent estimates of visual sensitivity at all.

Heuristic-based visual search. Our results are in apparent conflict with the results of Najemnik & Geisler [7,9–10], which show good human performance in selection of saccades in visual search. One possibility is that the visual system has accurate information concerning visual sensitivity to different targets as a function of eccentricity but that this information is unavailable for the sort of comparative judgments we considered here.

But a second possibility is that human visual search is actually based on simple heuristics plus a qualitative understanding of one's visual sensitivity map. Such a heuristic-based approach may approximate ideal performance in some tasks while failing utterly in others. The experimenter who considers performance in a limited range of scenes may record behavior that approximates optimal but is in fact no more than a lucky coincidence of a heuristic rule and experimental conditions. Such "apparent optimality" is not rare in behavioral studies of animals [38] or humans [39].

The task used by Najemnik & Geisler [7,9–10] involved detection of a Gabor patch in a $1/f$ field of noise and only the

overall pattern of saccades was considered in evaluating the model. In contrast, we designed our simple task so that the visual system can *only* succeed if it has access to estimates of visual sensitivity for the range of contrasts and eccentricities we considered.

Dominance was the only test where subjects predominantly succeeded and their success could be due to a preference for higher contrast [34] or a preference for locations closer to the fovea [32]. According to the errors in the equivalence test, among the seven observers that failed transitivity, five observers' decisions could result from a bias toward selecting the nearer target. In the context of eye movements, this bias would correspond to a preference for shorter saccades over longer.

We conjecture that this preference for short saccades could be an oculo-motor heuristic serving to integrate the visual sensitivity map into saccade selection. A key prediction of Najemnik & Geisler's model [7,9–10] is exactly that observers will prefer short saccades. Tatler and Vincent [40] presented compelling evidence that saccade selection could be better predicted by oculo-motor preferences than by visual information or task.

If human saccade decisions are based on such heuristics rather than on a computation that requires knowledge of visual sensitivity maps, we would expect a failure of adjustment when one's visual sensitivity map is changed. In fact, when observers' foveae were artificially shifted with gaze-contingent techniques, their performances in visual search were significantly worse than predicted by the ideal-observer model [41].

There is evidence that saccade selection and explicit perceptual decision in visual search pick the same location [42–43]. This was previously understood as evidence that saccade selection and explicit decision use the visual sensitivity map in the same way. However, if humans have little knowledge of their own visual sensitivity map, as our results suggest, and their saccades are chosen through oculo-motor heuristics, it might be their explicit decision actually comes from their saccade behavior. If so, the observer might have no way of finding out which combination of eccentricity and contrast offered the higher probability of successful detection than by observing his preferences among potential saccades, something like deciding what to have for lunch by waiting to see which sandwich your hand selects.

Supporting Information

Text S1 Proof for Equation 3.

Found at: doi:10.1371/journal.pcbi.1001023.s001 (0.06 MB DOC)

Author Contributions

Conceived and designed the experiments: HZ CM LTM. Performed the experiments: HZ. Analyzed the data: HZ. Wrote the paper: HZ CM LTM.

References

- Curcio CA, Sloan KR, Kalina RE, Hendrickson AE (1990) Human photoreceptor topography. *J Comp Neurol* 292: 497–523.
- Pelli DG, Tillman KA (2008) The uncrowded window of object recognition. *Nat Neurosci*, pp 1129–1135.
- Legge GE, Kersten D (1987) Contrast discrimination in peripheral vision. *J Opt Soc Am A* 4: 1594–1598.
- Virsu V, Rovamo J (1979) Visual resolution, contrast sensitivity, and the cortical magnification factor. *Exp Brain Res* 37: 475–494.
- Johnston A (1987) Spatial scaling of central and peripheral contrast-sensitivity functions. *J Opt Soc Am A* 4: 1583–1593.
- Farrell JE, Desmarais M (1990) Equating character-identification performance across the visual field. *J Opt Soc Am A* 7: 152–159.
- Geisler WS, Perry JS, Najemnik J (2006) Visual search: The role of peripheral information measured using gaze-contingent displays. *J Vis* 6: 1, 858–873.
- Renninger LW, Verghese P, Coughlan J (2007) Where to look next? Eye movements reduce local uncertainty. *J Vis* 7: 6, 1–17.
- Najemnik J, Geisler WS (2005) Optimal eye movement strategies in visual search. *Nature* 434: 387–391.
- Najemnik J, Geisler WS (2008) Eye movement statistics in humans are consistent with an optimal search strategy. *J Vis* 8: 1–14.
- Kahneman D, Tversky A, eds. *Choices, values, and frames*. Cambridge, UK: Cambridge University Press.
- Gonzalez R, Wu G (1999) On the shape of the probability weighting function. *Cogn Psychol* 38: 129–166.
- Luce RD (2000) *Weighting functions. Utility of gains and losses: Measurement-theoretical and experimental approaches*. London: Lawrence Erlbaum, pp 84–108.
- Luce RD, Tukey JW (1964) Simultaneous conjoint measurement: A new type of fundamental measurement. *J Math Psychol* 1: 1–27.

15. Krantz DH, Luce RD, Suppes P, Tversky A (1971) Foundations of measurement, vol. 1: Additive and polynomial representations. New York: Academic Press.