

Failure of Intuition When Choosing Whether to Invest in a Single Goal or Split Resources Between Two Goals



Alasdair D. F. Clarke and Amelia R. Hunt

School of Psychology, University of Aberdeen

Psychological Science
2016, Vol. 27(1) 64–74
© The Author(s) 2015
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0956797615611933
pss.sagepub.com
 SAGE

Abstract

In a series of related experiments, we asked people to choose whether to split their attention between two equally likely potential tasks or to prioritize one task at the expense of the other. In such a choice, when the tasks are easy, the best strategy is to prepare for both of them. As difficulty increases beyond the point at which people can perform both tasks accurately, they should switch strategy and focus on one task at the expense of the other. Across three very different tasks (target detection, throwing, and memory), none of the participants switched their strategy at the correct point. Moreover, the majority consistently failed to modify their strategy in response to changes in task difficulty. This failure may have been related to uncertainty about their own ability, because in a version of the experiment in which there was no uncertainty, participants uniformly switched at an optimal point.

Keywords

decision making, optimal behavior, open data

Received 3/12/15; Revision accepted 9/23/15

A goalie can choose to stand in the middle of the goal or to one side or the other. A student can choose to study all the course material or to focus on deeper learning of a subset. 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 to achieve them all. Otherwise, people are better off focusing their resources on only one task or goal at the expense of the others. We report results from three experiments that all converge—despite their very different methods—on the conclusion that humans are surprisingly deficient at achieving optimal outcomes. When presented with a choice between dividing available resources between two goals and investing all their resources in one goal, the participants were poor at choosing the best strategy, even though the factors they needed to take into account were relatively stable and limited.

In this article, we define *optimal decisions* as those that achieve the best possible outcome while minimizing energy expenditure and risk. There are many examples of optimal or near-optimal decision making in humans (e.g., Kibbe & Kowler, 2011; Kording & Wolpert, 2004; Najemnik & Geisler, 2005; Oruc, Maloney, & Landy, 2003). Wolpert and Landy (2012), for example, have argued that motor control successfully maximizes action consequences given constraints of task, motor, and sensory uncertainty. However, other researchers have demonstrated human failures to maximize expected gain in more deliberative human decisions (Gardner, 1959; Kahneman & Tversky, 1984; Morvan & Maloney, 2012; Vulkan, 2000; Zhang, Morvan, Etezzad-Heydari, & Maloney, 2012).

Our interest in optimal decision making began with an intriguing contradiction in the visual-search literature.

Corresponding Author:

Amelia R. Hunt, School of Psychology, University of Aberdeen,
Aberdeen AB24 3XB, Scotland
E-mail: a.hunt@abdn.ac.uk

One influential model of search (Najemnik & 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. However, Morvan and Maloney (2012) recently provided striking evidence that human observers do not reliably fixate locations that maximize 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 observers take into account the probability of detecting the target at a given eccentricity when deciding where to fixate, they should fixate a location in between the two boxes when the boxes are relatively close together. As the boxes move farther apart, they will reach an eccentricity at which it is no longer possible to discriminate the target in either box at a level above chance. At this point, observers should switch to a strategy of fixating one box or the other, because this will yield accuracy close to 100% if the target happens to appear inside the fixated box and chance-level accuracy if it does not. Surprisingly, Morvan and Maloney found that all 4 of their observers failed to maximize their target discrimination performance; their choice of whether to fixate the center or a target box did not vary with the distance between the boxes. Morvan and Maloney proposed that saccade target selection is based largely on heuristics, such as a tendency to make saccades in particular directions, rather than on visual sensitivity and uncertainty.

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 behavior that depend on a mechanism that maximizes information gain (Hayhoe & Ballard, 2014; Najemnik & Geisler, 2005). In our first experiment, therefore, we replicated the method of Morvan and Maloney (2012) using a larger sample. We then were interested in establishing whether the failure of optimal behavior that they reported could be considered specific to the context of eye movements and detection of targets or was attributable to a larger problem that pervades 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 evidence that rapid motor responses achieve different outcomes than deliberative decisions do (Hunt & Klein, 2002; Wu, Delgado, & Maloney, 2009). The aim of our study was to investigate whether participants would 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. Although the experiments varied in terms of task and response modality (detection, throwing, memory, and

reaching), they all involved the same essential decision-making paradigm: experimentally creating a point at which, to achieve the best possible outcome, it is necessary to switch between dividing available resources across two goals and 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 or goal. The purpose of this session was to characterize each participant's performance across difficulty as well as to facilitate the participant's awareness of his or her 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 favor of the other. Participants' choices were compared with individualized estimates of what they should have chosen, according to their performance in the first session.

Method

The motivation and logic of the four experiments were similar, so we report the method for all four together. Morvan and Maloney (2012) found choice behavior that was idiosyncratic but clearly not optimal in the 4 observers they tested. We expected to replicate this pattern, but we 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 each experiment. An Eyelink 1000 (SR Research, Mississauga, Ontario, Canada) was used to record eye movements in the detection and memory experiments described later in this section.

Participants

Forty-eight undergraduates at the University of Aberdeen were recruited to participate, 12 for each experiment. All participants were recruited via word of mouth, were naive to the aims of the experiment, and gave informed consent to participate. The experiments were reviewed and approved by the School of Psychology ethics committee.

Materials and procedure

Detection experiment. The aim of Session 1 of the experiment was to obtain a psychometric function for each participant for target detection, and to allow the participant to practice the detection task and gain familiarity with their own level of performance. The stimulus consisted of two gray boxes (length = 1.8° of visual angle) on

either side of a central fixation cross. The distance from the fixation point to the boxes varied (2.9° , 5.7° , 7.1° , 8.3° , 9.4° , 10.5° , 11.7° , or 12.8°). After the participant had maintained a stable central fixation for 1 s, the target (a small white circle) was presented at the top or bottom of one of the two boxes. The target was displayed for 500 ms, after which a blank gray screen was displayed. The trial was immediately cancelled if the participant broke central fixation. Participants reported whether the target had been presented at the top or bottom of one of the two boxes by pressing the up or down arrow key on a keyboard; they were instructed to 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 distance. No feedback was given. Individual performance was modeled in R (Version 3.2.1; R Development Core Team, 2015) using a generalized linear model with the *mafc*-probit link function from the *psyphy* package (Version 0.1-9; Knoblauch, 2014). This function was labeled $\phi(\delta)$, with δ defined as the distance from the fixation point to either box.

In Session 2, 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 right-most gray boxes (see Fig. 1a for an illustration of the trial display). This meant that the cross's position varied with the separation. We used the eight distances from Session 1, 48 repetitions of each, presented in random order. This gave a total of 384 trials. After a fixation was detected inside one of the three boxes, the target was presented in either the left box or the right box. Participants were told the target would never appear in the center box. As in Session 1, participants used a keyboard to report whether the target had appeared at the top or bottom of a box.

We used ϕ from Session 1 to derive each participant's optimal strategy and predicted accuracy. In this task, when the separation between the boxes is small, participants can direct their saccades toward the central box and have a good chance of detecting the target in either location. Once δ increases to the point at which $\phi(\delta)$ is less than .75, the participants should switch strategies and fixate either the left box or the right box. Once the participants fixate one of the boxes, there is a 50% chance that the target will appear at fixation (in which case, they have about a 100% chance of responding correctly) and a 50% chance that the target will appear at the other box (in which case, they have to guess and would be correct 50% of the time). Together, these values give an expected accuracy of 75%.

Throwing experiment. This experiment was analogous to the detection experiment, except that the task was to

throw a beanbag into one of two hoops. The participants did not know which hoop would be designated as the target when they were asked to choose a place to stand. For two hoops close to one another, the ideal position would be halfway between them. For hoops too distant for participants to throw accurately from a central location, however, the optimal behavior would be to stand close to one hoop, which would yield a success rate of 50%.

The experiment took place in a sheltered area of concrete slabs (for a picture of the area, see Fig. 1b). The slabs were 0.46×0.61 m, which made them useful markers for placing hoops and recording standing positions. In Session 1, participants stood in the center slab of the area marked with black tape. Flat hoops with a diameter of 0.40 m were placed at six different distances from the participants (1.88, 3.22, 4.14, 5.06, 6.90, and 8.74 m). Each participant tossed 12 beanbags for each hoop distance in one direction (i.e., left or right of the tape, from closest to farthest), and then repeated the tosses in the opposite direction (counterbalanced), for a total of 144 trials. The beanbag was cleared from the area after each toss. A trial was recorded as correct if the final resting place of the beanbag was inside or touching the specified hoop. No differences in direction were found, so we ignored this factor in subsequent analyses. Each participant's accuracy was modeled using logistic regression with a fixed intercept of -0.99 . That is, we assumed that participants were 99% accurate if they stood right next to a hoop. Each participant's curve (modeled the same way as in the detection experiment) was used to select six slabs on which to place hoops in Session 2. We based the distances on the slab at which participants were closest to 50% accurate—that is, where $\phi(\delta)$ was equal to .50—which we called Slab M. This was the point at which, to maximize accuracy, participants should have switched between standing in the center to standing closer to one hoop or the other.

In both blocks of Session 2, red hoops were always closest to the center, yellow hoops were next, and blue hoops were farthest away. For the first block of Session 2, six hoops were taped down onto six different slabs, three in each direction relative to an unmarked center point: Slab M - 1 (red hoops), and Slab M + 1 (yellow hoops), and the slab at which participants had expected accuracy of 10% (blue hoops). The second block had the same configuration, but the hoops were taped down on the slab at which participants had expected accuracy of 90% (red hoops), Slab M (yellow hoops), and Slab M + 2 (blue hoops). 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,

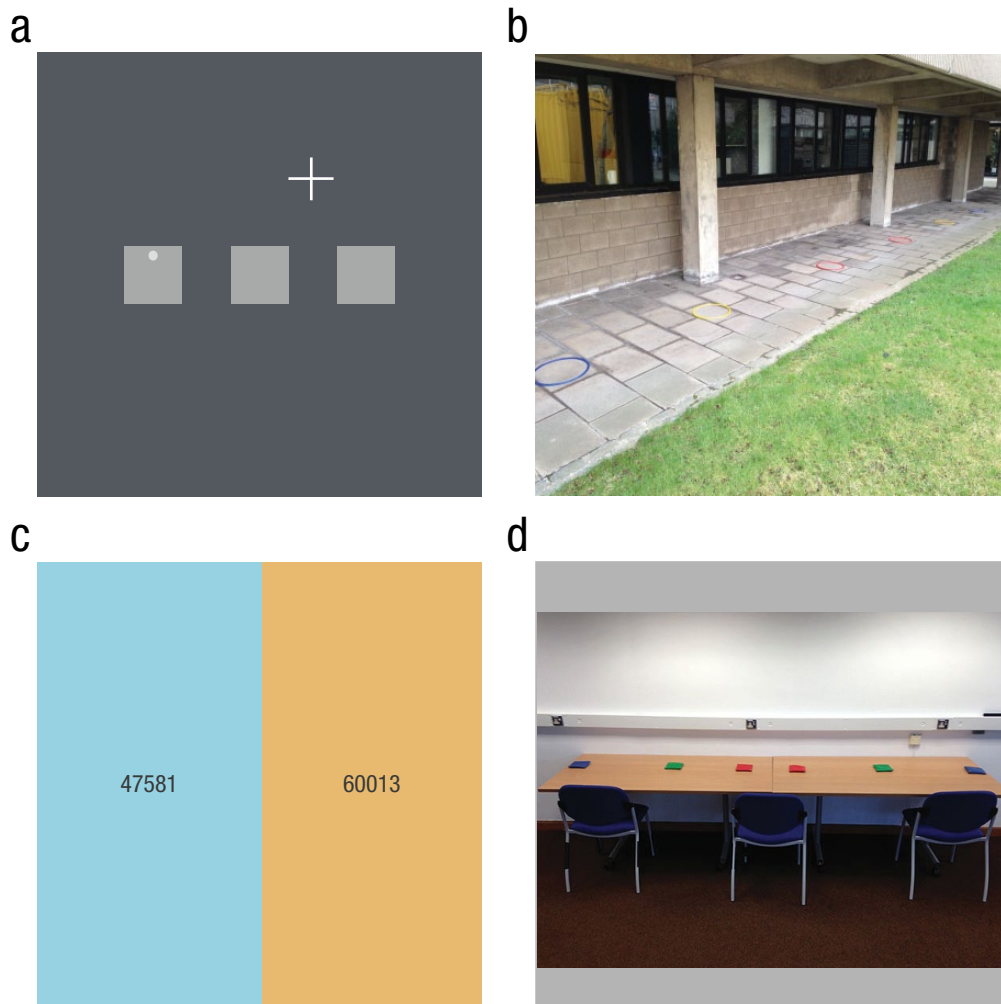


Fig. 1. Examples of the stimuli and setups for Session 2 of the detection, memory, and throwing experiments and the single session of the reaching experiment. On each trial of the detection experiment (a), participants fixated a cross and then shifted their gaze to one of three boxes. Next, the target (a small white circle) appeared inside the left or right box, shown here in the left box. Participants then reported whether the target had appeared at the top or bottom of the box. The eccentricity of the boxes varied across trials. In the throwing experiment (b), six hoops (two of each of three colors) were placed on the ground. Participants were given a beanbag matching the target color and chose a place to stand. They were then told which of the two hoops of the target color they were to try to toss the beanbag into. In the memory experiment (c), participants were asked to remember for later report one of two numbers presented on either side of a monitor. The number of digits varied from trial to trial. In the reaching experiment (d), six colored beanbags (two of each of three colors) were placed on a long table. Participants were told the color of the beanbag they would need to pick up and then chose which chair they wanted to sit in.

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 and throws in each block (16 trials for

each hoop color in a random order). The main experimenter stood on the grass to the side of the paved area, and participants returned to them after each trial to receive a new beanbag. The other experimenter cleared the beanbags and recorded tossing 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 to enable quick and subtle recording of standing position). The order of colors and direction of throw was randomized separately for each participant.

Memory experiment. In this experiment, participants were shown two numbers and were later asked to report only one of them. At the time of presentation, the participants did not know which number they would have to report. In this task, if the two numbers have a small number of digits and are therefore easy to remember, the ideal behavior is to look at, and memorize, both numbers. However, as the number of digits increases to the limits of working memory, the optimal behavior is to focus on just one of the numbers and ignore the other one.

In Session 1, we measured each participant's digit span. Each stimulus consisted of a randomly generated sequence of n digits (between 2 and 12). On each trial, this digit sequence was displayed on the left side or the right side of the screen. Each of the sides had one of two differently colored backgrounds. The background colors swapped randomly on a trial-to-trial basis. The number was displayed for 5 s, followed by a gray screen for 3 s. Finally, a response screen was presented with the same background colors as were presented with the digit sequence, with the prompt "please enter the number" displayed where the digits had appeared. Participants then entered the sequence as they remembered it, using the number keypad on a computer keyboard. A response was considered correct if all the digits were typed in the correct order. Each digit length was used nine times, for a total of 99 trials. Trials were presented in a random order, and participants were given a break halfway through. As for the earlier experiments, we modeled each participant's accuracy using logistic regression; $\phi(n)$ was the probability of remembering an n -digit number.

Session 2 of the experiment was similar to Session 1. The main differences were that participants' eye movements were tracked while carrying out the task, and they were presented with two sequences of n random digits (2 to 12, as in Session 1) to memorize (Fig. 1c). In any given trial, the digit sequences were equal in length. When the response screen was shown, the location of the text indicated whether participants should report the left or the right digit sequence—if the text appeared on the left, participants were to report the digits that previously appeared on the left, and if it appeared on the right, participants were to report the digits that previously appeared on the right. The colored background was used as an additional prompt. Random digit sequences were presented 15 times for each value of n , for a total of 165 trials. Trials were presented in random order.

Eye-tracking data were analyzed by assigning fixations to one of two $14^\circ \times 2.8^\circ$ areas of interest centered on the two numbers. Fixations that fell outside of these areas were discarded. Attentional split was then defined as the proportion of time spent fixating the area of interest that received more attention. Thus, a value of .5 indicated that

a participant spent equal time looking at the two digit sequences, whereas a value of 1.0 indicated that a participant fixated one digit sequence for the entire time. In this experiment, deriving predicted accuracy given an optimal strategy was not as straightforward as it was in the two previous experiments. We could estimate the probability of remembering both numbers as $\phi(2n)$, but the data showed that this underestimated performance for small values of n (see the Supplemental Material available online), presumably because of chunking (Miller, 1956). We assume that for small values of n , our participants memorized both numbers. However, as n increased, the task became increasingly difficult, and participants should have changed strategy and attempted to remember only one number; the probability of a correct trial would then become $.5\phi(n)$.

Reaching experiment. In the final experiment, we took participants' basic choice to a trivially simple level. We wanted to be certain that our previous results were not a consequence of participants misunderstanding the instructions. Six beanbags were placed on a long table (Fig. 1d). Two red beanbags were near the center, a green beanbag was placed halfway to each end, and a blue beanbag was placed at each end. Three chairs were placed at the table (left, center, and right). Participants were asked to sit in the center chair and to try and reach, with their backs still touching the chair, the red, green, and blue beanbags (thus demonstrating their own reach span, as in Session 1 of the previous experiments). Participants were then asked to stand, and the experimenter told them that they would be picking up a beanbag of a specified color. They were then asked to choose one of the seats and to sit in it. Participants were not told which of the two beanbags of the specified color they would have to pick up until after they had selected a chair. Participants selected a chair once for each of the three colors. The order of specified colors and which beanbag was to be picked up was randomized for each participant.

Results

The top row of Figure 2 shows proportion correct in Session 1 for typical participants in the detection, throwing, and memory experiments. Accuracy data for all participants are available in the Supplemental Material. In the bottom row of Figure 2, the same participants' decision behavior in the second session is compared with the optimal strategy derived from their Session 1 performance (the dashed line).

The top row in Figure 3 shows the decision behavior of all participants in all four experiments. In the first three experiments, the overwhelming majority of participants failed to systematically change their behavior as

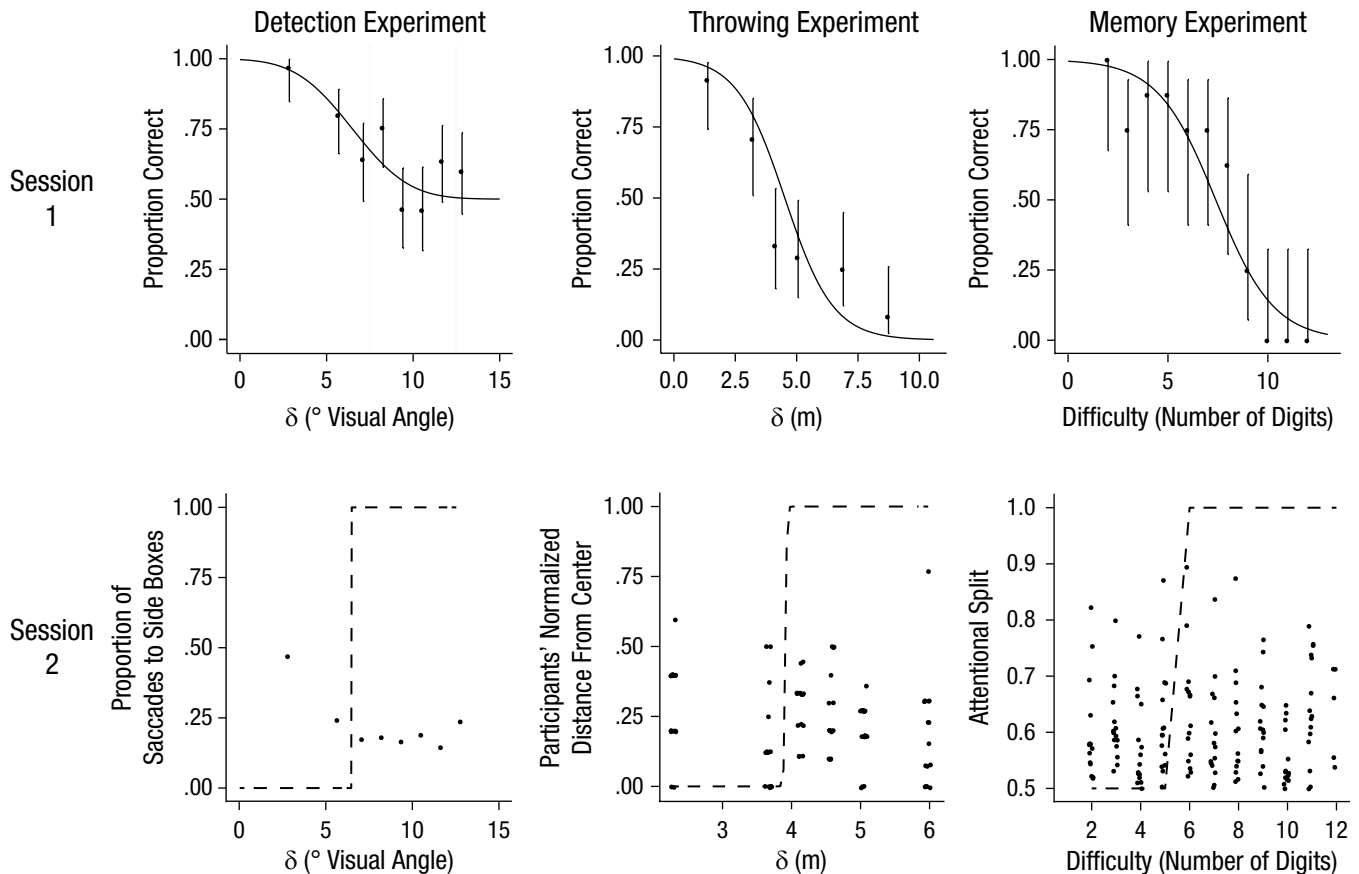


Fig. 2. Results from representative participants in the detection, throwing, and memory experiments. The top row shows proportion correct in Session 1 as a function of distance (δ) from fixation (detection experiment), δ from center to hoop (throwing experiment), or difficulty (memory experiment). The curves indicate regression lines. Error bars indicate 95% binomial proportion confidence intervals. Because the detection experiment involved a two-alternative forced choice, chance performance was 50%. The bottom row shows decision behavior in Session 2 for the same participants whose accuracy is graphed in the top row. Proportion of saccades to the side boxes (detection experiment) and participants' normalized distance from center (throwing experiment) are graphed as a function of δ . Attentional split is graphed as a function of difficulty (memory experiment). In the detection experiment, the participant's choice was binary; hence, each dot represents the proportion of trials on which the participant fixated the side box in each condition. For the other two experiments, each dot represents behavior on a single trial. For the throwing experiment, participant's distance from the center was normalized such that a distance of 1 indicates a position at the hoop of the target color. In each graph, the dashed line indicates the optimal strategy that the participant should have adopted, given his or her performance in Session 1. Results from the other participants were similar to those shown here and are available in the Supplemental Material.

task difficulty increased. In the detection experiment, participants selected their own individual strategy (as in Morvan & Maloney, 2012), and they tended not to vary this strategy as difficulty increased. Only 1 of the 12 participants exhibited behavior that approached an optimal strategy.¹ For the throwing experiment, there were fewer trials than in the other experiments, and individual strategies were less consistent; nonetheless, in aggregate, the participants stood just as close to the center when throwing to hoops that were far away as they did when the hoops were close together. For this experiment, we also examined sequence effects at the trial level to determine whether participants had a tendency to learn or to persist over time with one strategy over another. There was no consistent pattern (see Fig. S4 in the Supplemental

Material). For the memory experiment, participants were as likely to fixate both digit sequences as they were to fixate only one, regardless of sequence length. However, several participants came closer to adopting a strategy that was optimal in this experiment, a detail we return to in the Discussion. In the reaching experiment, which was designed to check that participants could correctly understand the instructions used in the preceding studies, participants' behavior was uniformly optimal.

Each participant's choice behavior can be modeled by fitting a step function: $y = c_1$ for all $x \leq s$, $y = c_2$ for all $x > s$, where s is the point at which the participant switched strategies (e.g., from a center to a side strategy). A linear model would not be appropriate, given the nature of the optimal strategy depicted in Figure 2. We fit

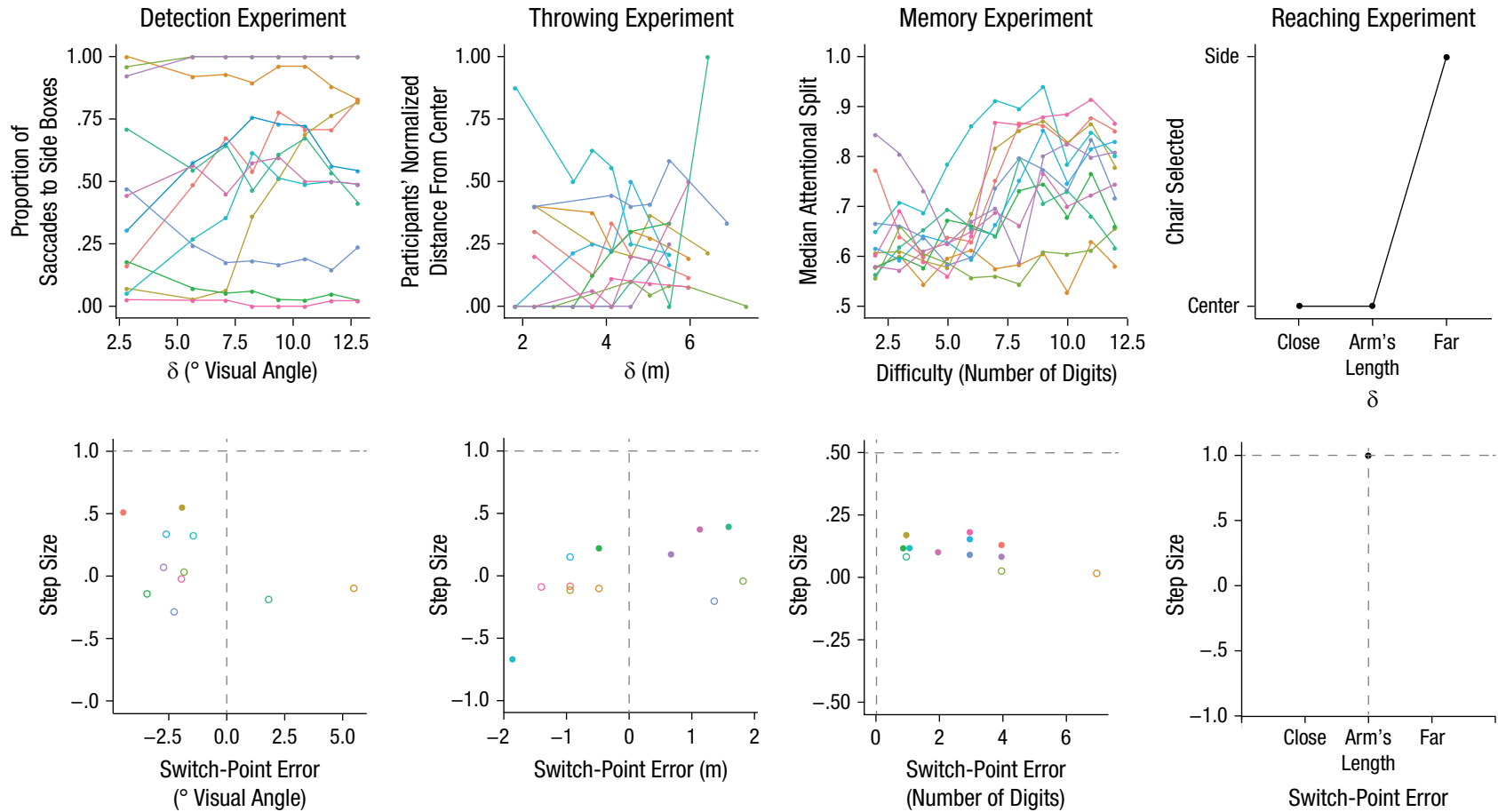


Fig. 3. Choice behavior in the four experiments (top row) and results of analyses in which step functions were fitted to these data (bottom row). In the top row, proportion of saccades to the side boxes (detection experiment), participants' normalized distance from center (throwing experiment), and chair selected (reaching experiment) are graphed as a function of distance (δ), and median attentional split is graphed as a function of difficulty (memory experiment). For the detection, throwing, and memory experiments, each line represents the behavior of a different participant. For the reaching experiment, the line indicates the behavior of all 12 participants. In the bottom row, step size is plotted as a function of switch-point error (i.e., the position of the step relative to the optimal location) for each experiment. For the detection, throwing, and memory experiments, each circle represents the behavior of a different participant. On each plot, optimal behavior in each dimension is indicated by a dashed line. Filled circles indicate a model-fit R^2 greater than .1, and open circles indicate a model-fit R^2 less than .1. For the reaching experiment, the black circle indicates the behavior of all 12 participants, whose performance was uniformly optimal.

s , c_1 , and c_2 to the data using least squares regression. From this model fit, we categorized behavior into four rough patterns:

1. Perfectly rational behavior would lead to a c_1 of 0 and a c_2 of 1 ($c_1 = .5$ in the memory experiment); s would be equal to the point predicted by the performance of each individual.
2. A participant might behave rationally, but with a biased or noisy estimate of their own ability. This would lead to a c_1 approximately equal to 0 and a c_2 approximately equal to 1, but the value of s would not match the optimal switch point. There may also be variance not explained by the step function (i.e., R^2 may be low).
3. If a participant failed to behave rationally but still modified his or her strategy according to task difficulty, c_2 should be larger than c_1 but fall short of the maximum step size. For example, in the detection experiment, a step with a c_1 of 0 and a c_2 of .5 (which would lead to a step size of .5) would mean that the participant always fixated the central box when the boxes were close together and fixated the side box on half of the trials with a large separation.
4. Participants might not modify their behavior at all or might do so in the wrong direction, which would lead to no step or a reversed step (indicated by s and R^2 being close to 0).

All the model fits from these analyses are given in Table S1 in the Supplemental Material and are summarized in the second row of Figure 3. We focus in this figure on the size and position of the step. The size of the step should be 1 (.5 in the memory experiment), and the position of the step is illustrated relative to its predicted location under an optimal strategy. As the figure shows, all participants in the reaching experiment were perfectly described by the step function (Category 1 in the preceding list). In the three other studies, no participants could be described as being in Categories 1 or 2. Only in the memory experiment was the step function a reasonable model for participants' behavior (9 participants had an $R^2 > .1$, as indicated by the filled circles, which puts them in Category 3). In the other two studies, except possibly in the case of a few participants, step size, direction, and location were generally not consistent with choice behavior that was modified by task difficulty (Category 4).

In Figure 4, we compare overall accuracy in Session 2 ("observed" on the x -axis) with the accuracy expected if an optimal strategy or a simple reference strategy had been adopted by each participant. The accuracy of the optimal strategy was calculated for each distance (or number of digits) as described in the Method section.

This value was calculated individually for each participant on the basis of their psychometric curves. Expected optimal accuracy as shown in Figure 4 represents the mean over all the distances or difficulties tested. For the detection, throwing, and reaching experiments, we used expected performance from the central location as a reference ("central"); in the memory experiment, the reference was performance expected from looking only at one number ("single"). We can see that for the detection, throwing, and memory experiments, participants managed to out-perform the reference, but the majority of them failed to achieve optimal performance. This difference was significant when evaluated in a paired t test comparing observed accuracy and optimal accuracy—detection experiment: $t(11) = -2.45$, $p = .017$; throwing experiment: $t(11) = -3.65$, $p = .002$; memory experiment: $t(11) = -3.98$, $p = .001$. It should be noted that this way of illustrating the magnitude of the differences between observed and optimal strategies downplays our effect; if we had excluded conditions in which the reference and optimal behaviors were identical (e.g., very short distances in the throwing experiment, in which the optimal behavior was to stand halfway between the targets), the effect would be intensified. (In the reaching task, the observed accuracy is higher than optimal because 8 out of 12 participants happened to choose the seat in front of the randomly selected reach target, which is within the variation we would expect based on chance.)

For 1 participant in the detection experiment, there was no difference between the central and optimal strategies, because his visual acuity was good enough to perform above chance even at the largest eccentricity. Likewise, across the detection, throwing, and memory experiments, several participants achieved higher accuracy in Session 2 than had been predicted from their Session 1 performance, probably because of practice effects. This suggests that our estimate of optimal accuracy was conservative.

Discussion

We observed a striking failure to make optimal decisions in three of the four experiments presented. In the fourth, in which 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 could understand the instructions and the constraints on their decision well enough to behave sensibly when the task was trivial. Why were participants seemingly unable to make this decision in the other situations? The possible explanations fall into three general categories.

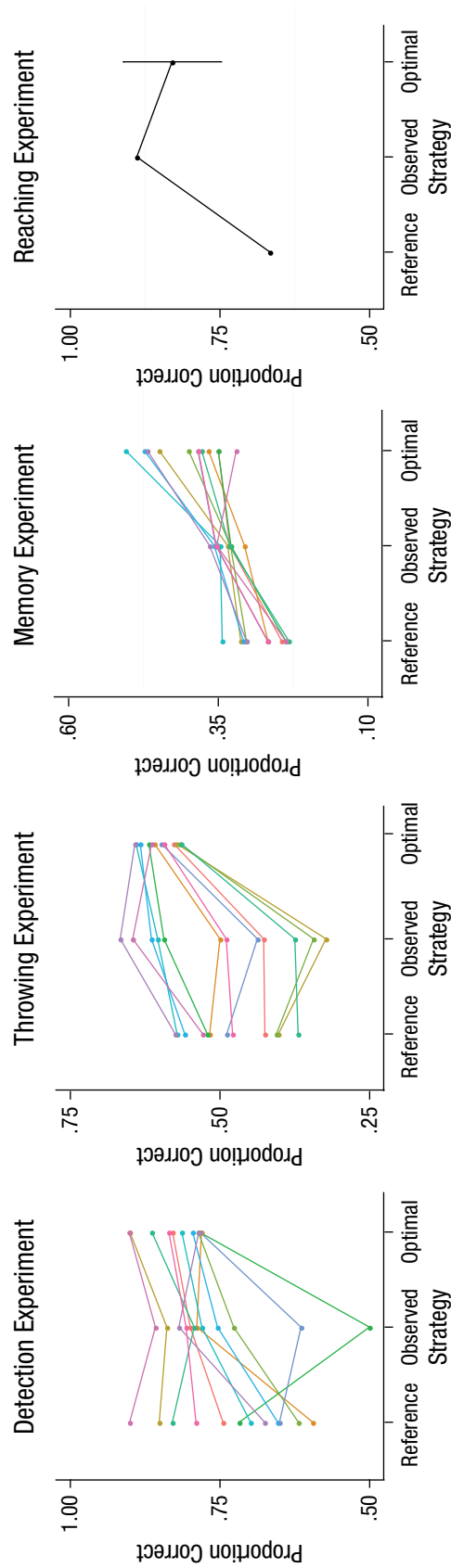


Fig. 4. Comparison of observed accuracy in Session 2 and expected accuracy if participants had followed the reference and optimal strategies. For the detection, throwing, and reaching experiments, the reference strategy was to remain at the central location; in the memory experiment, the reference strategy was to look only at one number. Performance was averaged over a range of stimuli, even when there was no expected difference in performance between reference and optimal accuracy; hence, the graphs underestimate the size of the differences between strategies. For the detection, throwing, and memory experiments, each line represents a different participant. For the reaching experiment, the black line summarizes observed and expected accuracy for all 12 participants.

First, participants may have failed to estimate their own performance accurately. The reaching task was different from the other tasks in that participants' choice depended on estimating the length of their arms, rather than on learning and remembering the limitations of their own visual acuity, throwing skill, and memory—abilities that are 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. Performance changes across these manipulations were stable and systematic (this is clear from the individual curves presented in the Supplemental Material), and people have been shown previously to be able to accurately estimate and make decisions on the basis of expected performance (e.g., Barthelme & Mamassian, 2009; Paunonen & Hong, 2010). Moreover, if people were estimating performance incorrectly or imprecisely, we would expect there to be a switch in strategies at some point as task difficulty increased, but not a consistent switch or a switch 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 2 behavior in the results of the model fit. No participants fit well into this category, which suggests a more global failure.

Second, participants may have failed to frame the decision correctly. Achieving an optimal strategy required the participants to make a logical decision (i.e., whether to invest in one option or both), followed sometimes by an entirely arbitrary decision (i.e., which option to invest in). Participants could make this pair of decisions effectively in the reaching task, which demonstrated that they were 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 distracted participants from framing the task appropriately. Alternatively, maybe trial-to-trial changes in task difficulty prevented participants from setting a single threshold at which to switch between strategies. During debriefing, we asked participants in the throwing experiment how they arrived at their decisions; some made arbitrary decisions, whereas others became focused on finding some pattern in the order of targets selected (even though they were told that selection was random). More participants fell into Category 3 (i.e., they modified their behavior with difficulty) in the memory experiment than in the detection or throwing experiment. In the other experiments, decisions were discrete; in the memory experiment, behavior progressed over a 5-s interval, so the participants may have been able to learn more effectively from the cumulative effect of their choice behavior.

A third possible explanation is that participants were prioritizing something other than accuracy in the task. Most (but not all) participants were biased toward investing in both potential targets rather than focusing on one. Performing the task in difficult circumstances might be seen either as a challenge or as an opportunity for learning, whereas selecting one option or the other takes away the challenge and puts the outcome in the hands of chance, which might be seen as a failure to be responsible for the outcome. Similarly, it might be a particularly unpleasant experience to guess incorrectly and have the nonselected option turn out to be the target; participants who happened 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, but that 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 the notion that some of them may have decided to prioritize their own interest or pleasure in the task over accuracy.

There are many ways to be suboptimal, and the fact that no single explanation 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 did participants demonstrate behavior 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 the 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 payoff diminishes (e.g., DeMiguel, Garlappi, & Uppal, 2009). We suggest that, because of the complexity involved in deciding between multiple goals in most situations, people in general fail to use 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. Only as the number of tasks decreases and the situation becomes simpler does a failure to adopt a sensible strategy become both more apparent and more detrimental.

Author Contributions

A. D. F. Clarke and A. R. Hunt developed the study concept and design. A. D. F. Clarke performed the data analysis. A. D. F. Clarke and A. R. Hunt cowrote the manuscript and approved the final version for submission.

Acknowledgments

We thank Alex Irvine, Melissa Spilioti, and a group of 3rd-year students for their help with data collection, and Peter Neri, Rob McIntosh, Sandrina Ritzmann, and James Sheils for their valuable input.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Funding

This work was supported by a James S. McDonnell Foundation Scholar Award (to A. R. Hunt).

Supplemental Material

Additional supporting information can be found at <http://pss.sagepub.com/content/by/supplemental-data>

Open Practices



All data have been made publicly available via Open Science Framework and can be accessed at <https://osf.io/btkjw>. The complete Open Practices Disclosure for this article can be found at <http://pss.sagepub.com/content/by/supplemental-data>. This article has received the badge for Open Data. More information about the Open Practices badges can be found at <https://osf.io/tvyxz/wiki/1.%20View%20the%20Badges/> and <http://pss.sagepub.com/content/25/1/3.full>.

Note

1. See results for Participant 3 in Fig. S1 in the Supplemental Material. 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.

References

- Barthelme, S., & Mamassian, P. (2009). Evaluation of objective uncertainty in the visual system. *PLoS Computational Biology*, 5(9), Article e1000504. doi:10.1371/journal.pcbi.1000504
- DeMiguel, V., Garlappi, L., & Uppal, R. (2009). Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *Review of Financial Studies*, 22, 1915–1953.
- Gardner, M. (1959, October). Mathematical games: Problems involving questions of probability and ambiguity. *Scientific American*, 201(4), 180–182.
- Hayhoe, M., & Ballard, D. (2014). Modelling task control of eye movements. *Current Biology*, 24, 622–628.
- Hunt, A., & Klein, R. (2002). Eliminating the costs of task set reconfiguration. *Memory & Cognition*, 30, 529–539.
- Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *American Psychologist*, 39, 341–350.
- Kibbe, M., & Kowler, E. (2011). Visual search for category sets: Tradeoffs between exploration and memory. *Journal of Vision*, 11(3), Article 14. doi:10.1167/11.3.14
- Knoblauch, K. (2014). psyphy: Functions for analyzing psychophysical data in R (Version 0.1-9) [Software]. Retrieved from <https://cran.r-project.org/web/packages/psyphy/index.html>
- Kording, K., & Wolpert, D. (2004). Bayesian integration in sensorimotor learning. *Nature*, 427, 244–247.
- Miller, G. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63, 81–97.
- Morvan, C., & Maloney, L. (2012). Human visual search does not maximize the post-saccadic probability of identifying targets. *PLoS Computational Biology*, 8(2), Article e1002342. doi:10.1371/journal.pcbi.1002342
- Najemnik, J., & Geisler, W. S. (2005). Optimal eye movement strategies in visual search. *Nature*, 434, 348–391.
- Oruc, I., Maloney, L., & Landy, M. (2003). Weighted linear cue combination with possibly correlated error. *Vision Research*, 43, 2451–2468.
- Paunonen, S., & Hong, R. (2010). Self-efficacy and the prediction of domain-specific cognitive abilities. *Journal of Personality*, 78, 339–360.
- R Development Core Team. (2015). R: A language and environment for statistical computing (Version 3.2.2) [Computer software]. Retrieved from <https://www.r-project.org/index.html>
- Simon, H. (1991). Bounded rationality and organizational learning. *Organizational Science*, 2, 125–134.
- Vulkan, N. (2000). An economist's perspective on probability matching. *Journal of Economic Surveys*, 14, 101–118.
- Wolpert, D., & Landy, M. (2012). Motor control is decision-making. *Current Opinion in Neurobiology*, 22, 996–1003.
- Wu, S., Delgado, M., & Maloney, L. (2009). Economic decision-making compared with an equivalent motor task. *Proceedings of the National Academy of Sciences, USA*, 106, 6088–6093.
- Zhang, H., Morvan, C., Etezzad-Heydari, L., & Maloney, L. (2012). Very slow search and reach: Failure to maximize expected gain in an eye-hand coordination task. *PLoS Computational Biology*, 8(10), Article e1002718. doi:10.1371/journal.pcbi.1002718