

# Inefficient eye movements: Gamification improves task execution, but not fixation strategy

Warren R. G. James <sup>1</sup>, Josephine Reuther <sup>1</sup>, Ellen Angus <sup>1</sup>, Alasdair D. F. Clarke <sup>2</sup> , and Amelia R. Hunt <sup>1,\*</sup> 

<sup>1</sup> School of Psychology, University of Aberdeen

<sup>2</sup> Department of Psychology, University of Essex

\* Correspondence: a.hunt@abdn.ac.uk; Tel.: +44 (0)1224 273139

Version September 8, 2019 submitted to Vision

**Abstract:** Decisions about where to fixate are highly variable and often inefficient. In the current study, we investigated whether such decisions would improve with increased motivation. Presented with two boxes in varying distances, participants had to choose where to fixate to detect a discrimination target, which would then appear in one of the boxes once they fixated a box. To maximize their chances of being able to see the target, participants would simply need to fixate between the two boxes when they were close together, and on one of the two boxes when they were far apart. We “gamified” this task, giving participants easy-to-track rewards that were contingent on discrimination accuracy. Their decisions and performance were compared to previous results that were gathered in the absence of this additional motivation. We used a Bayesian Beta Regression model to estimate the size of the effect and associated variance. The results demonstrate that discrimination accuracy does indeed improve in the presence of performance-related rewards. However, there was no difference in eye movement strategy between the two groups, suggesting this improvement in accuracy was not due to the participants making more optimal eye movement decisions. Instead, the motivation encouraged participants to expend more effort on other aspects of the task, such as paying more attention to the boxes and making fewer response errors.

**Keywords:** visual search; eye movements; attention; strategy; decision

## 1. Introduction

We use eye movements to solve many problems, such as avoiding obstacles, or to find an item of importance. The ability to direct our vision is invaluable. In some situations, humans appear able to use eye movements to direct the visual system in an optimal manner [1], while in other situations, people do not make eye movements that are consistent with an optimal strategy [2,3]. Najemnik and Geisler [1] derived an “ideal observer” model of eye movements in visual search, which minimises the number of fixations needed to find a target by deploying fixations that reduce overall uncertainty of information in the search space. To do this, the model accounts for fixation history, as well as differences in acuity across the retina, to select fixations that provide the most new information. They found that humans matched the ideal model in terms of the number of fixations needed to find the target, suggesting that we make decisions about where to fixate based on maximizing information gain. However, calculating expected information gain for all possible fixations is a resource-intensive process, considering we tend to select new fixations 2-3 times each second during search. With this in mind, Clarke *et al.* [4] suggest a simpler alternative model, that human performance in a visual search task can be characterised as a stochastic process. The stochastic model randomly selects each eye movement during search from a population of saccade vectors executed from that region of the search array, until the target is “found” (that is, until the model’s fixation lands close enough to the target

that it would be detected, given empirical estimates of target visibility over eccentricity). Though the population from which the saccades were selected may have been shaped by factors such as the target visibility and search history, the model itself did not use this information in selecting each new fixation location. Like [Najemnik and Geisler \[1\]](#), [Clarke and Hunt \[2\]](#) also find that their model matches human observers in terms of the number of fixations needed to find the target, despite the selection process being random instead of driven by expected information gain.

This leaves us with a puzzle: two models with very different underlying mechanisms that both describe human search equivalently well. On the one hand, humans may make each fixation based upon current scene statistics and fixation history, and on the other, their fixation selection may be based on a more generic set of perceptual and motor biases. The assumption that differentiates these two models is whether fixations are selected based on a principle of expected information gain.

Experiments by [Nowakowska \*et al.\* \[5\]](#) directly test this assumption and suggest both models may be correct. They presented participants with search arrays that had two halves: on one half, the target would pop out from the distractors and could be easily seen using peripheral vision, and on the other side, the distractors were more variable and the target would be harder to spot. An optimal searcher would direct all fixations to the more difficult side, because no new information would be gained by fixating on the easy side. The results showed large individual differences. A few participants conformed to an optimal strategy, but a wide range of other strategies were observed. These strategy differences were closely correlated with the efficiency of search performance, and [Clarke \*et al.\* \[6\]](#) showed they were also stable over time (see also [\[7\]](#)). These studies suggest that no single model can describe fixation selection during search: variability driven by contexts and individuals will need to be accounted for.

Another paradigm that addresses the extent to which eye movements are driven by expected information gain was devised by [Morvan and Maloney \[3\]](#). In their task, participants had to make a single eye movement in order to discriminate a small target dot that could appear in one of two locations. In the beginning of each trial, participants were presented with three boxes. One in the centre, with two boxes either side. Participants were instructed to fixate one of these boxes and then the target would appear in either of the two side boxes. Throughout the experiment, these two side boxes would appear at different eccentricities, sometimes being placed very close to the central box, and at other times a large distance away. When the boxes were closely spaced, the optimal strategy would be to fixate the central box, using peripheral vision to detect the target when it appears to the left or right of fixation. However, as the distance between the boxes gets larger, performance from a central fixation position begins to decrease towards chance (50%). Therefore, beyond a certain box separation, the optimal strategy is to fixate either of the two side boxes, as this would achieve an expected accuracy of around 75% (100% for targets appearing in the fixated box, and 50% for targets appearing in the other box). If participants were to select fixations according to expected information gain, given the limitations of their own visual acuity, they would fixate the centre box when expected accuracy at the given box separation is greater than 75%, and one of the side boxes when expected accuracy from the centre is less than 75%. The results of Morvan and Maloney's experiment clearly demonstrate that people failed to select fixations according to this optimal strategy. This failure went well beyond not quite switching at the optimal point, or not consistently carrying out the optimal strategy. In fact, none of the participants in their study modified their choice of which box to fixate according to the distance between the boxes.

This failure to fixate in a way that optimizes information gain during search does not appear to be a limitation that is specific to eye movement and visual search. [Clarke and Hunt \[2\]](#) demonstrated the same striking decision failure across a range of different tasks (choosing where to stand to throw objects at targets, choosing whether to try and memorize one or two digit strings for later report). They characterised the decision failure as an inability to use an expectation of task difficulty to guide decisions about whether to attempt to accomplish multiple goals or to focus resources on a single task.

Sub-optimal behaviour is prevalent in many other tasks [8]. Potential explanations for the observed patterns have been discussed by [Kahneman and Tversky](#) [9], as well as [Gigerenzer and Gaissmaier](#) [10]. One important consideration in the decision dilemma posed in the current study is that participants might not have an accurate representation of their own ability, and hence, are unable to adopt the optimal strategy, which requires them to use that information to decide whether to pursue one goal or two. This would be an example of Bounded Rationality [11] where human behaviour is argued to be optimal given some constraints. The constraints in this case would come from imprecise information about one's own skill, or a lack of confidence in the accuracy of that information. This explanation seems unlikely, because participants in these experiments first complete a session in which their ability with one target is recorded. The purpose of this session is to calculate their switch point in the second session, and to calibrate the second session's difficulty level to match the skill of the participant, but a secondary benefit of this session is that the participant gains experience doing the task and observing their own chances of success at different levels of task difficulty. To check if participants successfully learned this information, [James et al.](#) [12] asked participants to provide estimates of their chances of successfully hitting a target at a range of different distances away. They found that participants' estimates were very close to their actual accuracy in performing the task. This suggests that participants do have the information necessary to perform optimally, but do not make use of this when making decisions.

Other explanations for suboptimal decisions in this task could be based on other classic cognitive biases, such as risk or loss aversion [9] or achievement motivation [13]. However, there is a great deal of variation in the decision behaviour of individual participants, with some consistently trying to accomplish both goals, and others consistently focusing on one, and most showing a blend of both strategies (but rarely in a way that varies systematically with task difficulty). As such, one single bias or heuristic might be able to explain the behaviour of some individuals on some trials, but does not provide a satisfying explanation for sub-optimal performance in the experiment overall.

The potential explanation for sub-optimal performance that is addressed in the current experiment is a lack of motivation. The idiosyncratic and variable decisions the participants tend to make in this paradigm may be due to participants prioritizing something other than accuracy in the task when making their decisions. Some other studies that added a financial incentive have shown that participants begin to make use of more optimal strategies [14,15]. A clear motivation for success in the form of a financial gain can lead participants to move from employing strategies that attempted to be successful for all outcomes to focus more on those that were more likely. It is important to note that the participants in the experiments of [Morvan and Maloney](#) [3] (N=4 in Experiment 1, N=2 in Experiment 2) were financially rewarded for correct responses, up to a total of \$20. However, the monetary gain was spread out over thousands of trials, which may have diluted its effects. In [2] there were more participants and the experiments were much shorter, but participants were offered no monetary reward for being successful, relying instead on the value placed on success by each participant (which may in some cases have been non-existent). Therefore, it is possible that at least some of the participants were simply not motivated to succeed.

Even with an added financial incentive, the goals of a participant may still be misaligned with the desires of the experimenter. For example, even though in [Morvan and Maloney](#) [3] 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 and Palmer](#) [16] suggest that game-like experiments can encourage participants to prioritise success rather than speed of completion. They posit that enjoyable experiments intrinsically motivate participants to focus on 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 not try and rush through it or be distracted.

In the present study, we gamified the decision problem with the aim to motivate participants to adopt a more success-oriented attitude. We hypothesise that participants taking part in a more engaging version of the task would be more likely to use a more optimal strategy, resulting in a greater rate of success.

## 2. Materials and Methods

### 2.1. Participants

Participants (N=18, 13 female) were recruited via word of mouth at the University of Aberdeen. The age range was 20-23 (mean of 20). The control group data (N=12) was drawn from another study using the same paradigm but without a game context (Experiment 3 in [17]).

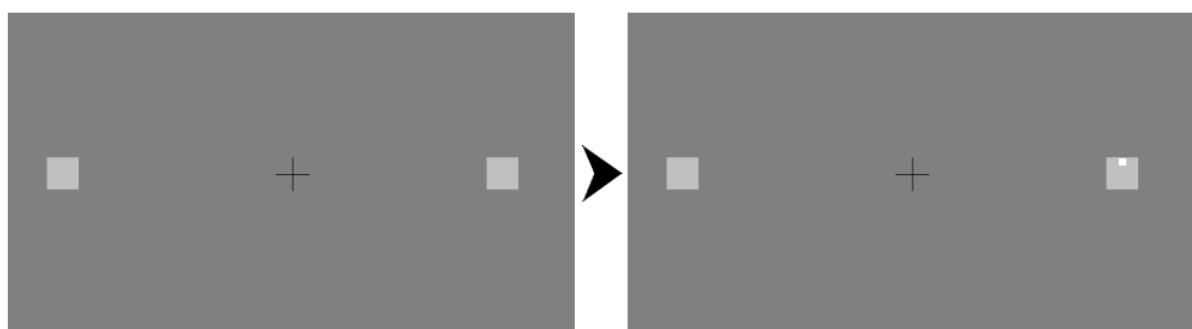
### 2.2. Procedure

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

Both sessions made use of an Eyelink 1000 (version 4.594, SR Research Ltd, Mississauga, Ontario, Canada) which recorded eye position at 1000 Hz. Participants were seated  $\approx 53$ cm from the screen, which was maintained throughout the experiment by the use of an adjustable chin rest. A Sony Trimaster EL OLED was used to display the stimuli. The resolution was 1920 x 1080p (54cm wide, 31 cm tall) with a refresh rate of 60Hz. The experiment was programmed for and run in Matlab 7.9.0 (R2009b) with Psychtoolbox [18] and EyelinkToolbox functions [19]. 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 recalibrated if they had failed to fixate cumulatively 10 times since the last calibration, or if there were 5 fixation errors in a row.

### 2.3. Visual Acuity

In the first session, participants were presented with a fixation cross with two boxes placed either side (Figure 1). They were told that their task was to identify whether a small dot ( $\approx 0.3^\circ$ ) had appeared in the top or the bottom half of one of the boxes (each  $\approx 1^\circ$ ), irrespective of the box it had appeared in 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. If participants broke fixation during any stage of this process, the screen would turn red for 700ms to indicate that they had broken fixation.

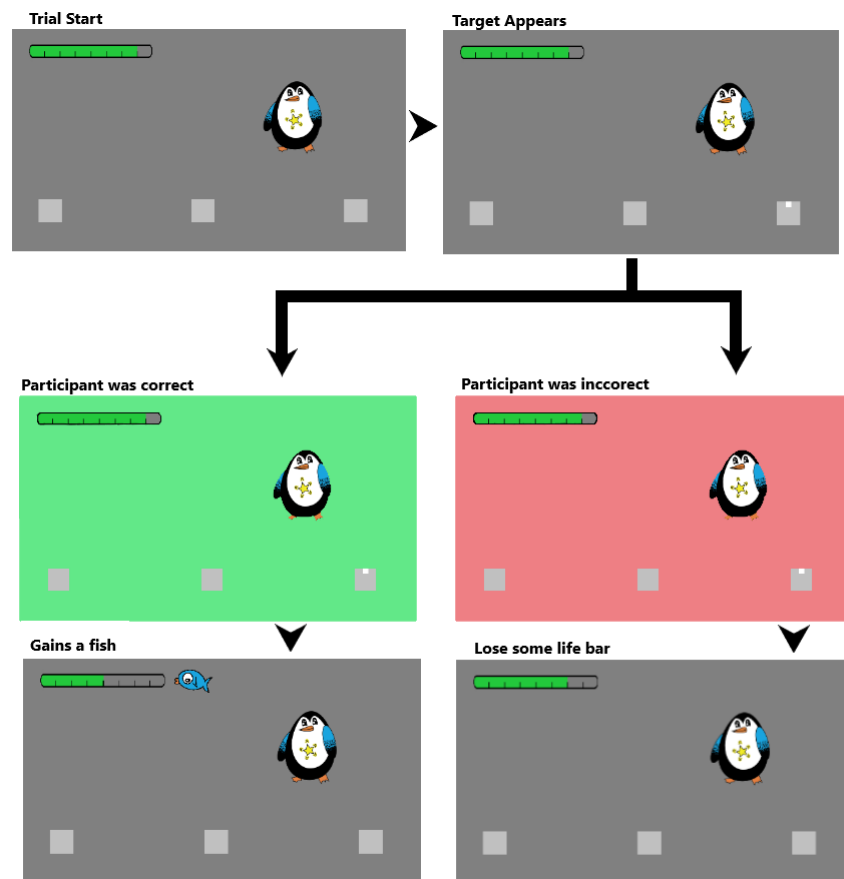


**Figure 1.** The setup for the visual acuity task

The boxes were presented at varying separations from the central fixation cross ( $2.7^\circ$ ,  $3.9^\circ$ ,  $5.2^\circ$ ,  $6.8^\circ$ ,  $8.4^\circ$ ,  $10.1^\circ$ ,  $11.4^\circ$ ,  $12.5^\circ$ ). There were 4 blocks of 96 trials each. Each different separation was presented 12 times in a pseudo-random order where all 12 trials per distance were presented in succession. Data from this session was then used to calculate the separation between boxes at which participants were 75% accurate in detecting the target. This value was subsequently used in the second part to tailor the experiment for each individual.

#### 2.4. Decision Task

For the second part, participants were introduced to Pugadoo the Penguin. Pugadoo was approximately  $5.7^\circ$  tall and  $4.9^\circ$  wide, and had a star on his belly, which served as the fixation mark. Participants were told that their task was to collect as many fish as possible for Pugadoo. To do so, they needed to choose one of three boxes, fixate it and correctly report whether the target dot was presented in the upper or lower half of one of the two side boxes (Figure 2). 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 (selected at random). On each trial, Pugadoo was placed in the top half of the screen so that the star on his belly would end up either midway between the centre and the left, or the centre and the right box, respectively (with an equal chance). This placement matched the placement of the fixation cross used in the comparison set of data [17]. The three boxes were placed along the horizontal meridian, one box in the centre of the screen, and two boxes equidistant on either side of the central box. Boxes were presented with one of nine distances, of which seven were based on each participant's individual switch point that was calculated based on each participant's performance in the first part. These distances were defined as the switch point itself, and then the switch point  $\pm 3^\circ$ ,  $\pm 2^\circ$ , and  $\pm 1^\circ$ . A further two distances were constant across all participants ( $8^\circ$  and  $18^\circ$ ). In the left corner of the screen, participants were presented with a bar indicating the number of correct responses needed to earn the next fish. The bar was presented in the top left corner of the screen throughout the duration of the task, and was approximately  $12^\circ$  by  $0.75^\circ$  in size. The bar had 10 "ticks" and started half way filled (5 ticks) at the beginning of each block. One "tick" was added, or deducted, for each correct or incorrect response, respectively. If the bar was filled to 10 ticks, one fish would be added to the right hand side of the life bar. If the bar decreased to have 0 ticks, then the number of fish would decrease by one. After a change in fish count, the bar would go back to being half filled. Fish were presented next to the bar and remained visible for the duration of the block.



**Figure 2.** The setup for the decision task.

At the beginning of each trial, participants would fixate the star on Puggadoo's belly (Figure 2) and press the spacebar. After a stable fixation of 700 ms, three boxes were displayed on screen. Participants then executed a saccade to the box of their choice (left, centre or right). After 50ms of a stable fixation in the selected box, a target dot would appear in one of the side boxes for 500ms. Participants indicated its position within the box (top or bottom) by pressing the "up" or "down" arrow keys. If the response was correct, the background turned green and the bar would fill by one tick. If the response was incorrect, the background would turn red and the bar would be emptied by one tick. Filling the bar completely was accompanied by a sound cue of Homer Simpson celebrating in his traditional manner of "Woohoo!" to signify that they had gained a fish. When fully depleting the bar, a fish would be taken away from the total at which point a separate sound cue was presented (Homer Simpson saying "Doh!"), signifying the loss of a fish. The constant presentation of the bar and the fish, together with the trial-by-trial feedback allowed participants to monitor their performance while completing the task.

### 2.5. Comparison Data

Additional data were retrieved from [https:// osf.io/t6c5q/](https://osf.io/t6c5q/) in order to compare the current study's data to that of non-motivated participants (*Control* group) and participants that were guided as to where they should fixate (*Optimal* group). The setup for the comparison data was the same as in the current study. The data consisted of 24 participants who were equally split between a control group and an instructed group. The *Control* group carried out the task as per [2] while the *Optimal* group were instructed as to where they should fixate on each trial so as to enforce the use of the optimal strategy. These instructions came in the form of one of the boxes blinking which signalled that participants were to fixate that box for this trial. Which box to fixate was decided upon by using the participant's session



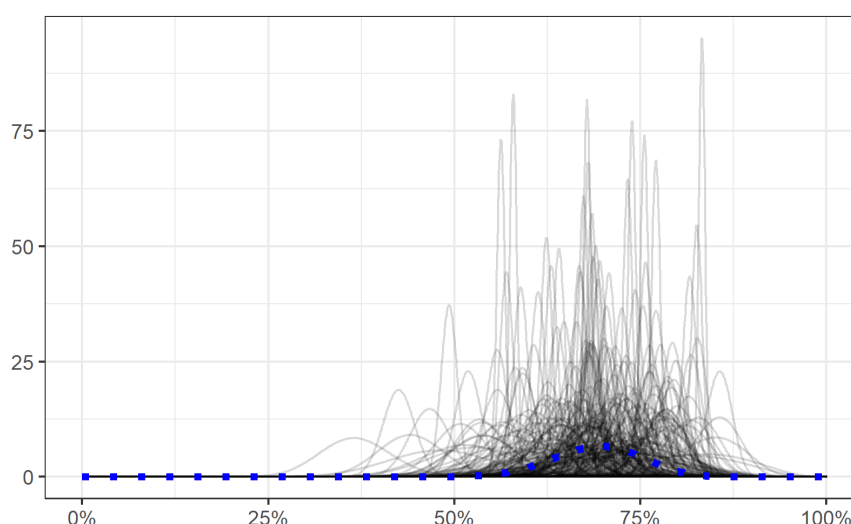
1 performance and choosing the central box if their expected accuracy for that distance was above 75%, or randomly selecting between one of the two side boxes if their expected accuracy was below 75%.

These acted as comparison groups for the current study; if motivation does encourage people to make more optimal decisions, then the participants in the current study should differ from the *Control* group and be more similar to the *Optimal* group. In the original study, participants completed eight blocks of decisions. However, the optimal group were only guided for the first four blocks. As such, only the first four blocks for each participant were used in the analysis section here. Additionally, the current study only used four blocks and as such taking the first four blocks ensures that any learning effects are equalised due to the same level of experience with the task.

## 2.6. Analysis

All analysis was carried out in R version 3.4.3 [20] using rStan [21]. To model these data, we used a Bayesian Beta Regression model [22] with logit link in order to compare the different rates of success between the three groups. The reason for using a Bayesian approach is that we are more interested in looking at the effect size, and associated variance and uncertainty than examining whether the difference is significantly different from zero. We opted to use Beta regression rather than logistic regression, as this allows us to directly compare the observed rates of success with each participant's expected rate of success (which is explained in a later section).

A prior was calculated for the *Control* group based on the performance of participants in Session 1. To estimate these, a *Beta* distribution was fit to average accuracy data for participants in part 1. From the fit shape parameters, measurements for the mean and precision were calculated and the appropriate transformations applied in order to set the priors for the parameters in the model. The prior for the coefficient for the mean was normally distributed about 0.8 (SD = 0.4) and the coefficient for the precision was normally distributed about 5.6 (SD = 1.4). Several samples from these priors can be seen in Figure 3. Priors for the effects of being in the *Motivated* or *Optimal* group were both centred on 0 (SD = 1) so as to avoid assuming that group membership had any effect.



**Figure 3.** This plot shows 250 samples from the prior in black. The dotted blue line shows the prior using the values obtained from session 1 data.

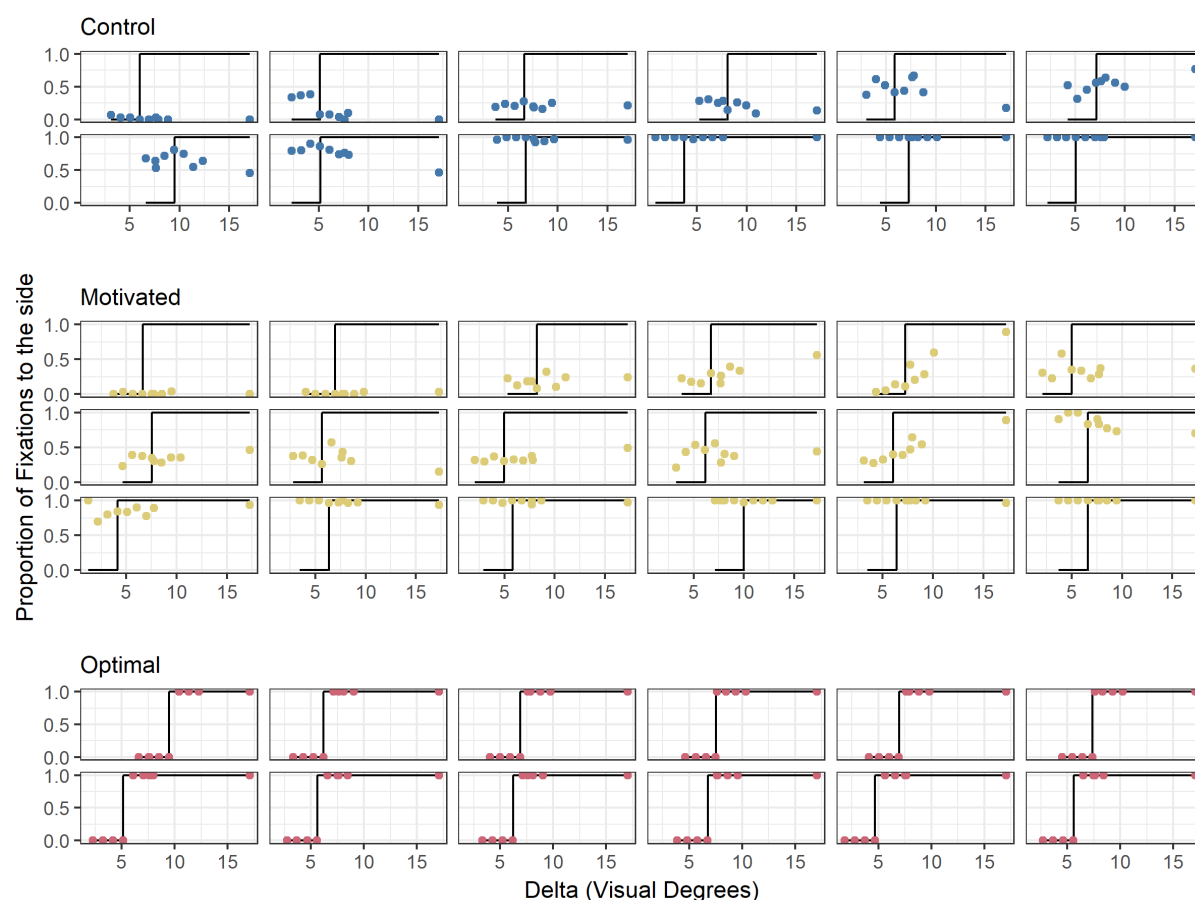
### 3. Results

#### 3.1. Fixation Proportions

On each trial, participants were given three possible locations to fixate, and the target did not appear until a fixation was detected in one of these locations. 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 as a function of box distance.

Figure 4 shows the proportion of fixations to one of the side boxes for each Delta (box separation in Visual Degrees) that was tested for each participant. This is split to show the different conditions under which the participants completed the task, with one panel for the data of the control group, one panel for the data of the motivated participants, and one panel for the participants that were guided to make optimal decisions. From an initial inspection of Figure 4, participants in the motivated group do not appear to differ from the Control group. Both the control and the motivated group differ dramatically from the optimal group, who were instructed where to fixate on each trial. In Figure 4, the plots have been arranged by the proportion of fixations each participant made to the central box. This helps to demonstrate that there is a wide range of behaviours that can be, and are, observed in this data. There does not appear to be any systematic difference between the *Motivated* and *Control* groups. Instead, the *Motivated* group are as variable as their counterparts in the *Control* group.



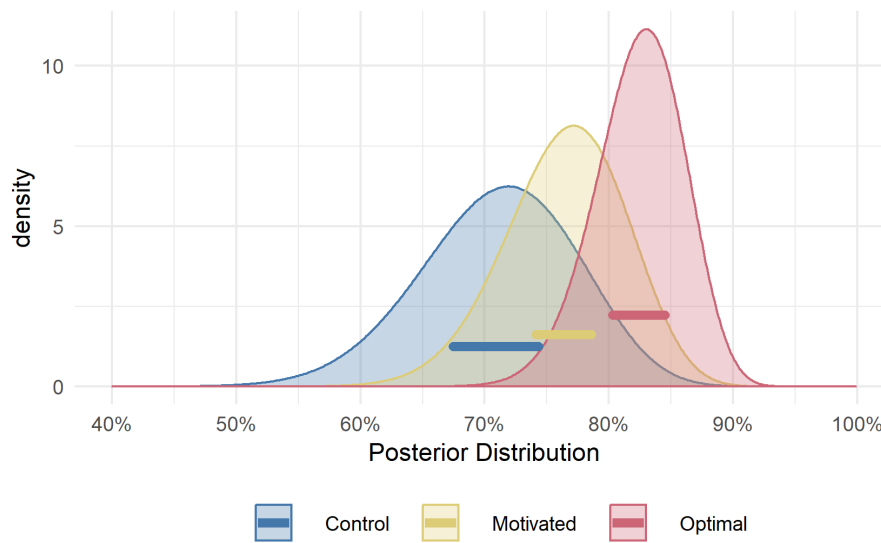


**Figure 4.** Plot of proportion of fixations to the side box (y-axis) with increasing separation between the boxes (Delta, x-axis). Each separate plot shows the behaviour for an individual participant. The bottom set of plots (in red) demonstrate the behaviour of participants who were guided to employ an optimal fixation strategy. The middle (in yellow) plots show the *Motivated* group, and the top (in blue) plots show the *Control* group. The black line shows where participants would have fixated, had they made use of the optimal strategy.

### 3.2. Rate of Success

Initial analysis for rate of success examined the overall accuracy for each participant. To do this, an average was calculated for each participant which was then modelled as described above. For each model, the average value from the posterior will be reported as well as the 95% Highest Posterior Density Interval (HPDI) for this value. The HPDI is the narrowest region of the posterior distribution that contains, in this case, 95% of the estimates for the mean.

Rate of success was modelled using group as a predictor variable. As the primary interest of this paper is the overall rate of success, models including Delta (Visual Angle) are not included in the main text, but are available in the supplementary material.



**Figure 5.** Posterior distributions of success rate in target detection. The lines highlight the 95% HPDI for the mean of each group's predicted rate of success.

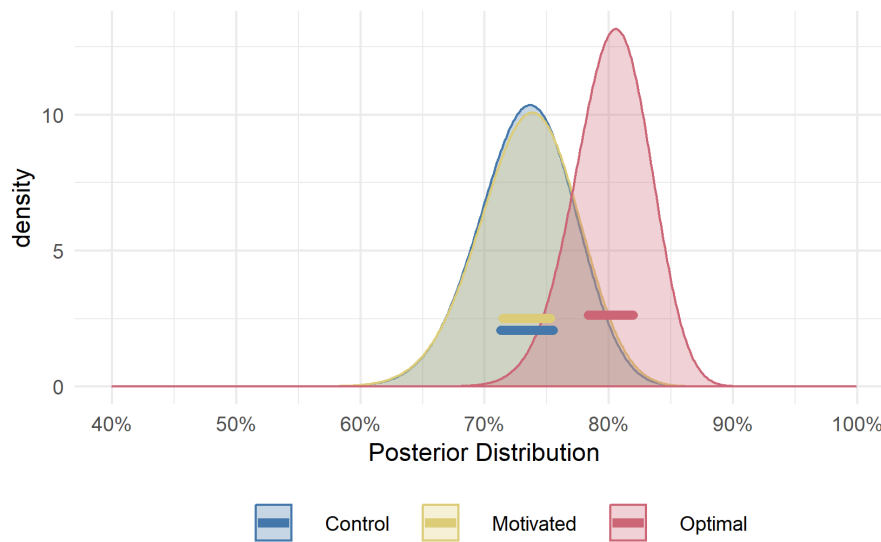
Figure 5 shows the posterior distributions for success rate in target detection. The *Motivated* group were on average 5.4% (95% HPDI of |1.4%, 9.4%|) more successful than the *Control* group. Neither group managed to reach the same rate of success as the *Optimal* group who were on average 11.3% (95% HPDI of |7.3%, 15.5%|) more successful than the *Control* group, and 5.9% (95% HPDI of |2.7%, 9%|) more accurate than the *Motivated* group. This suggests that our manipulation worked: while the motivated group still behave sub-optimally, they do manage to outperform the control group. The results also demonstrate that the optimal fixation decision strategy does substantially improve success (i.e. detection accuracy), and that the motivated group, while better than the control group, does not achieve the detection success rate of the optimal group.

### 3.3. Expected Rate of Success Given Fixation Strategies

The higher rate of success in the discrimination task for the motivated group relative to the controls may initially be surprising, given the lack of any clear differences in the fixation strategies between the motivated and control groups that can be observed in Figure 4. Given that the distances in the decision task were calibrated to each individual's performance in the first session, the same overall accuracy can be expected across groups if their fixation strategies are similar. The likely explanation is that the introduction of Pugadoo in the second part of the experiment caused the thereby motivated group to try harder to see the dot, and to make more careful responses in the discrimination task, relative to the control group, who performed the task without any additional incentive to respond correctly.

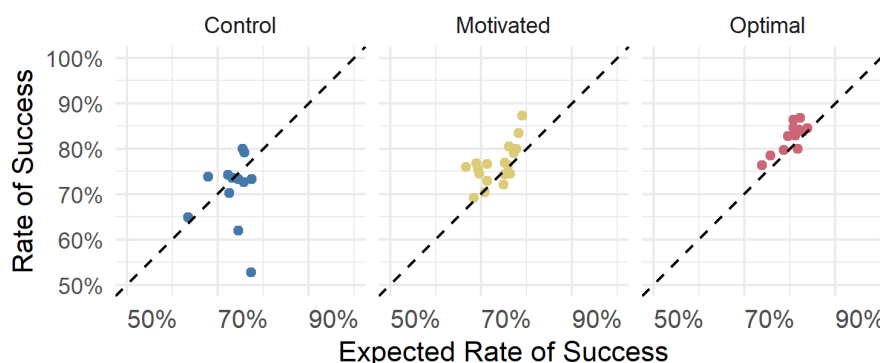
To investigate what proportion of the differences between groups could be explained by a more optimal fixation strategy, rather than general improvements in the discrimination task, each participant's expected success rate was calculated given the fixations choices they had made on each trial. This calculation of expected success was based on both group's Part 1 performance, before the penguin character was introduced to the motivated group. Using the participants' performance in Part 1, it was possible to estimate how likely each participant was to detect the target at all possible ranges that were present in Part 2. To get a measure of expected accuracy, we calculated the participants accuracy for the box on the left ( $B_l$ ) and the box on the right ( $B_r$ ) given the distance from fixation. As each box was likely to contain the target, simply taking the average of these two values provides the expected accuracy for that trial. For example, if the boxes were far apart and the participant had fixated the left hand box, they would have a  $\approx 100\%$  chance of success for the left hand box ( $B_l$ ), and  $\approx 50\%$

chance of success for the right hand box ( $B_r$ ) as this would be imperceptible but they would have a 50% 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  $(B_l + B_r) \div 2$ . This estimate of accuracy given actual fixation choice was calculated for each participant, on each trial, and then averaged over all trials give an average expected accuracy per participant. This “estimated success rate” removes the effect of improvements in discrimination performance from part 1 to part 2, and also removes the effects of any runs of good or bad luck for any specific participants (e.g. choosing the target or non-target side on more than half the trials, or seeing the target but pressing the wrong key). As such, this measure provides an estimate of the effect on success rate of fixation decisions alone; any improvements to performance in the motivated group compared to controls in this measure will reflect a more effective fixation strategy.



**Figure 6.** Posterior distributions of estimated success rate in target detection given fixation decisions. This estimate of success is based on each participant’s visual acuity as measured in session 1, so isolates the effect of fixation decisions on detection performance. The line highlights the 95% HPDI for the mean of each group’s expected rate of success.

It is clear from Figure 6, which shows the posterior distributions for expected success rate, that the control and motivated groups are now almost perfectly overlapping. In other words, the difference that was observed when modelling participant’s raw success rate (Figure 5) entirely disappeared: the difference between the average expected rate of success between the Motivated and Control groups was  $\approx 0\%$  (95% HPDI of  $[-0.25\%, 0.3\%]$ ) suggesting that the Motivated and Control group were very similar, if not identical, in terms of their expected success rate. These results suggest that the previously observed greater rate of success experienced by the Motivated group was not due to them making better, more strategic decisions about where to fixate. Instead, this greater rate of success is due to other factors, like trying harder to see the target and trying not to make any button-pressing errors (or perhaps they were simply more lucky than their non-motivated counterparts). As can be seen in Figure 7, the Motivated and Optimal groups appear to perform the task in such a way as to reach a similar level of success as would be expected given their choices, however, the Control group appears to have some participants whose performance falls short of what would be expected.



**Figure 7.** These scatter plots show a point for each participant's Rate of success and how it relates to their Expected Rate of success. The dashed line is a line with slope 1 going through the origin which would represent a 1:1 relationship (i.e., they performed as well as would be expected).

#### 4. Discussion

The results from this experiment are consistent with previous findings in that participants are highly variable and suboptimal in their fixation decisions in this task [2,3]. Despite the presence of a strong motivation to be optimal, that demonstrably increased participants' efforts to do well in the task, none of the participants adopted the optimal strategy. The novel contribution from the current experiment is that the decisions about where to fixate (Figure 4) show no discernible difference in strategy between the Motivated and Control groups. This would suggest that having a clear motivation for success does not cause people to use more optimal strategies. This result rules out a failure to adequately motivate participants as a likely explanation for suboptimal decisions.

The motivated group did have higher overall accuracy in the task than the control group, demonstrating that the motivation manipulation was effective. The boost in accuracy in the absence of an improved strategy suggests participants were expending more effort to detect the target and respond correctly. This result is consistent with several studies demonstrating that visual sensitivity, attention, and effort are all affected by the presence of rewards [16,23,24]. Participants in the Motivated group may have made less erroneous button presses (that is, pressed the key they intended to) than the Control group as a result of more sustained attention/lower lapse rates [25]. They also may have made more of an effort to detect the target during the decision phase, in which the "game" was introduced, and therefore were better able to detect the target than would be expected given their initially measured visual acuity [16].

This finding rules out a general lack of motivation as an explanation for why participants fail to discern or follow an easily implemented decision rule that would increase their accuracy. This can be added to a growing list of other simple explanations. As noted in the introduction, James *et al.* [12] ruled out a lack of self-awareness of ability as a viable explanation. Another possible explanation is that the optimal strategy may in fact be more difficult to implement than it initially seems. Hunt *et al.* [17] ruled this out as well by asking non-naïve participants to perform both the Throwing and Detection task from Clarke and Hunt [2]. The results from these participants were very close to optimal. This demonstrates that it is possible to implement the strategy when it is known.

As noted in the introduction, calculating optimal fixations in the context of visual search can be quite effortful (as demonstrated by the ideal search model of Najemnik and Geisler [1]), and does not always result in large gains in terms of utility or efficiency relative to making "random" fixations [4]. This may go some way towards explaining why participants do not employ the optimal strategy in

the simpler context of the current experiment. They may not know, *a priori*, that the strategy in this particular case is simple and will benefit performance, but they may know that the gains that can usually be achieved by figuring out an optimal strategy would not outweigh the cost of figuring out the strategy in the first place. This could be considered an example of satisficing, in that participants reached a point at which their strategy met their own moderate standards [11]. The participants in the motivated condition performed better in terms of their rate of success by presumably making more of an effort to respond accurately and detect the target. That is, the presence of a motivating element might have raised their standard above that of the *Control* group. It is possible that, had they explicitly recognized the additional gains they could be making with better fixation strategies, they may have re-examined fixation strategies and made more efficient decisions.

A striking contrast exists between the results from this task, in which participants are highly variable and sub-optimal, and examples from visually-guided reaching and eye movements for tasks in which participants appear able to execute near-optimal movement plans (e.g.[26,27]). In these studies, participants are able to take into account their own variability in movement production to target locations that balance risk and reward to optimize expected gain. The key difference between those experiments and the one presented here (and in [2,3,12,17]) is the directness of the relationship between the strategy and the rewarded outcome. In the current task, participants need to select a place to fixate that optimizes their accuracy in the detection task, but it is performance on the detection task that is ultimately rewarded. This highlights an interesting distinction between an optimal decision about how to prepare for an upcoming task, relative to optimal decisions about how to execute the task when it comes. In the current study, the second component (executing the detection task) responded to the motivation manipulation and showed clear improvements, but the first component (deciding where to look in preparation for detection) was unchanged. A similar dissociation is reported by [28], in which participants performed rapid reaches towards pairs of lights, one of which was only revealed to be the target after the movement had begun. In these experiments, participants were able to weigh up possible movement plans and select those which minimized mechanical effort. However, they completely failed to flexibly adapt their reaching strategies to maximize expected success, similar to the decision failure reported in the current experiment. It seems likely that there is an important role for reinforcement learning in shaping the optimal performance seen in the task execution component of these decisions. An interesting question for future research to address is why this reinforcement does not improve the strategic preparation aspects of these decisions.

Although in the current study participants were clearly motivated by the game context to perform better than the control group, it is possible that even the motivated participants had set themselves a standard for performance that they were able to meet without changing their fixation strategy. Their goal was simply to get as many fish as they could, and all the participants completed the experiment with a net positive number of fish, so this may have been sufficient for them to feel that they did their best. A clearly defined goal, and an ongoing representation of how close they are to attaining that goal, may help participants benchmark their “success” and try harder to improve. In other words, we cannot rule out that a very strong manipulation of motivation would not elicit more efficient fixation decisions from participants. However, we can conclude that motivation at levels that clearly improve overall task effort does not also lead to improved fixation strategies in this task. This demonstrates that the general failure to implement good fixation strategies in this task is not due to failure in trying to do well.

## 5. Conclusions

In conclusion, the results of this experiment demonstrate that having a clear motivation to succeed does not facilitate the use of more optimal strategies when performing a task. The pattern of results obtained from the Motivated group are in line with those obtained in previous research. The improvement in their detection performance despite the failure to adopt better strategies demonstrate

they are trying harder to successfully complete the task. In other words, motivation caused participants to “work harder, not smarter”.

**Author Contributions:** Conceptualization, Josephine Reuther, Ellen Angus and Amelia Hunt; Data curation, Josephine Reuther and Ellen Angus; Formal analysis, Warren James and Alasdair Clarke; Methodology, Josephine Reuther and Amelia Hunt; Project administration, Amelia Hunt; Supervision, Josephine Reuther and Amelia Hunt; Visualization, Warren James; Writing – original draft, Warren James; Writing – review & editing, Warren James, Josephine Reuther, Alasdair Clarke and Amelia Hunt.

**Funding:** This research was funded by the James S. McDonnell Foundation (scholar award to Amelia Hunt)

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

1. Najemnik, J.; Geisler, W.S. Optimal eye movement strategies in visual search. *Nature* **2005**, *434*, 387–391.
2. Clarke, A.D.F.; Hunt, A.R. Failure of intuition when choosing whether to invest in a single goal or split resources between two goals. *Psychological Science* **2016**, *27*, 64–74.
3. Morvan, C.; Maloney, L.T. Human visual search does not maximize the post-saccadic probability of identifying targets. *PLoS Comput Biol* **2012**, *8*, e1002342.
4. Clarke, A.D.F.; Green, P.; Chantler, M.J.; Hunt, A.R. Human search for a target on a textured background is consistent with a stochastic model. *Journal of Vision* **2016**, *16*, 4–4. doi:10.1167/16.7.4.
5. Nowakowska, A.; Clarke, A.D.; Hunt, A.R. Human visual search behaviour is far from ideal. *Proceedings of the Royal Society B: Biological Sciences* **2017**, *284*, 20162767.
6. Clarke, A.D.F.; Irons, J.; James, W.R.G.; Leber, A.B.; Hunt, A.R. Stable individual differences in strategies within, but not between, visual search tasks **2018**.
7. Nowakowska, A.; Clarke, A.D.; Sahraie, A.; Hunt, A.R. Practice-related changes in eye movement strategy in healthy adults with simulated hemianopia. *Neuropsychologia* **2019**, *128*, 232–240.
8. Rahnev, D.; Denison, R.N. Suboptimality in perceptual decision making. *Behavioral and Brain Sciences* **2018**, *41*.
9. Kahneman, D.; Tversky, A. Choices, Values, and Frames. *American Psychologist* **1984**, *39*, 341–350.
10. Gigerenzer, G.; Gaissmaier, W. Heuristic Decision Making. *Annual Review of Psychology* **2011**, *62*, 451–482. doi:10.1146/annurev-psych-120709-145346.
11. Simon, H.A. Invariants of human behavior. *Annual review of psychology* **1990**, *41*, 1–20.
12. James, W.R.G.; Clarke, A.D.F.; Hunt, A.R. Failure to use probability of success when deciding whether to pursue one goal or two. Proceedings of the 39th annual meeting of the Cognitive Science Society. Cogsci, 2017.
13. Atkinson, J.W.; Litwin, G.H. Achievement motive and test anxiety conceived as motive to approach success and motive to avoid failure. *The journal of abnormal and social psychology* **1960**, *60*, 52.
14. Goodnow, J.J. Determinants of Choice-Distribution in Two-Choice Situations. *The American Journal of Psychology* **1955**, *68*, 106–116.
15. Phillips, L.D.; Edwards, W. Conservatism in a simple probability inference task. *Journal of experimental psychology* **1966**, *72*, 346.
16. Miranda, A.T.; Palmer, E.M. Intrinsic motivation and attentional capture from gamelike features in a visual search task. *Behavior Research Methods* **2014**, *46*, 159–172.
17. Hunt, A.R.; James, W.R.G.; Mackay, E.; Reuther, J.; Clarke, A.D.F. Can we learn to make optimal choices? Submitted.
18. Pelli, D.G. The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial vision* **1997**, *10*, 437–442.
19. Cornelissen, F.W.; Peters, E.M.; Palmer, J. The Eyelink Toolbox: eye tracking with MATLAB and the Psychophysics Toolbox. *Behavior Research Methods, Instruments, & Computers* **2002**, *34*, 613–617.
20. R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2016.
21. Stan Development Team. RStan: the R interface to Stan, 2018. R package version 2.18.2.

22. Ferrari, S.; Cribari-Neto, F. Beta regression for modelling rates and proportions. *Journal of applied statistics* **2004**, *31*, 799–815.
23. Engelmann, J.B.; Pessoa, L. "Motivation sharpens exogenous spatial attention": Correction to Engelmann and Pessoa (2007). *Emotion* **2007**, *7*, 875.
24. Manohar, S.G.; Finzi, R.D.; Drew, D.; Husain, M. Distinct motivational effects of contingent and noncontingent rewards. *Psychological science* **2017**, *28*, 1016–1026.
25. Massar, S.A.; Lim, J.; Sasmita, K.; Chee, M.W. Rewards boost sustained attention through higher effort: A value-based decision making approach. *Biological Psychology* **2016**, *120*, 21–27.
26. Trommershäuser, J.; Landy, M.S.; Maloney, L.T. Humans rapidly estimate expected gain in movement planning. *Psychological science* **2006**, *17*, 981–988.
27. Trommershäuser, J.; Gepshtein, S.; Maloney, L.T.; Landy, M.S.; Banks, M.S. Optimal compensation for changes in task-relevant movement variability. *Journal of Neuroscience* **2005**, *25*, 7169–7178.
28. Hesse, C.; Kangur, K.; Hunt, A.R. Decision making in slow and rapid reaching: Sacrificing maximum task success to minimize effort. Submitted.

© 2019 by the authors. Submitted to *Vision* for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).