Dear Professor Melvyn Goodale,

We appreciated your and the reviewers’ helpful and clear comments and suggestions for the manuscript we submitted to Proceedings of the Royal Society of London: B. We appreciate the overall positive assessment of our study, and have made revisions and refinements to the manuscript as suggested by the reviewers. These have been highlighted in the text. We feel the manuscript has greatly improved as a result and we look forward to your decision and any further comments.

Thank you once again for considering our manuscript for publication in Proceedings of the Royal Society of London: B. Our replies to reviewer comments are below.

Sincerely,

Anna Nowakowska

Responses to Reviewers:

Associate Editor, Professor Melvyn Goodale  
Board Member: 1  
Comments to Author:  
We now have comments from two reviewers -- and both are positive (one more than the other).  I think the manuscript would be stronger if you could address the 'why' question more thoroughly.  I am recommending that you revise manuscript before we move forward.  Please attend carefully to the comments of the two reviewers, particularly reviewer 2.  
  
  
Reviewer(s)' Comments to Author:  
  
Referee: 1  
  
Comments to the Author(s)  
This is excellent behavioral work demonstrating that visual search is much less efficient than a rational mind would imply.  It follows on the heels of previous work from this group showing a similar effect in locomotion (Clarke, A.D.F. & Hunt, A.R. Psychol Sci. 2016; 27: 64-74).  
  
The empirical work is elegant, the results are dramatic, and I’m fully in favor of publication in some form.  Researchers who study eye movements, visual search, and scene perception will surely want access to these data and these ideas.

We are grateful for these very positive comments.   
  
Minor revisions.  The work could be better situated in the previous literature in several respects:  
1.    research on eye movements in reading went over similar ideas in the 1970’s and continue to this day, arguing whether eyes are guided by the visual form, semantics, phonetics etc and asking what are the time course of these effects.  The null hypothesis there is that eye movements in reading follow a random walk (sound familiar?) and sometimes there is support for it, though more often the evidence favor some kind of guidance.

Thank you for pointing this out. We were not aware of this parallel debate in the reading literature. We’ve added references in the introduction to Fend (2006) and Engbert and Kleigl, 2001, who both propose stochastic models for eye movement control in reading, as well as to reviews of models of eye movement control by Rayner & McConkie (1975) and Rayner (1979).

2.    Ray Klein (Klein, R., & Farrell, M. (1989). Search performance without eye movements. Perception & Psychophysics, 46(5), 476-482.)’s work hinted at the inefficiency of eye movements in search long ago, in their comparison of search with and without eyemovements (a brief flash).   
3.    Others studying strategies in visual search noted that naïve observers (and those specifically instructed to search “actively” made many completely unnecessary eye movements

(Watson, M.R., Brennan, A.A., Kingstone, A., Enns, J.T. (2010). Looking versus seeing: Strategies alter eye movements during visual search. Psychonomic Bulletin & Review, 17(4), 543-549.)

We’ve added these references to the discussion of Experiment 1, and appreciate this reviewer drawing them to our attention.  
  
But on balance, this is very nice work that chips at an important theoretical framework in the literature on visual search (Najemnik & Geisler). What’s to replace it?  Well, this study doesn’t say, but it lays the groundwork for the debate  to surely follow.

Minor revisions.  In that vein, I would welcome a bolder proposal with respect to the “why?” of random walk processes in search.  If the authors were to take a page from the older reading debates it might along the lines of random walks are “good enough” for dumb, slave-like implicit processes that don’t take much energy (like moving the eyes around a scene).  But of course, that would beg the question of why the same thing happens when people are able to contemplate and move their bodies more deliberately, as in Clarke et al. Psychol Sci. 2016.  Its not clear how the account there would help in the present demonstration of similar phenomenon with unconscious eye movements (i.e., from the abstract:  “This failure may have been related to uncertainty about their own ability, because in a version of the experiment in which there was no uncertainty, participants uniformly switched at an optimal point.)

We were perhaps too subtle here: We believe the stochastic search model described in the introduction is a reasonable alternative to the ideal searcher. The basic idea is that selection of each saccade (and perhaps many other decisions) is “random with constraints” - the constraints may be biological limitations or learned/evolved biases or tendencies that make some saccades and sequences of saccades more likely than others. Taking eye movements as a whole, without regard for individual differences, they do not seem to be driven preferentially to locations that produce the most information, but just select a saccade direction seemingly at random. That said, when you take the variability across individuals into account, it looks like different models might be required to fit different individuals. Perhaps a few participants would match an ideal observer model, but the majority do not. We have added to our discussion here to make our conclusions clearer and help to frame the ensuing debate.  
  
  
  
Referee: 2  
  
Comments to the Author(s)  
Nowakowska, A., Clarke, A. D. F., & Hunt, A. R. (2016). Human visual search behaviour is far from ideal. Proceedings of the Royal Society B, ms.  
  
In this paper, Os see displays divided into an easy search on one side of the screen and a hard search on the other. Targets are present on 50% of trials. The authors argue that Os should not fixate into the easy search field if they are being ideal. They should go collect info from the hard side. However, they do go to the easy side almost half the time; hence, they are not ideal.  
  
Well, I am biased to believe this because I think that the Najemnik & Geisler information foraging ideas are really only going to work for a very simplified set of stimuli (e.g. where you cannot tell where an item is, let alone a target…as in their lovely experiments). That said, I am not sure how excited I should be by these results. Why not?  
  
1)    The side of the hetero stimuli is randomized – right? So, suppose I always pre-program my first saccade into the middle of one side, at random. Half the time, I will guess right and go into the hetero side but the other half will go to the ‘wrong’ side. What does the timing of that first saccade look like?

This explanation suggests the second (and third, fourth, etc), rather than the first, saccade would be directed to the heterogeneous side, as these would certainly be stimulus-driven. But you can see from Figure 2 that subsequent saccades don’t, on balance, do much better than the first one. To directly address the question about timing though:  When the target is absent it takes participants around 307 ms to start moving their eyes. This time is similar when target is present on the hard side (300ms) and a bit faster when it is present on the easy side (284 ms). The timing suggests these saccades are stimulus-driven at least most of the time, as does the fact that they’re faster when the target is present on the easy side.   
  
2)    There is no penalty for making a fixation in this experiment, right? (Well, yes there is a 350 msec RT cost but I, the observer, really really do not care about a fraction of a second….maybe I should….but I don’t). So, let’s make a fixation to the target, when present, just to confirm its presence. Why not?

To address this last question first, we’d like to emphasize that the data we use in analyzing proportion of saccades directed towards one side or the other come only from target absent trials. None of them are target-directed because there is no target. We’ve clarified this in the text. As for the first point (do observers care about RT), we are testing a theory (cited 500+ times, discussed in many textbooks) that is proposed as the mechanism for how we search, irrespective of how much we care about the outcome.  
  
3)    That would be two homo-side fixations on many trials. And that is what you find in your data, right? Where do those two fixations go anyway? Some sort location information might be useful. It would be useful to know if the data  include saccades to the target. I don’t think those were excluded, were they?

As noted above, target present trials were excluded to avoid having to figure out which saccades might be driven by the target. We have a clear hypothesis as to what an optimal visual system would do when the target is absent.

4)    Similarly, on absent trials, there is no obvious penalty, so if I didn’t find anything on the hetero side, why not put a fixation on the homo side before I quit? Might be a good idea….in general.

We agree this isn’t a bad idea, but the data don’t really bear this out as the dominant strategy (see Figures 3 and 4). Many of our participants explore the homogeneous side first and only then thoroughly explore the heterogeneous side.

5)    Suppose you showed the stimulus for 300 msec and explicitly told Os that you are only going to get one fixation. And suppose fixations on the target didn’t count. Now what do you think the results would look like?

We have some basis to believe the results would look similar: Morvan and Maloney (2012) demonstrated that people do not use single saccades (even though it is the only one they can make to identify the target) in an optimal way.

But more generally, this question is getting at the idea that perhaps people are just making more eye movements than they need to. Experiment 3 was designed to test exactly this. That is, we wanted to see if people’s eye movements are as reasonable as could be expected in the split screen displays, given that people are inefficient with their eye movements or are driven to satisfy a overly conservative certainty criterion. But we find that even when you take an individual’s inefficiency in searching uniform displays into account, they still perform worse than you would predict on the split displays.  This is why we think that we would not make participants optimal by allowing them to make just one eye movement.   
  
6)    P4 I don’t have a problem with the sample size and I think I know what you are saying but this line is confusing “Previous seminal experiments on this topic had a very small numbers of participants (e.g. N=2 in (6); N=4 in(9)) but report results from individuals separately rather than averaging them. Our sample is larger, but we maintain the approach of reporting individual differences (as in (10)), making a very large sample impractical and unnecessary.”

We’ve simplified this sentence and hope it is clearer now.  
  
7)    Fig 2: Distributions are clearly different but could you give some mean RTs or something?

Means and SD for the three conditions are now included in text.

Parallel 1.75( ± .13)

Serial 3.94 (± 2.19)

Absent 7.00 (± 4.51)  
  
8)    Why are Os “close to 100% correct with no eye movements and only 87% with eye movements in Exp 2?

Thanks for pointing this out - we hadn’t broken these numbers down into all the relevant conditions. We’ve added a table to present the full summary of RT and accuracy results across conditions.

|  |  |  |  |
| --- | --- | --- | --- |
| Target Side | Reaction Time (±SD) | Accuracy (±SD) |  |
| Heterogeneous full | 3.28 (2.23) | 55.00 (20.55) | |
| Homogeneous full | 1.77 (.13) | 97.32 (3.32) | |
| Heterogeneous split | 3.42 (2.42) | 48.10 (23.99) | |
| Homogeneous split | 1.84 (.17) | 97.42 (4.60) | |
| Absent split | 6.03 (3.28) | 95 (6.36) | |
| Absent heterogeneous full | 6.94 (4.55) | 93.39 (6.09) | |
| Absent homogeneous full | 2.84 (.73) | 97.86 (3.91) | |

P11 Mean number of fixations seems low (for example, 4 fixations in 2.15 seconds). Any reason?

Thank you for pointing out this! There was an error in the reporting these. We thoroughly checked the numbers and they are as follow (this has also been rectified in the text).

Mean number of fixations on target absent trials:

Heterogeneous array 21.64 (±15.36)

Homogeneous array 7.57 (±2.81)

Split screen array 18.82 (±11.42)

These numbers underscore the general inefficiency of search, as the number of fixation on split screen arrays should be half the numbers on the full heterogeneous array (no eye movements are necessary to inspect homogeneous display). Yet participants are making 8 additional fixations on the split screen trials, presumably to inspect the homogeneous side.   
  
10)    I assume you are up against some strict word count for the journal. If not, I would put the pilot data in the main body of the paper. It would make the paper clearer.

It was not a strict word count that caused us to choose to put the pilot data in the supplementary material. We just felt the pilot data divert the reader from the straightforward pair of experiments we would like to present. They simply establish that our stimuli are fit for purpose. We have added more detail about this experiment to the methods description in Experiment 1, but left the full description in the Supplementary material.  
  
So, it isn’t that I disagree with the findings or the conclusions but, I do wonder what it means to be less than ideal when the cost is so small and the behavior is sensible, in general.

If somebody wanted to build a model of visual search it would be nice to know what people are guided by. If someone is trying to build a model our data is a nice starting point. The data points out that the model of visual search is neither random nor optimal.  
  
Best  
  
Jeremy Wolfe, Signed review