Failure of intuition when presented with a choice between investing in a single goal or splitting resources between two goals

A. D. F. Clarke* and A. R. Hunt†

School of Psychology, University of Aberdeen, UK

September 18, 2015

Abstract

In a series of related experiments, we ask people to choose whether to split their attention between two equally likely potential tasks, or prioritise one task at the expense of the other. When the tasks are easy, the best strategy is to prepare for both of them. As difficulty increases beyond the point where participants can perform both tasks accurately, they should switch strategy and focus on one task at the expense of the other. Across three different modalities (target detection, throwing, and memory), none of the participants switch strategy at the correct point. Moreover, the majority consistently fail to modify their behaviour with task difficulty at all. This failure may be related to uncertainty about the trial outcome, because in a version of the experiment in which there is no uncertainty, participants were uniformly optimal. Keywords: decision making, optimal behaviour

1 Introduction

A goalie can choose to stand in the middle of the net or to stand to one side or the other. A student can choose to study all the course material, or to focus on learning a subset more deeply. A funding body can choose to divide resources across a large number of projects, or to focus resources on one or two especially promising ones. At its simplest, the choice that will lead to the best outcome in each of these scenarios depends on the likelihood of success given constraints of time and ability: Is it possible to achieve multiple goals given these constraints? If so, it makes sense to try and achieve them all. Otherwise, we are better off focusing our resources on only one task or goal at the expense of the others. Here we report results from three experiments that – despite using very different methods – all converge on the conclusion that humans are surprisingly deficient at achieving optimal outcomes. When presented with a choice between dividing available resources between two goals versus investing all their resources in one goal, the participants were poor at choosing the best strategy even though the factors they need to take into account are relatively stable and limited. A definition of optimal decisions is those that achieve the best possible outcome while minimising energy expenditure and risk. There are many examples of optimal or near-optimal decisions in humans (e.g. Kibbe and Kowler, 2011, Körding and Wolpert, 2004, Najemnik and Geisler, 2005, Oruç, Maloney, and Landy, 2003). Wolpert and Landy (2012), for example, have argued that motor control can be viewed as a decision-making problem of maximising movement outcome dependent on task, motor and sensory uncertainty. However, others have demonstrated human failures to maximise expected gain in more deliberative human decisions Gardner (1959), Kahneman and Tversky (1984), Morvan and Maloney (2012), Vulkan (2000), Zhang, Morvan, Etezad-Heydari, and Maloney (2012).

Our interest in optimal decision-making began with an intriguing contradiction in the visual search literature. One influential model of search (Najemnik and Geisler, 2005) proposes that each eye movement during search is directed to the location that decreases uncertainty about the target location by the maximum amount possible. How-

*a.clarke@abdn.ac.uk
†a.hunt@abdn.ac.uk

ever, Morvan and Maloney (2012) recently provided striking evidence that human observers do not reliably fixate locations that maximise their probability of detecting a target. In their study, observers had to choose where to fixate, and then a low-contrast discrimination target would appear inside one of two boxes. If humans take the probability of detecting the target at a given eccentricity into account when deciding where to fixate, they should fixate a location in between the two boxes when they are relatively close together. As the boxes move further apart, they will reach an eccentricity at which it is no longer possible to discriminate the target in either box above chance. At this point, observers should switch to a strategy of fixating one or the other box, since this will yield accuracy close to 100% if the target happens to appear inside the fixated box, and an at-chance accuracy if it does not. Surprisingly, not only did all four of the observers fail to maximise their target discrimination performance, their tendency to fixate between versus on the target boxes did not vary with the distance between the boxes at all. Morvan and Maloney (2012) propose that saccade target selection is based largely on heuristics, such as a tendency to saccade in particular directions, rather than taking visual sensitivity and uncertainty into account.

The failure to adjust fixation strategy in response to this very simple change in spatial configuration is surprising, and difficult to reconcile with models of fixation behaviour that depend on a mechanism that maximises information gain (Hayhoe and Ballard, 2014, Najemnik and Geisler, 2005). Our first experiment therefore replicated Morvan and Maloney (2012) in a larger sample. We then were interested in establishing whether this failure of optimal behaviour could be considered specific to the context of eye movements and detection of targets, or a larger problem pervading human decisions in general. Saccadic eye movements are rapid, energy-efficient, and frequently not under voluntary control, so the decisional processes involved may not generalize to other modalities. Indeed, there is precedent for rapid motor responses to achieve different outcomes than deliberative decisions (Hunt and Klein, 2002, Wu, Delgado, and Maloney, 2009). The aim of our study is investigate whether participants exhibit more strategic decision-making in tasks involving deliberate, high-stakes decisions that have more tangible outcomes.

We carried out a series of four related experiments. While each experiment varied in terms of task and modality (detection, throwing, memory and reaching), crucially they all involved the same decision-making paradigm: to experimentally create a point at which, to achieve the best possible outcome, it is necessary to switch between dividing available resources across two goals versus investing all resources in one goal. All experiments were conducted in two sessions. In the first session participants performed the task with only one target/goal. The purpose of this session was to both characterise each participant's performance across difficulty, as well as to facilitate the participant's awareness of their own level of skill across difficulty. In the second session, participants repeated the task, but this time there were two potential targets, and the participant had to make a decision about whether to divide possible success evenly between the two targets, or to abandon one target in favour of the other. Each participants' choices can be compared to individualized estimates of what they should have chosen.

2 Method

The motivation and logic of the four experiments were similar, so we report the methods and results for all four together.

2.1 Participants

Forty-eight undergraduates at the University of Aberdeen were recruited to participate, twelve for each experiment. All participants were recruited via word of mouth and were naive to the aims of the experiment. All gave informed consent to participate in the experiments, which were reviewed and approved by the School of Psychology ethics committee.

2.2 Materials and Procedures

Stimuli and setup: Stimuli and layout for each of the four experiments are shown in Figure 1. Specific details for each of the four experiments are described below.

Figure 1: Example stimuli and set-up of the second session of each experiment. **A** The detection experiment: participants start the trial by fixating the cross. After they shift their gaze to one of the three boxes, the target (a small white circle) appears inside the left or right box. Eccentricity of the boxes varied. **B** The throwing experiment: participants were told what color their target hoop will be. They choose a place to stand and are then told which of the two hoops of that color is their target. **C** The memory experiment: Example of a trial with 5-digit numbers. Participants do not know which number they will have to report. **D** The reaching experiment: Participants are told which color beanbag they will need to pick up, and asked to select a chair.

2.2.1 Detection Experiment

Morvan and Maloney (2012) found choice behavior that was idiosyncratic but clearly not optimal in the four observers they tested. We expected to replicate this pattern, but wanted to ensure it held true in a larger sample. Given that each observer is compared against an individualized estimate of their own optimal strategy, however, a very large sample is unnecessary, so we decided on a sample size of 12 for this and all subsequent experiments. An Eyelink 1000 (SR Research, Canada) was used to record eye movements. The aim of the Session 1 of the experiment was to obtain a psychometric function for each participant for target detection, and for the participant to gain practice with the detection task and familiarity with their own level of performance. The stimulus consisted of two grey squares (length $= 1.8°$, measured in degrees of visual angle), equidistant ($\delta \in \{2.9°, 5.7°, 7.1°, 8.3°, 9.4°, 10.5°, 11.7°, 12.8°\}$) from a central fixation cross. After the participant had maintained a stable central fixation for 1 second, the target (a small white circle) was presented at the top or bottom of one of the two squares. The stimulus was displayed for 500ms, after which a blank grey screen was displayed. The trial was immediately cancelled if the participant broke central fixation. Participants responded via keypress (up/down arrows) as to whether the target had been presented at the top or bottom of one of the two squares, with the instruction to just guess if they were not sure. There were four blocks of 96 trials (384 trials in total). Within each block, trials were presented in order of increasing δ. No feedback was given. Individual performance was modelled in `R` using a generalised linear model with the mafc-probit link function from the `psyphy` package. Let $\phi(\delta)$ be this function, where δ is the distance from the fixation point to the target.

In the second session, which took place about a week later, participants fixated a crosshair above three boxes and were instructed to choose one box to fixate. The crosshair was presented above the targets and positioned so that it was equidistant from the central and rightmost grey square (see Figure 1A for an illustration of the trial setup). This meant that the cross's position varied with the separation. The same eight values of δ as above were used, with 48 repetitions of each. This gave a total of 384 trials, which were presented in a random order. After a fixation was detected inside one of the three boxes, the target was presented in either the left or right box and, as in Part 1, the participant had to simply report whether the target had appeared up or down. Participants were told the target would never appear in the center box. We use ϕ from Session 1 to derive each participant's optimal strategy and predicted accuracy. When the separation between the boxes is small, participants can direct their saccade towards the central box and have a good chance of detecting the target in either location. Once δ increases to the point where $\phi(\delta) < 0.75$ the participant should switch strategies and fixate either the left or right box. When fixating the left or right box, there is a 50% chance that the target will appear at fixation, giving them $\approx 100\%$ chance of correctly responding up or down, and a 50% chance that the target will appear at the other box, in which case they will be correct 50% of the time by guessing. Together this gives an expected accuracy of 75%.

2.2.2 Throwing Experiment

This experiment is analogous to the detection experiment, except the task is to get a beanbag into one of two hoops. It is not known which hoop will be designated as the target, and participants are asked to choose a standing position. For two hoops close to one another, the ideal position to stand is halfway between them. If the distance between hoops is too large to throw accurately from the center, however, the optimal behaviour is to stand close to one hoop, giving a success rate of 50%.

The experiment took place in a sheltered area of concrete slabs (see Figure 1B). Each slab was $0.46\times0.61m$, making them useful markers for placing hoops and recording standing positions. In the first session, participants stood in the center slab of the area, which was marked with black tape, and flat hoops with a diameter of 0.40m were placed at six different distances from them (1.88, 3.22, 4.14, 5.06, 6.90, and 8.74m). Each participant tossed 12 beanbags for each hoop distance, in order of increasing distance, with each beanbag cleared from the area after each toss. The participant then completed the same set of distances but tossing in the opposite direction (this was counterbalanced), for a total of 144 trials. A trial was recorded as "correct" if the final resting place of the beanbag was inside or touching the specified hoop. No differences between direction were found so we ignored this factor in subsequent analyses. Each participant's accuracy was modelled using logistic regression with a fixed intercept of (0, 0.99). That is, we assume that participants are 99% accurate if they stand right next to the hoop. Each participant's curve (modelled the same way as the detection experiment) was used to select six slabs on which to place hoops in Session 2. We based these around the slab at which participants were closest to 50% accurate (i.e. where $\phi(\delta) = 0.50$, which we will call slab M. This is the point where, to maximise accuracy, participants should switch between standing in the centre to standing close to one or the other hoop.

For the first block of Session 2, six hoops were taped down, three in each direction: at slabs $M+1$ $M-1$ and on the slab with expected accuracy of 90% (relative to a centre point, which was unmarked). The second block was the same configuration but with hoops taped down on slabs M, $M+2$, and on the slab with expected accuracy of 10%. Red hoops were always closest, yellow in the middle, and blue furthest away. Participants were told:

> "You will be given a beanbag. Your task is to get the beanbag into one of the two hoops of the same color. For example, if you are handed a yellow beanbag, this means you will have to get the beanbag into one of the two yellow hoops. I am not going to tell you which hoop yet. First, you need to select a place to stand. You can choose anywhere you like within the paved area, but remember your task is to get the beanbag in the hoop of the specified color. Once you are in position, tell the experimenter you are ready."

Participants received one practice trial, and then completed 48 decisions/throws in each block (16 trials for each distance condition in a random order). The main experimenter stood on the grass to the side of the paved area and the participant returned to them after each trial to receive a new beanbag, while the other experimenter cleared the beanbags and recorded accuracy and standing position (the numbers 1 to 40 had been chalked on the edge of the paved area from one end to the other, for quick and subtle recording of standing position). The order of colour and direction of throw was randomised separately for each participant.

2.2.3 Memory Experiment

In this experiment participants were shown two numbers, and later asked to report only one of them. At the time of presentation, the participant did not know which number would have to be reported. If the two numbers have a small number of digits, and are therefore easy to remember, the ideal behaviour is to look at, and memorise, both numbers. However, as the number of digits increases to the limits of your digit span, the optimal behaviour is to focus on just one of the numbers and ignore the other one.

In Session 1 we measured each participant's digit span. Stimuli consisted of a randomly-generated sequence of between 2 and 12 digits. On each trial this digit sequence was displayed on the left or the right hand side of the screen. The two halves of the screen had a different coloured background, with the colours swapping randomly on a trial to trial basis. The number was displayed for five seconds, after which it was replaced with a grey screen for three seconds. Finally, a response screen was presented, consisting of the same coloured background as earlier, with the prompt "please enter the number" displayed at the location where the sequence of digits had originally been placed. Participants then typed the number in using the keypad. A response was considered correct if all the digits were typed in the correct order. There were nine repetitions at each number length, giving 99 trials in total. Trials were presented in a random order and participants were given a break halfway through. As with the earlier experiments, we model each participant's accuracy using logistic regression, giving $\phi(n)$ as the probability of remembering an n-digit number.

The second session of the experiment was similar to the first. The main differences were that participants were eye-tracked while carrying out the task (with an Eyelink 1000 as in Experiment 1 above), and they were presented with two sequences of digits to memorize (Figure 1C). The two sequences of digits were equal in length. When the response screen was shown, participants were prompted to report either the left or the right number. The coloured background was used as an additional prompt. There were 165 trials: 15 repetitions for each value of n (2 to 12, as in Session 1). Trials were presented in random order.

Eye-tracking data were analysed by assigning fixations to one of two $14° \times 2.8°$ areas of interest centred on the two numbers. Fixations falling outside of these areas were discarded. Attentional split was then defined as the proportion of time spent fixating the area of interest that received the most attention. So a value of 0.5 indicates the participant spent equal time looking at both numbers, while a value of 1.0 tells us that participants spent all their time fixating one of the numbers. Unlike the previous two experiments, it is not as straightforward to derive the predicted accuracy given an optimal strategy. We can estimate the probability of remembering both numbers as $\phi(2n)$, but the data show that this underestimates performance for small n (see supplementary materials), presumably due to chunking (Miller, 1956). We assume that for small n, our participants were memorising both numbers. However, as n increases the task becomes increasingly difficult, participants should change strategy and only attempt to remember one number, with a probability of them getting the trial correct being $0.5\phi(n)$.

2.2.4 Reaching Experiment

To foreshadow, the results of the above three experiments indicated consistent failures in strategic thinking. In the final experiment we took the basic choice we asked of participants in the previous experiments to a trivially simple level, to ensure our results are not a consequence of participants failing to fully understand the decision we were asking them to make. Six beanbags were placed on a long table (Figure 1D), with two red beanbags near the centre, two green beanbags each placed halfway to the end, and a blue beanbag at each end. Participants were first asked to sit in a chair placed by the middle of the table and asked to try and reach, with their back still touching the seat, the red, green, and blue beanbags (demonstrating their own reach span, as in Session 1 of the previous experiments). Participants were then asked to stand, and the experimenter asked them to choose one of three chairs to sit in to pick up a beanbag of a specified color. As in the throwing experiment, they were not told which of the two beanbags of this color they would have to pick up until after they selected a chair. Participants selected a chair once for each of the three colours. The order of colors and which was to be picked up was randomised for each participant.

3 Results

A typical individual's data from Session 1 of each of the saccade, throwing, and memory experiments are shown in the top row of Figure 2. The full set of data from all participants can be found in the supplementary information. In the bottom row of Figure 2, the same participant's actual behaviour in the second session is compared to the optimal strategy derived from their Session 1 performance (the blue line).

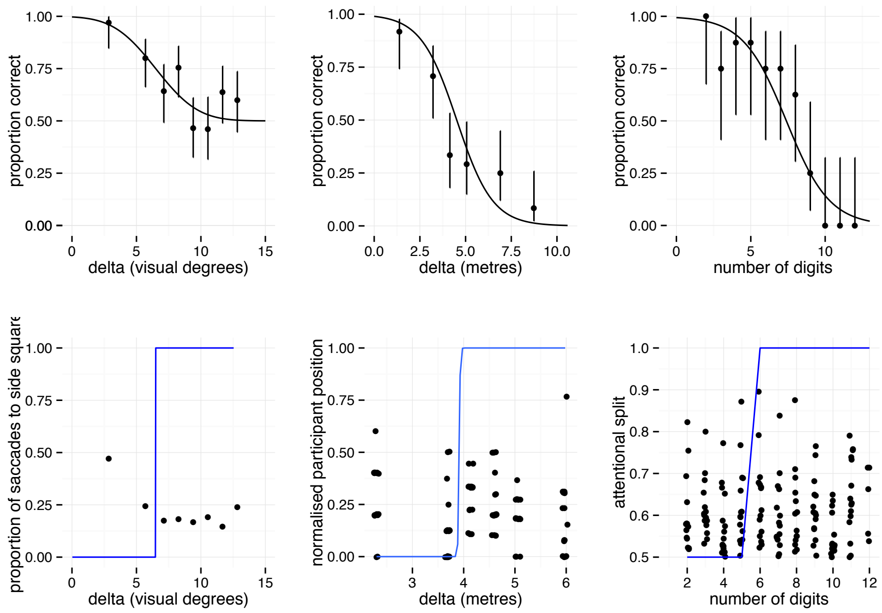

Figure 2: Results from one participant in each of the detection, throwing and memory experiments. Results from other participants were similar and are shown in the supplementary materials. Top row shows Session 1 accuracy. Accuracy decreased as the task difficulty increased, either by increasing distance (δ) or, in the case of the memory task, the number of digits. Error bars show 95% binomial proportion confidence intervals. As the detection task is two-alternative forced choice, chance performance is 50%. The bottom row shows decision behaviour from the same participants in Session 2 of the experiments. For the detection task, the participant's choice is binary and hence each dot represents the proportion of trials on which the participant fixated the side box in each condition. In the other two tasks, each dot represents behaviour on a single trial. For the throwing experiment, position "0" is the center and "1" is the distance to the target color on that trial. The blue line illustrates the optimal strategy that the participant should adopt, based on their performance in Part 1.

Figure 3 shows the decision behaviour of all individuals in all four experiments. In the first three experiments, the overwhelming majority of participants failed to systematically change their behaviour with increasing task difficulty. In the detection experiment, as in Morvan and Maloney (2012), each participant selects their own individual strategy, and they tend not to vary this strategy as difficulty increases. Only one of the 12 participants exhibited behaviour that approached an optimal strategy[1]. For the throwing experiment, there are fewer trials and individual strategies are less consistent; nonetheless the participants stood, in aggregate, just as close to the centre when throwing to hoops that were far away as when the hoops were close together. For this experiment we also examined sequence effects at the trial level to see if participants had a tendency to learn or to persist over time with one strategy over another. There was no consistent pattern here (see Supplementary Information, Figure SupMat.4). For the memory experiment, they were as likely to fixate both digit sequences when they were long as when they were short, although several participants came closer to adopting a strategy that was optimal in this experiment, a detail we will return to in the discussion. In the reaching task, designed to check that participants could correctly understand the instructions used in the preceding studies, participants were uniformly optimal.

Each participant's choice behaviour can be modelled by fitting a step function [$y = c_1$ for all $x <= s$, $y = c_2$ for all $x > s$, where s is the point at which the participant switch strategies (e.g. from a centre to a side strategy). A linear model would not be appropriate, given the nature of the optimal strategy as depicted in Figure 2. We fit s, c_1, and c_2 to the data using least-squares. From this model fit, four patterns of behaviour can be roughly categorized:

[1] See Supplementary material, Figure SupMat.1, Participant 3. This participant was the only member of our lab to participate in any of the studies, although she was naive to the aims of the experiment.

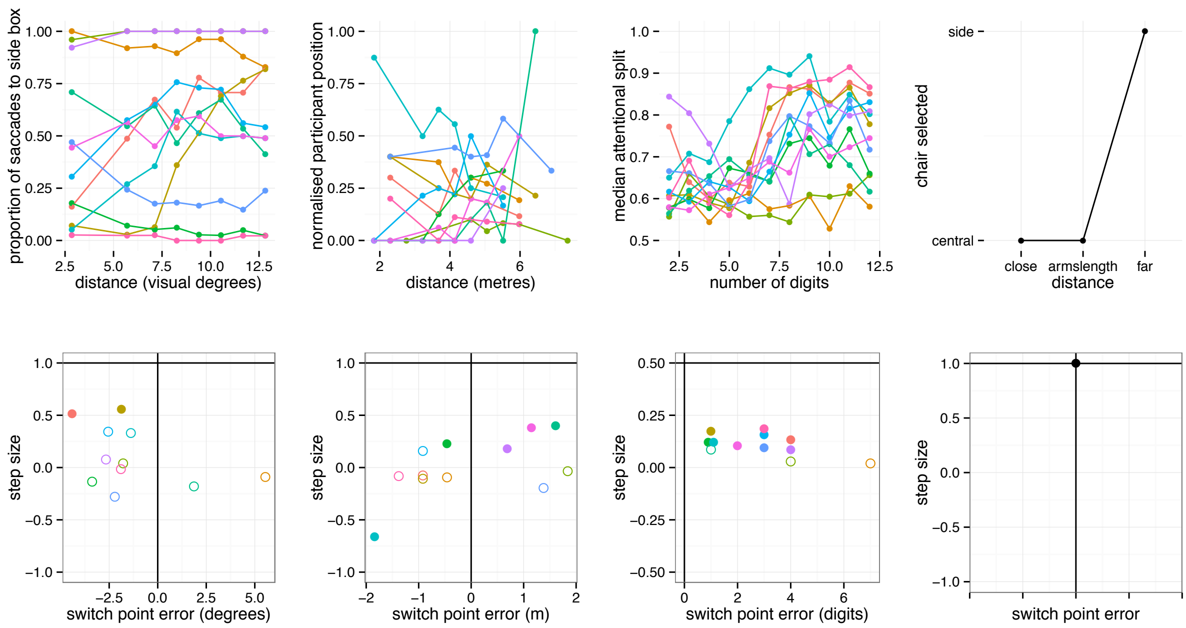

Figure 3: **First Row:** Choice behaviour for the four experiments. In the first three panels, each coloured line shows a different participant, while in the fourth panel, all twelve participants behaved in the same way, as illustrated by the black line. In each experiment, as task difficulty increases (moving along the x-axis), and optimal participant would exhibit a step function similar to that seen in the bottom row of Figure 2, from maximising their performance on both targets, to maximising their performance for one of the two targets. Participants generally failed to adopt this behaviour, except in the reaching task, where all participants behaved optimally. **Second row:** Results of analysis described in text in which step functions were fitted to the individual data shown above. The x-axis shows the position of the step relative to the optimal location, and the y-axis shows the size of the step. Optimal behaviour in each dimension is represented by the two black lines on each plot. Filled dots: participants' model fit $R^2 > .1$; Open dots: $R^2 < .1$.

A Perfectly rational behaviour would lead to $c_1 = 0$, $c_2 = 1$ [$c_1 = 0.5$ in the memory study], with s equal to the point predicted by the performance of each individual.

B A participant may behave rationally, but with a biased or noisy estimate of their own ability. This would lead to $c_1 \approx 0$ and $c_2 \approx 1$, but with a value of s that does not match the optimal switch point and/or unexplained variance (a low R^2).

C If participants fail to behave rationally, but still modify strategy with task difficulty, c_2 should be larger than c_1, but fall short of the maximum step size. For example, in the detection study, a step with $c_1 = 0$ and a $c_2 = 0.5$ (leading to a step size of 0.5) would mean the participant always fixated the central box when the boxes were close together, and fixated the side box on half the trials with a large separation.

D Participants may not modify behaviour at all or do so in the wrong direction, leading to no step or a reversed step (further indicated by an R^2 close to 0).

All the model fits from this analysis are given in Table 1 of the supplementary information, and summarized in Figure 3 (second row). We focus in this figure on the size of the step, which should be 1 (0.5 in memory study) and the position of the step relative to its predicted location under an optimal strategy. As can be seen in this figure, all participants in the reaching experiment are perfectly described by the step function (Category A of the scheme laid out above). Of the three other studies, no participants could be described as being in Categories A and B. Only in the memory experiment was the step function a reasonable model for participant behaviour (9 participants had an $R2 > .1$ as indicated by the filled circles, falling into Category C). In the other two studies, step size, direction, and location were generally not consistent with a choice behaviour that was modified by task difficulty (Category D), with the possible exception of a few participants.

In Figure 4 we compare overall accuracy in Session 2 (observed) to accuracy expected if an optimal strategy, as well a simple reference strategy, had been adopted by each participant. Accuracy of the optimal strategy was calculated for each distance (or number of digits) as described in the methods section above. This value was calculated individually for each participant based on their psychometric curve. Expected optimal accuracy as shown in Figure 4 represents the mean over all distances/difficulties. For saccades, throwing, and reaching, we used expected performance from the central location as a reference ("central"), and in the memory task the reference is expected performance from only looking at one number ("single"). We can see that for the saccade, throwing and memory tasks participants manage to outperform the reference, but the majority of them fail to achieve optimal performance. This difference is significant when evaluated in a paired t-test comparing "observed" accuracy to "optimal" (detection task: $t(11) = -2.45$, $p = 0.017$; throwing task: $t(11) = -3.65$, $p = 0.002$; memory task: $t(11) = -3.98$, $p = 0.001$). It should be noted that this way of illustrating the magnitude of the observed-optimal differences downplays our effect compared to if we had excluded conditions in which the reference and optimal behaviours are identical (such as very short distances in the throwing task where the optimal behaviour is to stand halfway between the targets).

For one participant in the detection task, there was no difference in their central and optimal strategy; this occurred because their visual acuity was good enough to perform above chance even at the largest eccentricity. Similarly, across detection, throwing, and memory tasks, several participants achieved accuracy in Session 2 that was higher than our predictions based on their Session 1 performance, likely due to practice effects. This suggests our estimate of optimal accuracy is conservative.

4 Discussion

We observed a striking failure to make optimal decisions in three of the four tasks presented above. In the fourth, where the task was to choose a seat from which to reach one of two beanbags, people were all able to select a chair close to one or the other beanbag when the two beanbags were too far away to reach from the central chair. This result demonstrates that our participants are able to understand the instructions and the constraints on their decision well enough to behave sensibly when the task is trivial. Why are they seemingly unable to make this decision in the other situations? The possible explanations fall into three general categories.

First, participants may fail to estimate their own performance accurately. A difference between the reaching task and the others is that in the former, the choice depends on the length of one's arm, while the others involve learning and remembering the limitations of one's own visual acuity, throwing skill, and memory - arguably more abstract and difficult to estimate. An inability to estimate performance seems unlikely, however, given the extensive practice participants had in the first session with the range of distance and difficulty levels presented in the second. Performance changes across these manipulations were stable and systematic (this is clear from the individual curves presented in the supplementary information), and people have been previously shown to be able to accurately estimate and make decisions based on expected performance (e.g. Barthelmé and Mamassian, 2009, Paunonen and Hong, 2010). Moreover, if people were estimating performance incorrectly or imprecisely, we would expect there to be a switch in strategies with increasing task difficulty at some point, but participants would not switch consistently or at the optimal level of difficulty (i.e. we would expect to see some evidence of bounded rationality, Simon, 1991). This describes what we denoted as Category B behaviour in the results of the model fit. No participants were well described by this category, suggesting a more global failure.

Second, participants may fail to frame the decision correctly. Achieving optimality requires the participants to make a logical decision (whether to invest in one option or both), followed sometimes by an entirely arbitrary decision (which target/option to invest in). Participants are able to make this pair of decisions effectively in the reaching task, demonstrating that they are capable of understanding the decision and its outcome in this very simple context. Perhaps the additional performance demands in the detection, throwing, and memory tasks distract participants from framing

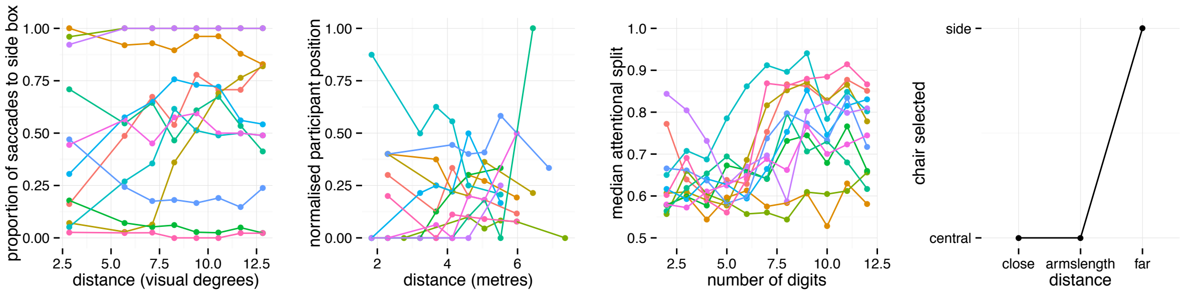

Figure 4: Accuracy of each participant in each experiment ("observed") relative to how accurate we would expect each observer to be given a simple reference strategy and an optimal strategy. Note: these figures take the average performance over a range of stimuli, including cases in which there is no difference in central/single and optimal performance. Hence these figures under-estimate the size of the effect. In the reaching task, the observed accuracy is higher than optimal because eight out of twelve participants happened to choose the seat in front of the randomly-selected reach target (within the variation we would expect based on chance)

the task appropriately, or trial-to-trial changes in task difficulty prevent participants from setting a single threshold at which to switch between strategies. During debriefing, we asked participants in the throwing experiment to tell us how they arrived at their decisions, and some were arbitrary, while others became focused on finding some pattern in the order of targets selected (even though they were told this was random). More participants fell into Category C in the memory task than in the detection and throwing tasks (i.e. they modified their behaviour with difficulty). Unlike in the other experiments, in which decisions were discrete, behaviour in the memory task unfolds over a 5-second interval, so the participants may have been able to learn from the cumulative effect of their choice behaviour more effectively.

A third possible explanation is that participants are prioritising something other than accuracy in the task. Most (but not all) participants were biased towards investing in both potential targets rather than focusing on one. Performing the task in more difficult circumstances may be seen as a challenge or an opportunity for learning, while selecting one or the other option takes away the challenge and puts the outcome in the hands of chance, which may be seen as a failure to be responsible for the outcome. Relatedly, it may be a particularly unpleasant experience to guess incorrectly and have the non-selected option turn out to be the target, and participants who happen to experience this loss on a series of trials in a row may have been discouraged from investing in a single option on subsequent trials. It should be noted that in the detection experiment of Morvan and Maloney (2012) participants were given substantial monetary rewards for accuracy, which does not seem to have made them any more likely to adopt an optimal strategy. Nonetheless, our participants were not explicitly rewarded for accuracy, so we cannot rule out that some of them may have decided to prioritise their own interest/pleasure in the task over accuracy.

There are many ways to be sub-optimal, and the fact that no single explanation from those listed above can account for all the results suggests that they all play a possible role to some extent and in some individuals. Nonetheless, only in the reaching experiment do participants demonstrate behaviour that could be classified as optimal. It is easy to imagine many scenarios in which the decision to invest all one's resources in one goal versus to divide resources between two goals would have serious consequences for an organism's survival (e.g. offspring investment; foraging). Given this, why is our ability to make a logical choice under these circumstances so easily disrupted? As situations become more complex, with increasing numbers of tasks and goals and decreasingly reliable ways of estimating likely success, the computations involved in determining the optimal strategy become more resource-intensive and time-consuming, and the potential pay-off diminishes (e.g. DeMiguel, Garlappi, and Uppal, 2009). It is our suggestion that, as a consequence of the complexity involved in deciding between multiple goals in most situations, people in general fail to employ a sensible strategy even when the required computations are extremely simple. This leads to the paradoxical conclusion that people's choices about how to allocate resources across multiple tasks are probably not optimal in principle, but they are usually adequate for complex situations. It is only as the tasks become fewer and the situation simpler that the failure to adopt a sensible strategy becomes both more apparent, and also more detrimental.

Author Contributions

ADFC and ARH developed the study concept and design. ADFC performed the data analysis. ADFC and ARH co-wrote the manuscript and approved the final version for submission.

Acknowledgments

The authors thank Alex Irvine, Melissa Spilioti, and a group of third-year students for their help with data collection, and Peter Neri, Rob Macintosh, Sandrina Ritzmann, and James Sheils for valuable input. This work was supported by a James S McDonnell Foundation Scholar Award (ARH).

References

Simon Barthelmé and Pascal Mamassian. Evaluation of objective uncertainty in the visual system. *PLoS Comput Biol*, 5(9):e1000504–e1000504, 2009.

Victor DeMiguel, Lorenzo Garlappi, and Raman Uppal. Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *Review of Financial Studies*, 22(5):1915–1953, 2009.

Martin Gardner. Matheatical games. *Scientific American*, October:180–182, 1959.

M. Hayhoe and D. Ballard. Modeling task control of eye movements. *Current Biology*, 24(13):R622–R628, 2014.

Amelia R Hunt and Raymond M Klein. Eliminating the cost of task set reconfiguration. *Memory & cognition*, 30(4):529–539, 2002.

Daniel Kahneman and Amos Tversky. Choices, values, and frames. *American psychologist*, 39(4):341, 1984.

Melissa M Kibbe and Eileen Kowler. Visual search for category sets: Tradeoffs between exploration and memory. *Journal of vision*, 11(3):14, 2011.

Konrad P Körding and Daniel M Wolpert. Bayesian integration in sensorimotor learning. *Nature*, 427 (6971):244–247, 2004.

George A Miller. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological review*, 63(2):81, 1956.

Camille Morvan and Laurence T. Maloney. Human visual search does not maximize the post-saccadic probability of identifying targets. *PLOS Computational Biology*, 8(2), 2012.

J. Najemnik and W. S. Geisler. Optimal eye move ment strategies in visual search. *Nature*, 434:387–391, 2005.

Ipek Oruç, Laurence T Maloney, and Michael S Landy. Weighted linear cue combination with possibly correlated error. *Vision research*, 43(23):2451–2468, 2003.

Sampo V Paunonen and Ryan Y Hong. Self-efficacy and the prediction of domain-specific cognitive abilities. *Journal of personality*, 78(1):339–360, 2010.

Herbert A Simon. Bounded rationality and organizational learning. *Organization science*, 2(1):125–134, 1991.

Nir Vulkan. An economists perspective on probability matching. *Journal of economic surveys*, 14(1):101–118, 2000.

Daniel M Wolpert and Michael S Landy. Motor control is decision-making. *Current opinion in neurobiology*, 22(6):996–1003, 2012.

Shih-Wei Wu, Mauricio R Delgado, and Laurence T Maloney. Economic decision-making compared with an equivalent motor task. *Proceedings of the National Academy of Sciences*, 106(15):6088–6093, 2009.

Hang Zhang, Camille Morvan, Louis-Alexandre Etezad-Heydari, and Laurence T Maloney. Very slow search and reach: Failure to maximize expected gain in an eye-hand coordination task. *PLoS computational biology*, 8(10):e1002718, 2012.

Supplementary Materials

See figures below.

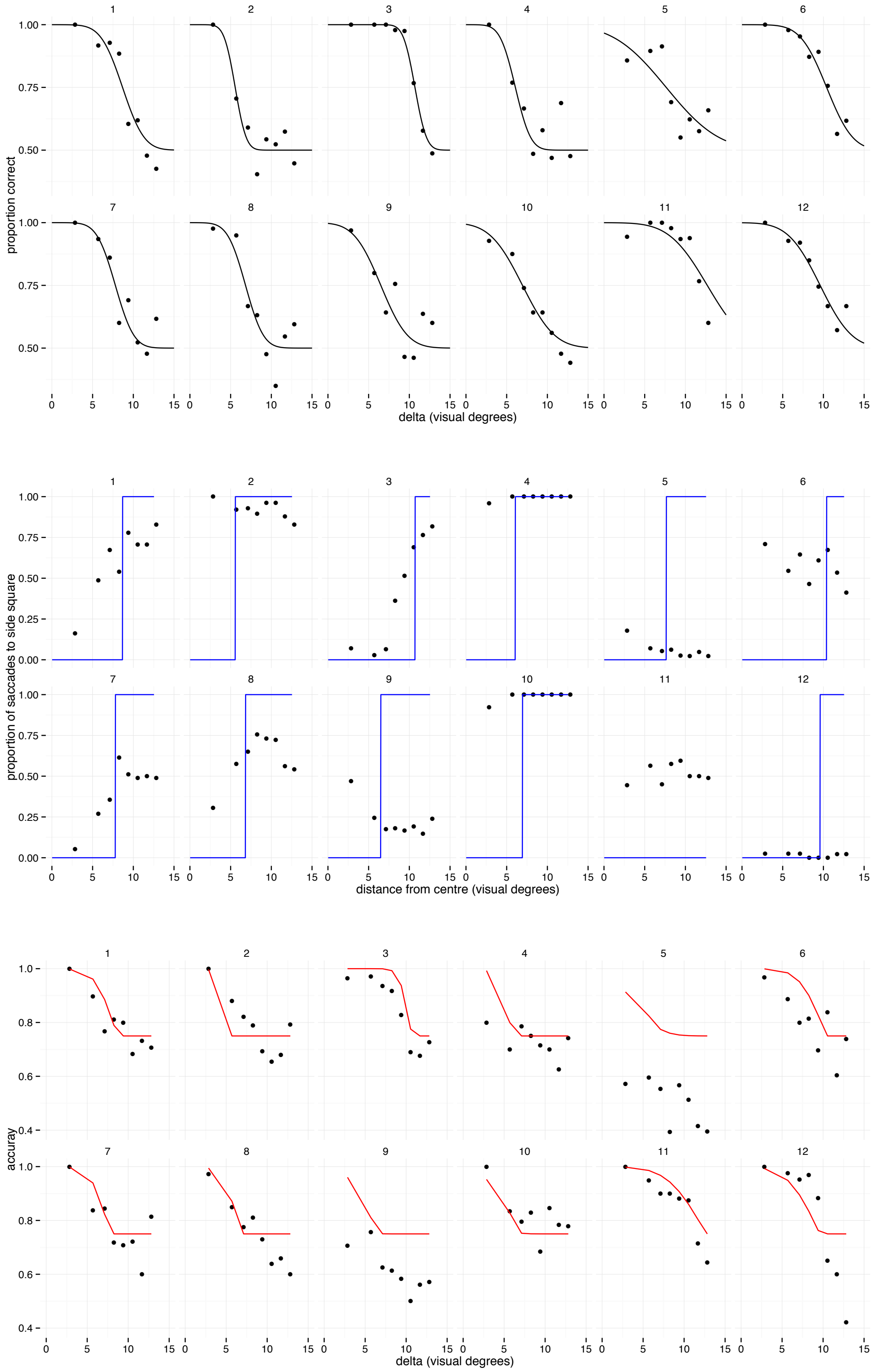

Figure SupMat.1: The complete results from the 12 participants in the detection experiment. Top set: results from Session 1. The black line is the psychometric curve fit to their data. Middle set: Proportion of saccades to the side square in Session 2 (black dots) compared to an estimate of optimal fixation behaviour based on Session 1 performance (blue line). Bottom set: Detection accuracy (black dots) compared to predicted accuracy under a baseline (blue line) and optimal (red line) strategy

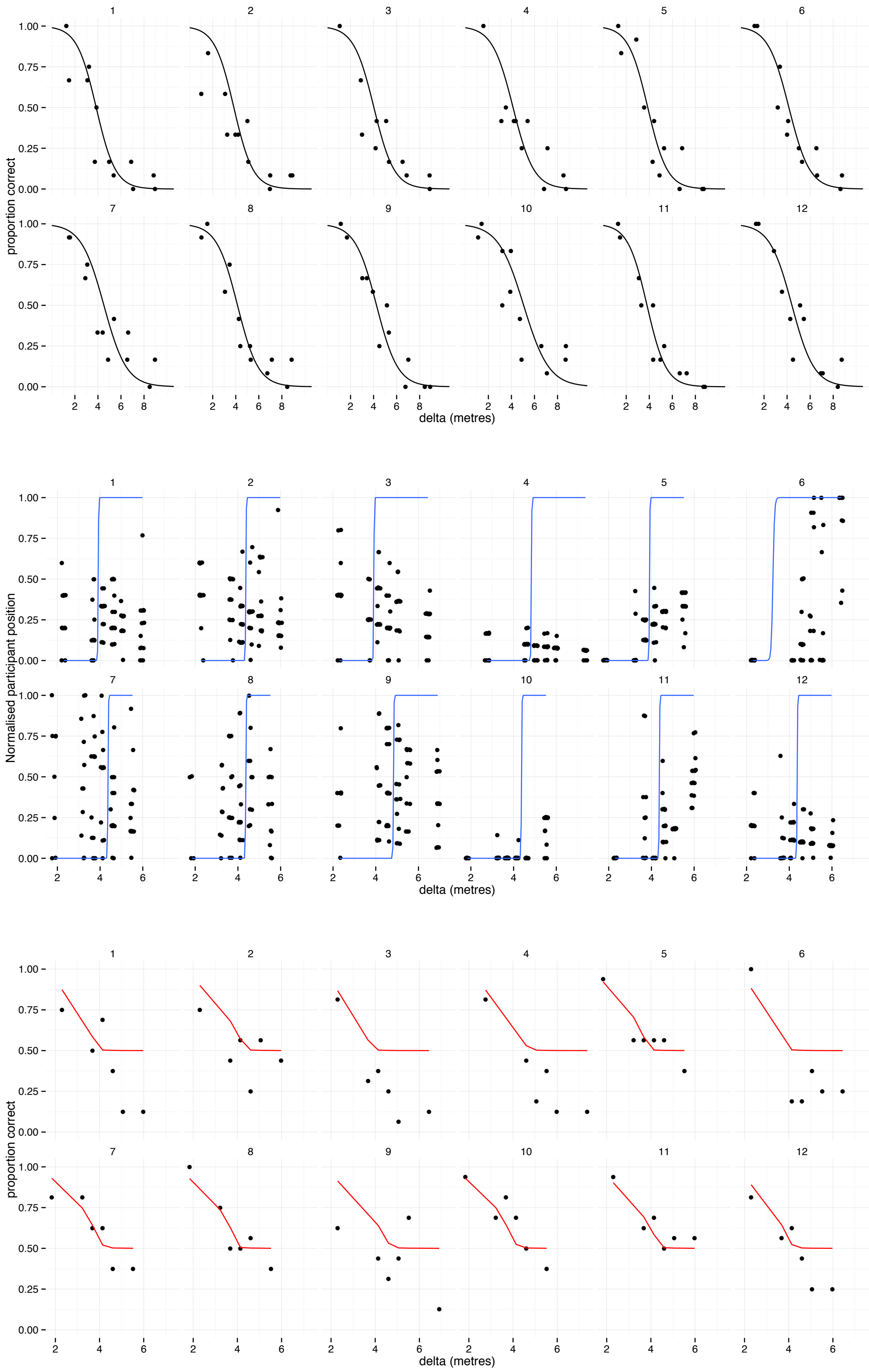

Figure SupMat.2: The complete results from the 12 participants in the throwing experiment. Top set: results from Session 1. The black line is the psychometric curve fit to their data. Middle set: standing position at each of six distances in Session 2 (black dots) compared to an estimate of optimal standing position based on Session 1 performance (blue line). Y-axis range has been restricted to 0-1, which removed data points for participants 7, 8, and 9, who sometimes stood outside the range of the hoops. Bottom set: Throwing accuracy (black dots) compared to predicted accuracy under an optimal (red line) strategy.

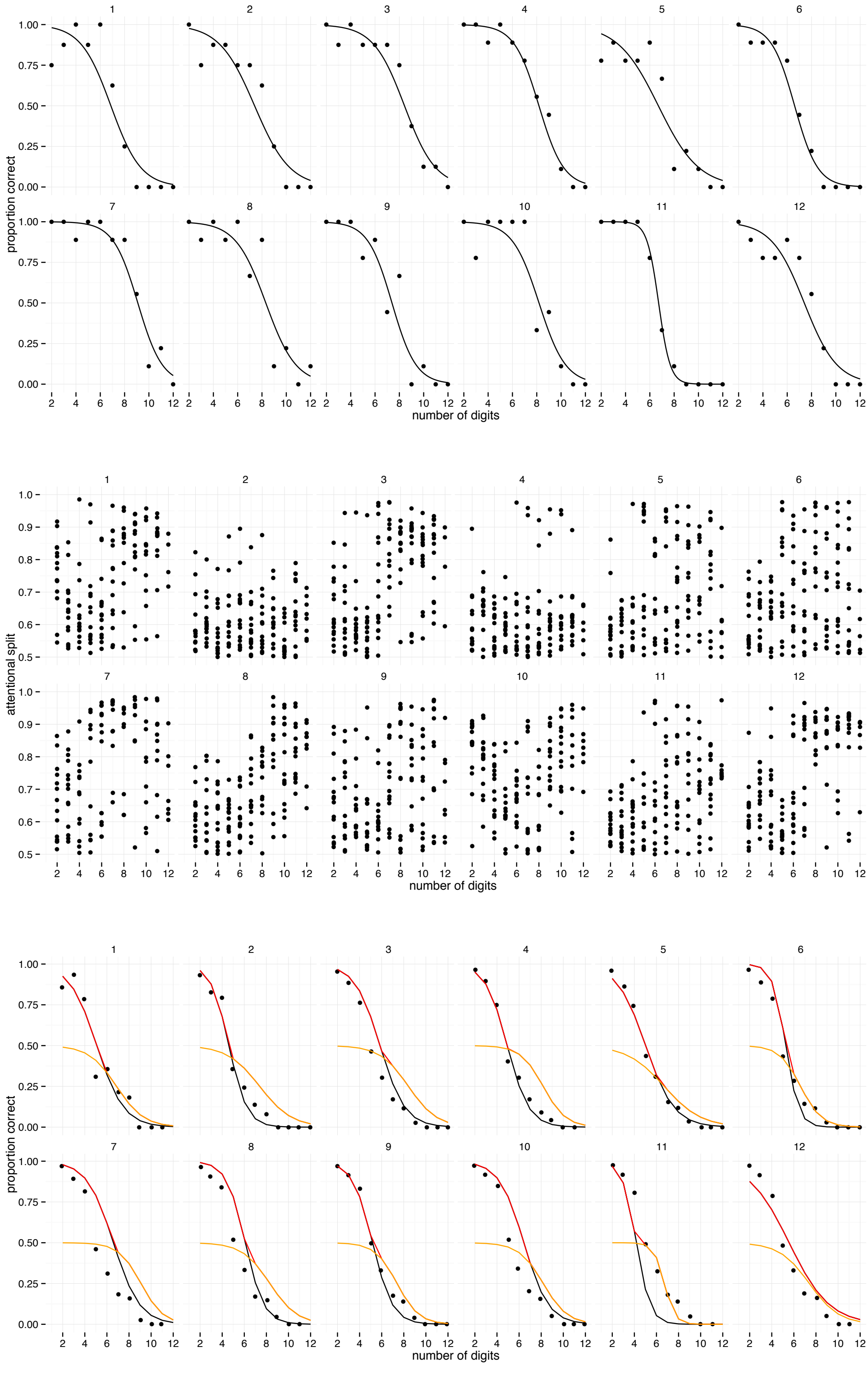

Figure SupMat.3: The complete results from the 12 participants in the memory experiment. Top set: results from Session 1. The black line is the psychometric curve fit to their data, giving $\phi(n)$ as the probability of remembering an n-digit number. Middle set: Attentional split (relative proportion of time spent on one digit, black dots) in Session 2. Bottom set: Memory accuracy in Session 2 (black dots) compared to predicted accuracy under an optimal strategy, calculated as the maximum of the red ($\phi(n)$) and orange ($0.5\phi(n)$) lines

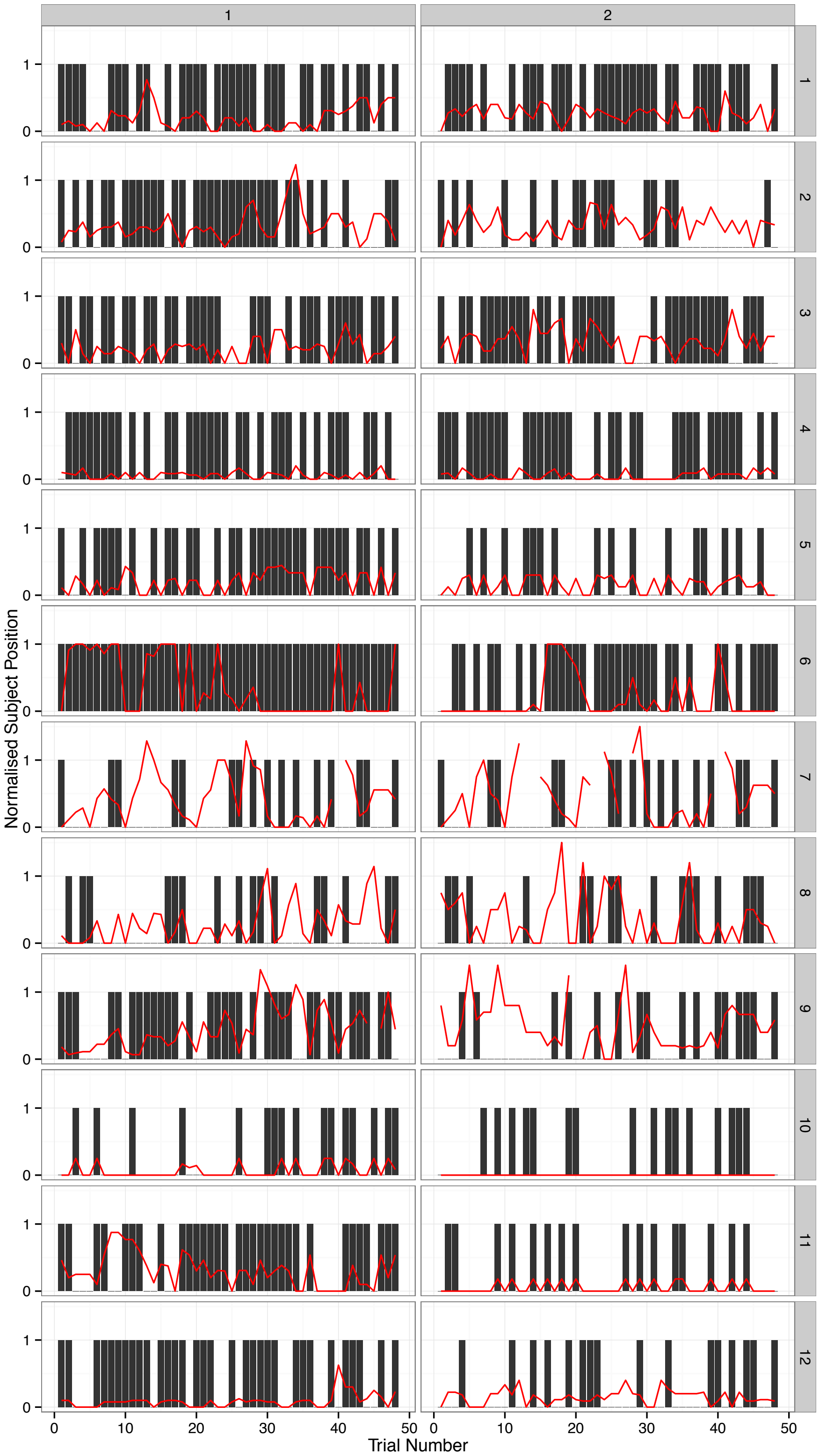

Figure SupMat.4: This figure shows trial-to-trial sequential variation in participant standing position (the red line) in the throwing experiment. The first column shows the first block, the second column shows the second block. Each row corresponds to an individual participant. The black bars indicate the trials in which the optimal strategy is to stand by the hoop, otherwise they should stand at the mid-point. Discontinuities in the red line correspond to trials in which the participant stood further than 1.5 normalized distance units from the center (Y axis range has been restricted to 1.5 to improve visibility).

Experiment	Participant	s	s_{opt}	c_1	c_2	$c_2 - c_1$	R^2
detection	1	4.3	8.7	0.32	0.84	0.52	0.11
detection	2	11.1	5.6	0.15	0.06	-0.09	0.02
detection	3	8.8	10.7	0.31	0.86	0.56	0.32
detection	4	4.3	6.1	0.00	0.04	0.04	0.04
detection	5	4.3	7.7	0.96	0.82	-0.14	0.03
detection	6	12.3	10.4	0.59	0.41	-0.18	0.02
detection	7	6.4	7.8	0.51	0.84	0.33	0.08
detection	8	4.3	6.8	0.35	0.69	0.34	0.05
detection	9	4.3	6.5	0.81	0.53	-0.28	0.03
detection	10	4.3	6.9	0.00	0.08	0.08	0.07
detection	11	4.3	-	0.48	0.56	0.08	< 0.01
detection	12	7.7	9.6	0.99	0.98	-0.02	< 0.01
throwing	1	3.0	3.9	0.29	0.21	-0.08	0.03
throwing	2	3.9	4.4	0.39	0.29	-0.09	0.05
throwing	3	3.0	3.9	0.36	0.26	-0.11	0.05
throwing	4	6.7	4.8	0.06	0.03	-0.04	0.05
throwing	5	3.5	3.9	0.02	0.25	0.23	0.55
throwing	6	4.8	3.2	0.08	0.48	0.40	0.25
throwing	7	2.5	4.4	1.09	0.43	-0.66	0.21
throwing	8	3.5	4.4	0.21	0.37	0.16	0.05
throwing	9	6.2	4.8	0.54	0.35	-0.20	0.03
throwing	10	5.1	4.4	0.00	0.18	0.18	0.70
throwing	11	5.5	4.4	0.12	0.50	0.38	0.41
throwing	12	3.0	4.4	0.19	0.11	-0.08	0.08
memory	1	7.5	3.5	0.68	0.82	0.13	0.23
memory	2	10.5	3.5	0.60	0.62	0.02	0.01
memory	3	5.5	4.5	0.62	0.80	0.17	0.32
memory	4	8.5	4.5	0.61	0.64	0.03	0.02
memory	5	4.5	3.5	0.60	0.72	0.12	0.14
memory	6	4.5	3.5	0.63	0.71	0.09	0.08
memory	7	5.5	4.5	0.69	0.81	0.12	0.15
memory	8	7.5	4.5	0.63	0.79	0.16	0.38
memory	9	6.5	3.5	0.65	0.74	0.09	0.11
memory	10	8.5	4.5	0.71	0.80	0.09	0.11
memory	11	5.5	3.5	0.61	0.71	0.10	0.18
memory	12	6.5	3.5	0.65	0.83	0.19	0.39

Table 1: s=location of the step; S_{opt} = predicted value for s based on participants performance curve; c_1 and c_2 = value of the fitted model before and after the step. Minimum c_1 in the memory study is 0.5. Participant 11 in the detection experiment had unusually good peripheral vision, leading to no point at which they should switch from looking in the center (hence no S_opt). In the reaching experiment, for all participants, $s = s_opt$, $c_1 = 0$; $c_2 = 1$ and $R^2 = 1$.