Penguins are cool

Warren JAMES,

Dr. Josephine REUTHER,

Ellen ANGUS,

Dr. Amelia HUNT,

and

Dr. Alasdair CLARKE

Abstract

Contents

Introduction

Everyday we make use of our visual system in order to solve many problems. Whether these are to aid us in avoiding obstacles, or to find an item of importance, the ability to direct our attention proves invaluable. However, the extent to which we make these decisions in an optimal manner is subject to debate.

Najemnik & Geisler (2005) argue that human performance in a visual search task is predicted by that of a model ideal observer. This model aimed to reduce uncertainty as to the location of a target item by deploying fixations to locations likely to contain the target and reduce the overall uncertainty of the search space. To do this, the model accounted for the loss of acuity for space further from the central fixation and using this information to reduce the amount of uncertainty to an appropriate extent. In accounting for visual acuity, the model would be less likely to revisit locations that were already inspected to an appropriate extent.

This would suggest that we make decisions about where to fixate based on underlying scene statistics and account for our own limitations. However, this would seem to be a fairly resource intense process, which begs the question as to whether this would be an appropriate solution to the problem. With this in mind, Clarke, Green, Chantler, & Hunt (2016) suggest that human performance in a visual search task can be characterised by a stochastic process. Instead of engaging in a resource intense calculation of where to fixate, they suggest that we make eye movements that we like to make in a more stochastic process.

As such, this leaves us with the question of whether we make principled fixations based on various factors, or whether we simply do what we like to do and fixate locations in a more random fashion. One way in which this question could be addressed is by simply tasking people with making one sensible eye movement. If we fixate based on principle, then to complete a task such as this should prove to be trivial.

Much to the contrary it appears that people do in fact tend to struggle at making one sensible eye movement (Morvan & Maloney, 2012; Clarke & Hunt, 2016). Morvan & Maloney (2012) developed a paradigm in which participants were to make a single eye movement in order to give themselves the best chance of success at detecting whether a target dot was positioned at the top or the bottom of a box *(point to figure when you have one?)*. To do this, three boxes were presented on screen. One in the centre, with two boxes either side. Participants

were told that they were to fixate one of these boxes and that the target would then appear in one of the two side boxes. Throughout the experiment, these two side boxes would appear at different eccentricities, sometimes being placed very close to the centre box, and at other times they would be a large distance from the central box. When the boxes were all close together, the optimal strategy would be to fixate the central box and use one's peripheral vision to detect the target. However, as the boxes separated, performance would begin to decrease. In this case, participants engaging in the optimal strategy would switch to focussing one of the two side boxes as this would then give them the best chance of success.

The results of their experiment clearly demonstrated that people failed to perform this task in an optimal fashion. These results were further supported by Clarke & Hunt (2016) who replicated this study with the addition of including other versions of the task that were not specific to eye movements. This would suggest that this failure is not constrained to optimal search, but more a general failure to maximise the chance of success.

This same sentiment was demonstrated by Hudson, Maloney, & Landy (2007). In their study, participants were tasked with reaching for a target displayed on screen within a time limit. However, they were to begin their reaching motion before they knew exactly where the target would be. Instead, participants were shown a display of points that could be the target. The display consisted of thin bars made up of many small squares. At the beginning of a trial, some of these squares would turn white indicating they were potential targets. The distribution of these squares could be altered so that they were evenly spread across the bars on screen, or they could be concentrated in one are such as the centre of the display. Participants were then to reach towards the screen and, after crossing an invisible threshold, the target would be made clear by leaving only one square still illuminated. What they found was that when the target was most likely to appear to one side, participants would still begin their reach towards a more central point. This would suggest that participants were opting for a strategy that allowed for a small chance of success at locations that were unlikely at the expense of overall success.

Behaviour such as this is prevalent in many other tasks. Kahneman & Tversky (1984) and Gigerenzer & Gaissmaier (2011) both discuss many ways in which people demonstrate sub-optimal behaviours and some potential explanations for the observed patterns. One potential line of inquiry may be that of motivation. In Clarke & Hunt (2016) participants were offered little in the way of reward for being successful other than the value placed on success by

each participant. It is possible that participants simply had not real motivation to succeed. There are studies suggesting that by adding a financial incentive participants begin to make use of more optimal strategies (Goodnow, 1955; Phillips & Edwards, 1966). By adding a clear motivation for success in the form of a financial gain for success, participants moved from employing strategies that attempted to be successful for all outcomes and began to focus more on those that were more likely.

However, this did not seem to encourage participants to perform more optimally in Morvan & Maloney (2012), which may be due to this effect being more nuanced than simply giving people money for being successful (Camerer & Hogarth, 1999). As such, this does not rule out that participants performed poorly due to a lack of motivation.

Even with this added financial incentive, the goals of a participant may still be maligned with the desires of the experimenter. For example, even though in Morvan & Maloney (2012) participants were rewarded for success, they may have prioritised completing the task as quickly as possible rather than gaining the largest sum of money at the end of the experiment. This could be the result of participants wanting to maximise the rate of gain over time as opposed to solely focussing on the total gain.

Miranda & Palmer (2014) discuss the idea that by using a more gamelike design, it may be possible to encourage participants to prioritise success rather than speed of completion. They posit that by making an experiment a more enjoyable experience, there is an intrinsic motivation to take part in the task as opposed to simply performing the task to gain something external (such as money or course credit). In effect, this could align the goals of the participant and experimenter as participants would find the task more enjoyable and therefore be less likely to try and reduce the overall length of the experiment.

In the present study, we attempted to motivate participants to want to succeed by gamifying the problem. The hypothesis being that participants taking part in a more engaging version of the task would be more likely to complete the task by using a more optimal strategy and therefore enjoy a greater rate of success.

Methods

Participants

Participants were recruited via word of mouth at the University of Aberdeen. In total, 18 participants took part in the experiment (13 female) with an age range of 20-23 (mean of 20).

Procedure

The experiment took part over two sessions. The first session was to measure each participant's visual acuity in order to tailor the second session to each individual. The first session lasted approximately 30 minutes. In the second session, participants carried out the actual decision task, which lasted between 40 and 50 minutes.

Both sessions made use of and Eyelink 1000 (version 4.594) (SR Research ltd, Mississauga, Ontario, Canada) which recorded eye position at 1000 Hz. Participants were sat $\approx 45cm$ from the screen, which was maintained throughout the experiment by use of an adjustable chin rest. *Was this experiment on the OLED?*. The experiment was programmed and run in Matlab 7.9.0 (R2009b) with Psychtoolbox (Pelli, 1997) and EyelinkToolbox functions (Cornelissen, Peters, & Palmer, 2002). Prior to starting each session, a 5-point calibration was carried out. Participants were re-calibrated prior to starting a new block. Additionally, they were also re-calibrated if they had failed to fixate appropriately 10 times since the last calibration, or if there were 5 errors in a row.

Session 1

In the first session, participants were presented with a fixation cross with two boxes placed either side ***get fig***. They were told that their task was to identify whether a small dot ***get size of this dot*** had appeared in the top or the bottom half of one of the boxes (each $\approx 1°$). It did not matter which box it had appeared in, they were simply to report whether it was up or down by pressing the "up" or "down" arrow, respectively, on the keyboard. The dot had an equal chance of appearing in either box and in either position. To commence each trial, participants had to press the spacebar whilst fixating the central cross. After maintaining fixation for 700ms, the target dot would appear in one of the two boxes, at random, and was displayed for 700ms (***Need to double check all this***). If participants broke fixation during any stage of this process, the screen would turn red for 700ms to indicate that they had broken

fixation.

The boxes were presented at varying separations from the central fixation cross (2.7°, 3.9°, 5.2°, 6.8°, 8.4°, 10.1°, 11.4°, 12.5°). There were 4 blocks in total, with each different separation being presenter 12 times in a pseudo-random order. This meant participants completed 4 blocks of 96 trials. Data from this session was then used to work out the point at which participants were 75% accurate in detecting the target. This value was then used in the second session to tailor the experiment for each individual.

Session 2

For the second session, participants were introduced to *"Pugadoo"*. They were told that there task was to help Pugado collect as many fish as possible. In order to do this, they would be choosing a box to fixate in order to detect whether the target dot was in the upper or lower half of one of the two side boxes (***Get fig***). On each trial, participants would see Pugadoo and have to fixate the star on their stomach to start the trial. They would then press the spacebar and three boxes would appear. Participants were told they could fixate any one of the three boxes but that the target would only ever appear in one of the two side boxes at random. These boxes were placed at various distances from the centrally presented box. These distances were defined in relation to the switch point that was calculated based on their performance in the first session. Seven were defined as the switch point in the following way; switch point −3°, −2°, −1°, 0°, +1°, +2°, +3°. A further two distances were added so as to be constants across all participants (8° and 18°).

For every trial that the participants correctly identified the position of the target dot, the background turned green and the bar in the top left corner would fill by one tick. After having filled the bar to the top, Pugadoo was rewarded with a fish and the bar would return to being halfway filled. filling the bar was accompanied by a sound cue to signifying that they had gained a fish. If the participant was incorrect, the background would turn red and the bar would be emptied by one tick on the bar. If the bar was fully depleted, a fish would be taken away from the total at which point there was a separate sound cue signifying the loss of a fish. This allowed participants to monitor their performance while completing the task.

Results

All analysis was carried out in R version 3.4.3 (R Core Team, 2016) using Rstan (Stan Development Team, 2018; Carpenter et al., 2017). Additional data was retrieved from `https://osf.io/t6c5q/` in order to compare the current study's data to a group without the presence of a clear goal, but who otherwise carried out the same task. The full methods and analysis of this comparison data can be found by following the link provided.

For the purposes of this study, the first four blocks completed by participants in this comparison set were used. In the first four blocks of this other study, participants either carried out the standard task as described in Clarke & Hunt (2016), or they were effectively instructed as to where they should fixate on each trial in order to achieve the optimal accuracy (henceforth referred to as the "*control*", and the "*optimal*" group respectively).

Fixation Proportions

On each trial, participants were given three possible locations to fixate. Fixations were coded as being either a central fixation (i.e., they fixated the centre box), or as a side fixation (i.e., they fixated one of the side boxes). A proportion was then calculated for how often the participant fixated one of the side boxes within each of the distances mentioned in the methods section.

From an initial inspection of Figure 1, participants in the motivated group do not appear to be doing anything different to those in the Control group. They certainly do not look as though they have been instructed as to where to look if one compares their results to the Optimal group.

I do have model results for this... but I imagine this can be added to the supplementary material? Also, it's pretty uninformative as to what's happening and doesn't add more to what can already be seen in the figure anyway

Rate of Success

Initial analysis for rate of success examined overall accuracy for each participant. To do this, an average was calculated for each participant which was then modelled using a Bayesian Regression Model with a Beta family. All priors were left as uniform and therefore uninformative.
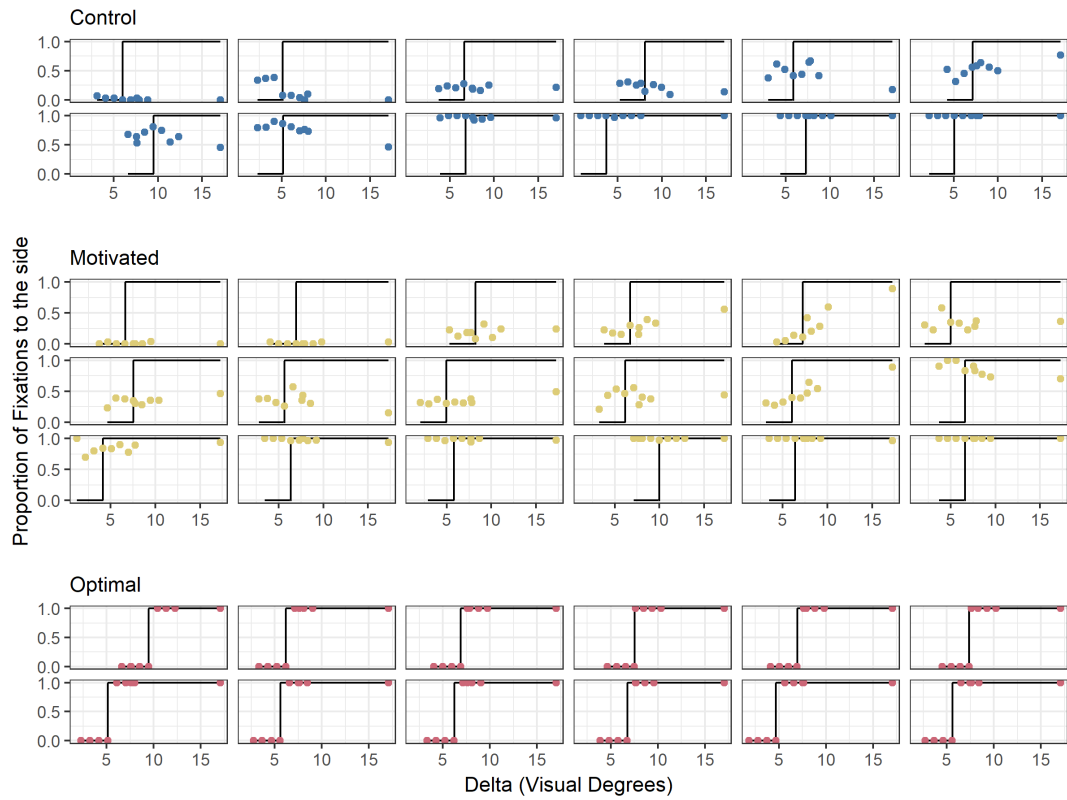
Figure 1: Plot of proportion of fixations to the side box (y-axis) with increasing Delta (x-axis). The bottom set of plots (in red) demonstrate what behaviour would look like if participants had employed an optimal fixation strategy

Rate of success was modelled using group as a predictor variable. Delta was not investigated as we are first and foremost interested in the participant's overall rate of success.

Should the analysis that included Delta be presented in the supplementary material? Pretty sure it does help with the model fit, but it's not entirely necessary for the conclusions of this paper?

In Figure 2, it can be seen that that the Motivated group were generally more successful than the Control group. The Control group had a predicted mean rate of success of 0.708 with a 95% HPDI of $|0.661, 0.747|$, and the Motivated group had a mean of 0.764 with a 95% HPDI of $|0.74, 0.789|$. However, they did not reach the same level of success as the optimal group which had a mean success rate of 0.825 with a 95% HPDI of $|0.805, 0.844|$.

Expected Rate of Success

The results described above appeared somewhat surprising given the fixation proportions that can be observed in Figure 1. Generally, one would not expect the motivated group to have
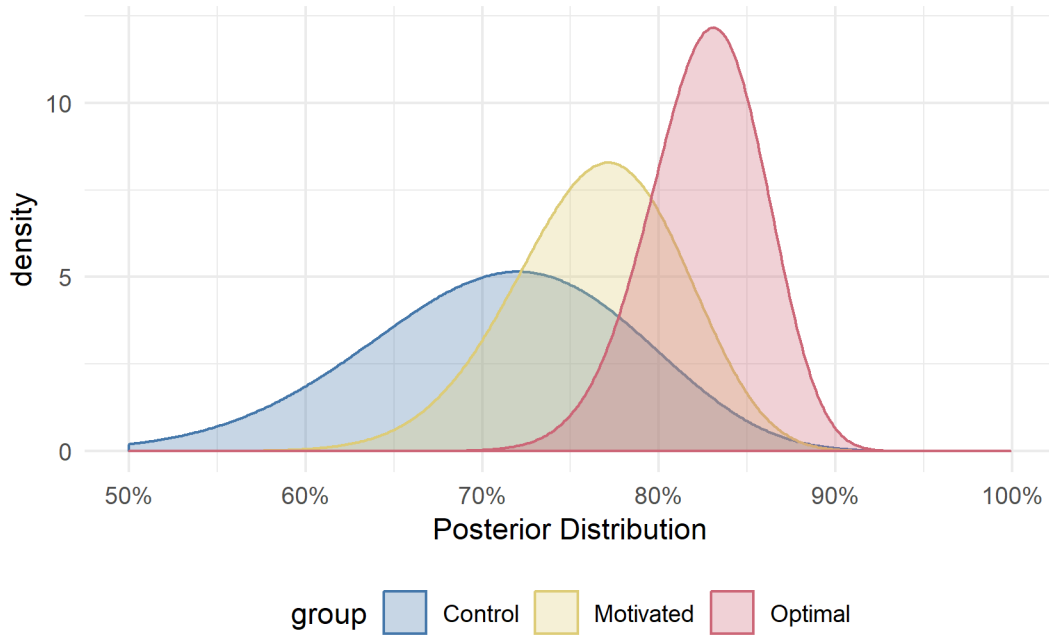
Figure 2: Posterior predictions for rate of success.

performed this much better at the task due to the task being calibrated for each individual participant. As such, additional analyses were carried out in order to make a fairer comparison of the groups. Each participant's expected success rate was calculated given the fixations choices they had made on each trial. For example, if the boxes were far apart and the participant had fixated one of the side boxes on that trial, they would have a $\approx 100\%$ chance of success at the fixated box (b_f) and $\approx 50\%$ chance of success at the non-fixated box (b_{nf}) as this would be imperceptible but they would have a 50% (c) chance due to the nature of the task. Each of these boxes would be equally likely to contain the target and so the expected chance of success on a trial like this would be 75% which is given by calculating $(b_f + b_{nf}) \times c$. This was calculated for each participant, on each trial. This was then averaged for each participant to give an average expected accuracy.

After accounting for chance, the difference that was once observed when modelling participant's raw success rate (Figure 2) appeared to entirely disappear. Inspecting Figure 3 would demonstrate that participants in the Motivated (mean of 0.735, 95% HPDI of $|0.715, 0.754|$) and Control (mean of 0.732, 95% HPDI of $|0.706, 0.755|$) groups almost identical expected success rates as can be seen by the almost entirely overlapping posterior predictions. These results would suggest that the previously observed greater rate of success experienced by the Motivated group was not due to them making better, strategic decisions about where to fixate. Instead, this greater rate of success would appear to be due to something other than a more
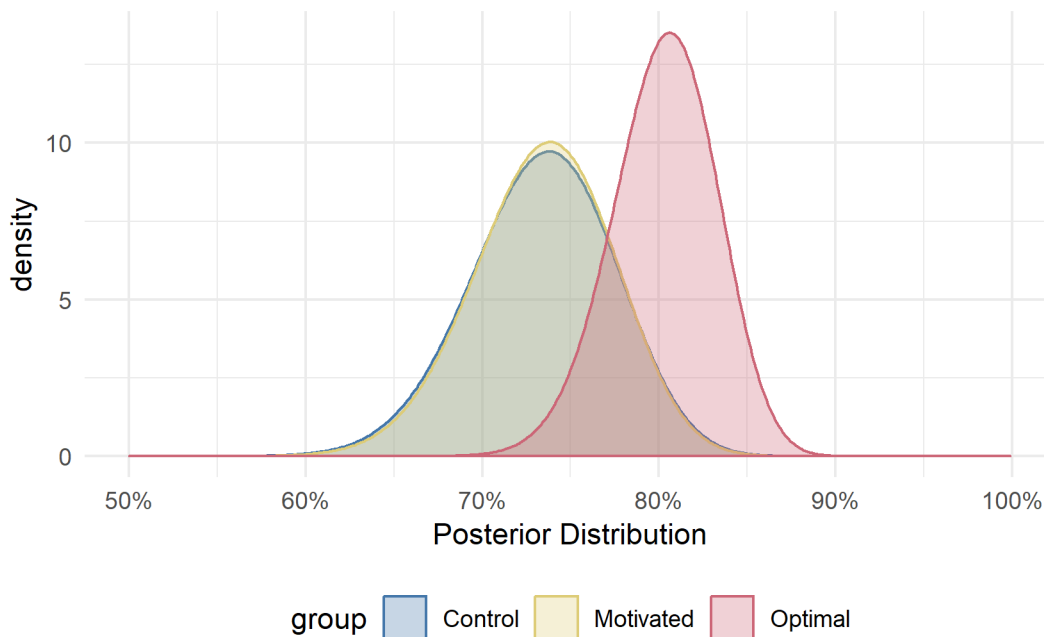
Figure 3: Posterior Predictions for expected success rate.

optimal strategy being utilised.

Discussion

In this experiment, we attempted to encourage the use of an optimal strategy by including a clear motivation for success. Upon examining Figure 1, there would appear to be no difference in strategy between the *Motivated* and *Control* groups. This would suggest that having a clear motivation for success does not cause people to begin to use optimal strategies. This could be expected given the similarity in results of Clarke & Hunt (2016) and Morvan & Maloney (2012) in which the former provided no incentive, and the latter rewarded participants with a monetary sum for each successful trial. The results from this experiment are broadly consistent with previous findings in that participants appear to be quite variable when approaching this task.

An interesting finding is present when looking at the participants' rate of success in each condition. As would be expected, the *Optimal* group had the greatest average rate of success. However, despite very appearing very similar in their strategy, the *Motivated* group appeared to have a much greater rate of success than the *Control* group (Figure 2). To investigate this further, we calculated a predicted rate of success for each participant given their choices throughout the task. After accounting for each choice, this difference all but disappeared. As

such, this increase in accuracy does not appear to have been due to a more optimal strategy use as the *Motivated* group now appeared almost identical in their expected rate of success (Figure 3).

This change in accuracy in the absence of an improved strategy could be explained by participants making a greater effort to be successful. Rather than employing a better strategy, it is possible that participants were instead merely expending more effort to detect the target. In an experiment by Engelmann & Pessoa (2007), they found that participants' sensitivity to detecting a small target increased when a greater monetary reward was associated with success. This idea that previously rewarded stimuli capture attention is also present in Miranda & Palmer (2014) who found that the presence of a distractor that had previously been a highly rewarded target would slow down a participant's response times. As such, this increase in performance could be due to rewarding stimuli capturing attention to a greater extent.

Another factor in how motivation may enhance performance is that of expending effort. Manohar, Finzi, Drew, & Husain (2017) suggest that when a reward is present, participants do in fact expend more effort in the completion of a task. They wanted to investigate whether the mere presence of reward would enhance effort, even when the earning the reward was not contingent on making a greater effort. In their task, participants were instructed to fixate a central cue. They were then informed about how the reward for success would be determined on that trial. A target would then appear either to the left or right of the central fixation. Participants were told prior to each trial how the reward would be determined. There were four ways in which this happened; participants would receive the reward if their response time was in the top 50% of response times, at random, a guaranteed amount, or a guarantee to not receive anything. To compare effort in this task, they measured the time taken to fixate the target. What they found was that participants were much quicker to fixate the target when rewards were contingent on performance and if there was a guarantee of receiving a reward, though the latter was to a lesser extent.

This demonstrates that even in a situation in which extra effort is not strictly necessary, the mere prospect of a reward can cause people to expend more effort in pursuit of the reward. As such, this could perhaps explain the results observed in the current study. Engaging with the task more and expending more effort could explain the difference in rate of success observed between the two groups. However, the main aim of the present study was to attempt to

encourage the use of a more optimal strategy in performing the task.

Conclusion

Although participants did not make better decisions, they may have been working harder to be successful in the task.

References

Camerer, C. F., & Hogarth, R. M. (1999). The effects of financial incentives in experiments: A review and capital-labor-production framework. *Journal of Risk and Uncertainty*, *19*(1), 7–42. Retrieved from `https://doi.org/10.1023/A:1007850605129` doi: 10.1023/A:1007850605129

Carpenter, B., Gelman, A., Hoffman, M., Lee, D., Goodrich, B., Betancourt, M., ... Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software, Articles*, *76*(1), 1–32. Retrieved from `https://www.jstatsoft.org/v076/i01` doi: 10.18637/jss.v076.i01

Clarke, A. D. F., Green, P., Chantler, M. J., & Hunt, A. R. (2016). Human search for a target on a textured background is consistent with a stochastic model. *Journal of Vision*, *16*(7), 4-4. Retrieved from `https://doi.org/10.1167/16.7.4` doi: 10.1167/16.7.4

Clarke, A. D. F., & Hunt, A. R. (2016). Failure of intuition when choosing whether to invest in a single goal or split resources between two goals. *Psychological Science*, *27*(1), 64-74.

Cornelissen, F. W., Peters, E. M., & Palmer, J. (2002). The eyelink toolbox: eye tracking with matlab and the psychophysics toolbox. *Behavior Research Methods, Instruments, & Computers*, *34*(4), 613–617.

Engelmann, J. B., & Pessoa, L. (2007). " motivation sharpens exogenous spatial attention": Correction to engelmann and pessoa (2007). *Emotion*, *7*(4), 875.

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology*, *62*(1), 451-482. doi: 10.1146/annurev-psych-120709-145346

Goodnow, J. J. (1955). Determinants of choice-distribution in two-choice situations. *The American Journal of Psychology*, *68*(1), 106–116.

Hudson, T. E., Maloney, L. T., & Landy, M. S. (2007). Movement planning with probabilistic target information. *Journal of Neurophysiology*, *98*(5), 3034-3046.

Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *American Psychologist*, *39*(4), 341-350.

Manohar, S. G., Finzi, R. D., Drew, D., & Husain, M. (2017). Distinct motivational effects of contingent and noncontingent rewards. *Psychological science*, *28*(7), 1016–1026.

Miranda, A. T., & Palmer, E. M. (2014). Intrinsic motivation and attentional capture from gamelike features in a visual search task. *Behavior Research Methods*, *46*(1), 159–172.

Morvan, C., & Maloney, L. T. (2012). Human visual search does not maximize the post-saccadic probability of identifying targets. *PLoS Comput Biol*, *8*(2), e1002342.

Najemnik, J., & Geisler, W. S. (2005). Optimal eye movement strategies in visual search. *Nature*, *434*(7031), 387–391.

Pelli, D. G. (1997). The videotoolbox software for visual psychophysics: Transforming numbers into movies. *Spatial vision*, *10*(4), 437–442.

Phillips, L. D., & Edwards, W. (1966). Conservatism in a simple probability inference task. *Journal of experimental psychology*, *72*(3), 346.

R Core Team. (2016). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from `https://www.R-project.org/`

Stan Development Team. (2018). *RStan: the R interface to Stan.* Retrieved from `http://mc-stan.org/` (R package version 2.18.2)