Variable and sub-optimal responses to a choice problem are a persistent default mode

Amelia R. Hunt1, Warren R.G. James1, Josephine Reuther2, Melissa Spilioti1 & Alasdair D.F. Clarke3

1. School of Psychology, University of Aberdeen
2. Department of Experimental Psychology, University of Göttingen
3. School of Psychology, University of Essex

Abstract

Choices can follow fixed, rational rules, but they can also rely on habits or guesses. In these experiments we ask healthy human participants to choose where to stand to throw a beanbag into one of two possible hoops (Experiment 1), and to choose where to fixate to detect a target that can appear in one of two possible locations (Experiment 2). In both cases, the optimal choice follows the same, simple logic: when targets are close together, standing at/fixating the midpoint is the best choice. When targets are far apart, standing/fixating close to one potential target will achieve better accuracy. People do not follow, or even approach, this optimal strategy, despite substantial potential benefits for performance. Our aim was to try to shift participants from sub-optimal, variable responses to following the simple fixed, rational strategy. First, we put participants into circumstances in which the solution was obvious. After participants correctly solved the problem there, we immediately presented the slightly-less-obvious context. Second, we instructed participants to make choices that followed an optimal strategy, and then removed this guidance and let them freely choose. Following both of these interventions, participants immediately returned to a variable, sub-optimal pattern of responding. We conclude that even when constructing and implementing rational decision rules is within reach, making variable and idiosyncratic responses to choice problems is a strong and persistent default mode. Borrowing concepts from classic animal learning studies, we speculate that this default may persist because choice variability can provide opportunities for reinforcement learning.

Keywords: problem-solving, decision, eye movements, variability, insight, training

Running head: Persistently sub-optimal choices

We are routinely faced with decisions about how to allocate limited resources to achieve competing goals. A rational principle to apply when faced with such decisions would be to use the resources required to achieve the goals decide how many goals to try and accomplish. Given two relatively easy goals, it makes sense to try and accomplish both. But if the combined task difficulty of achieving the two goals exceeds the available resources, focusing on accomplishing one is usually the better strategy. There are many examples of these “focus-or-divide” kinds of dilemmas in daily life, such as deciding which home improvement projects to embark on with a limited budget. You may desire to refurbish both your kitchen and your bathroom but attempting to do both at once is expensive. Having one of these projects fully completed is surely more desirable than running out of money halfway through and being left with only partially useable rooms. However, if you have a large enough budget, it is faster and more convenient to get all the work done at once. Clarke and Hunt (2016) posed this decision dilemma in a range of different contexts: throwing beanbags into hoops, detecting low contrast targets in peripheral vision, and memorizing sequences of digits. In each version of the dilemma, participants first performed the task multiple times over a range of difficulty levels to establish their skill level and limitations. For example, in the beanbag task participants were asked to throw the beanbag into hoops over a wide range of distances. When the hoop is close, this task is trivially easy, and when it is far away, the task becomes challenging. In the second phase of these experiments, participants were then given two *potential* goals, and they had to decide whether to try and accomplish both or to focus on one or the other. For example, the participants’ goal is to throw a beanbag into one of two hoops, but they are not told which of the two hoops is the target until after they choose where to stand. They could choose a spot equidistant between the two hoops, to maximize their chances of getting the beanbag in the hoops irrespective of which one is ultimately the target (divide). Alternatively, they could choose to stand next to one hoop or the other and hope it turns out to be the target (focus). Which of these strategies will lead to a higher probability of success depends on the distance between the hoops. If they are close together, the best choice is to stand at the midpoint between them. However, if the hoops are far enough apart that the expected odds of success from a central position is less than 50%, standing near one of the two hoop leads to better success rates: in the long term, they will be 100% accurate if the chosen hoop is the target, and 0% accurate at the other hoop, averaging out to 50%.

Although the logical solution to this dilemma is simple to understand and implement (in these cases at least), the results of Clarke and Hunt (2016) demonstrate that in this and two other tasks (visual detection and memorizing strings of digits), naive participants not only fail to maximize their potential accuracy on the task, most of them do not modify their focus-divide choices with manipulations of task difficulty at all. Some participants consistently stick to a “central” strategy in which they always split their resources between both goals, leading them to perform poorly as difficulty increases. The majority, however, exhibit a large amount of trial-to-trial variability, focusing on some trials and dividing on others, but not in a way that tracks with the demands of achieving the goals. This is important because it shows the limitation on performance in this task cannot be easily described or explained as one particular bias or error (e.g. to always divide). Other research using analogous methods (Morvan and Maloney, 2012; Hesse et al., 2020; James et al., 2017; 2019; 2023) produced the same failure through variability. The tentative explanation advanced by Clarke and Hunt (2016) is that in daily life, prioritization problems are usually difficult to solve because they involve more options and higher uncertainty. They suggest participants fail to recognize that the problem in this specific instance has a simple and easily implemented solution, so instead of trying to find a solution, participants default to heuristics, guesses, and trial-and-error strategies.

Recognizing that a problem has a tractable solution is a necessary first step for solving it, and this recognition could be considered an “insight”. Insight is associated with the “aha” moment that comes with solving some specific types of problems, classically known as *restructuring*, reflecting the idea that the underlying mechanism driving insight is a re-organisation that allows the same information to be represented in a different way. Weisberg (2006) argues that for a problem to be defined as an insight problem, there needs to be a discontinuity in the solution-finding process that is caused by restructuring, and this restructuring has to be necessary for a solution to be achieved. The focus-divide dilemma does not meet this definition of an insight problem, because participants can complete each decision with or without insight about the optimal strategy. That is, a failure to recognise the optimal solution does not cause an impasse in the focus-divide dilemma. Instead, participants can proceed through the experimental trials, repeatedly making choices and observing the consequences of those choices, but with a lower rate of success than they would have if they restructured the problem and implemented the optimal solution. This series of choices made in the absence of insight provide an interesting view on how people approach decision-making when they have no clear rule to follow.

The well-known “dual-systems” models (e.g. Stanovich & West, 2000) suggest cognition can take a rapid, automatic route as well as a more deliberative, logic-based route. These systems represent a very broad categorical distinction in cognitive reasoning and are normally applied to judgment and problem solving contexts. A more germane theoretical framework for the simple, repeated choices participants make in our tasks may come from classic animal learning studies, which provide a similar distinction between choices made with and without insight[[1]](#footnote-1). Maier (1940) examined the extent to which animals have the capacity to spontaneously adjust how they respond to stimuli based on knowledge. He argued for three different categories for solving problems: variability, equivalence reactions, and reasoning. *Variability* in choice responses appears to be a functional tendency that is assumed to support trial-and-error learning, and stands in contrast with repetitive, inflexible behaviour that is a hallmark of cortical damage. Krechevsky (1937) argues that varying choices allows the organism to explore “means-end readiness”, or in other words, the set of available options and their consequences within a particular problem space. These options can be refined into hypotheses, and then a solution, through a gradual process of reinforcement learning. In *Equivalence Reactions*, hypothesis learning can be transferred to other problems that share features with the original, though usually with some hesitancy. This process permits acceleration of the learning process when circumstances are in some way familiar, though if wrongly applied, it can interfere with learning. *Reasoning* differs from learning altogether, and refers to the ability to spontaneously integrate information from outside the experimental context. More generally, reasoning refers to the ability to use abstract information to guide behaviour, a concept later referred to as *teleological control* by Dickinson (1985). Dickinson similarly drew a stark distinction between *habits*, which are consistent responses elicited by the immediate environment, and *actions*, which are controlled by the animal’s knowledge about consequences. Habits will always follow from the same conditions, but actions allow an organism to change the choices it makes given what the organism knows. This dichotomy has held up well in the literature, with accruing evidence that actions and habits are subserved by distinct neural systems (e.g. O’Doherty, Cockburn & Pauli, 2017).

Viewing the series of choices in a focus-divide experiment from this learning perspective, the high degree of variability in choice observed in this task appears to be the default behaviour most participants engage in when they are not engaging in reasoning, consistent with a trial-and-error strategy. In contrast, the responses of informed participants (see the Appendix) are based on reasoning. Compared with the choices of naive participants, implementing the optimal strategy leads to choices that are more successful overall, but also less variable. With these two extremes as points of comparison, our question in this study is about whether and how we can transition people from variability to reasoning. One possibility is that equivalence reactions could provide a transition into spontaneous integration; that is, applying the correct solution in one circumstance could lead to its application in circumstances with similar properties. This generalization may prompt participants to restructure the successful choices as an abstract decision rule. On the other hand, applying a decision rule carries some risks. If applied to the wrong circumstances, the rule would not only lead to suboptimal outcomes, it would also restrict the variability of responses that is the basis for learning. Given these risks, taking a conservative approach to avoid over-generalization may be adaptive in the long run, leading to a high threshold of similarity needed for a hypothesis, or a strategy, to transfer across contexts. In two experiments, we explore this threshold by creating circumstances in which participants carry out the optimal strategy in one context, and then provide a similar (Experiment 1) and the same (Experiment 2) context to observe whether the strategy persists, or if participants return to variability. The aim was to test a potential explanation for the failure to solve the focus-divide dilemma, which is that it represents a rational tendency to apply a trial-and-error strategy in the absence of a clear decision rule. If so, when participants are in a circumstance where they can easily arrive at a clear decision rule to successfully solve the focus-divide problem, they should continue to apply that rule in future instances of the same problem. More generally, these experiments can shed light on our readiness (or lack thereof) to apply logical rules to new tasks.

**Experimental 1A: Transfer between contexts**

Clarke and Hunt (2016) demonstrated that participants fail to modify their focus-divide choices to take task difficulty into account. The same pattern of variable, sub-optimal behavior was exhibited across three diverse contexts: visual detection, throwing, and memorizing strings of digits. However, in a fourth experiment, Clarke and Hunt (2016) simplified the context to present a comparable, but trivial decision and observed uniformly optimal behavior. In this simplified version of the task participants were presented with a long table upon which three pairs of colored beanbags had been placed. The red beanbags were placed close to one another near the midpoint of the table, while the blue beanbags were placed at either end of the table. (The two green beanbags were placed at intermediate position between the blue and red bags.) Three chairs were placed at the table: one on the left end, one central and one on the right end (see Figure 1 for a photograph of this setup.) On each trial, participants were told to pick up a beanbag of one particular colour, which was selected at random. After being told which colour beanbag they would need to pick up, participants were asked to select and sit down in one of the three chairs. After they chose a chair, they were told which of the two beanbags (left or right) to pick up. Clearly, if the beanbags are within arm’s reach from the central chair, the participant should sit there. If they are too far apart to reach from the center, the participant should select a chair close to one of the beanbags, ensuring they can at least reach that one. Indeed, Clarke and Hunt (2016) found that all participants consistently used the same, optimal strategy in this version of the dilemma.

In this experiment we ask whether participants who have just performed the reaching task optimally would be primed to recognize the optimal solution in the throwing task, transferring the logic guiding their decisions from one task context to another. If so, the decisions of participants in the throwing task should be close to optimal if they just carried out optimal decisions in the reaching version of the task. The decision behaviour of this primed group was compared to a control group, and the experimenter who ran the throwing part of the session was blind to the group assignment of participants.

**Methods**

**Participants**

Thirty-two students from the University of Aberdeen were recruited either via ORSEE (Online Recruitment System for Economic Experiments) or SONA (Sona Systems Research Management System). Participants had no prior knowledge of the hypothesis of the experiment. They were randomly assigned to either the primed group (those who did the reaching task first and then the throwing task) or the control group (those who did only the throwing task), with 16 in each group. A £5 remuneration was given to the ORSEE participants, and appropriate academic credits awarded to the SONA participants. In this and all subsequent experiments, the experimental protocol was reviewed and approved by the Aberdeen Psychology Ethics Committee (except 1B below, which was approved by the Essex ethics board), and participants all gave signed, informed consent to participate.

**Justification of sample size**

If the intervention is effective, we predict a shift from a trial-and-error strategy to a strategy based on reasoning, which would produce a very large change in standing positions and accuracy. To detect this large effect would only require 2-3 participants if we followed standard conventions, but we set a higher bar and adopt the approach to power presented in James et al. (2022). Based the standing positions of participants in the beanbag throwing experiment (2) in Clarke and Hunt (2016), they simulated experiments with a sample size from 3 to 24 to detect a shift in standing positions of 0.05 of the normalised range. Estimates of the mean difference between conditions varied widely with a small N but stabilized after about 12 participants. This suggests group estimates of relative standing position change little as you add participants beyond a sample size of 12.

**Materials and Procedure**

Two experimenters conducted this study. The first experimenter carried out the reaching task (detailed below) with half of the participants. The other half remained with the first experimenter for five to ten minutes to ensure that the second experimenter would not be able to tell if the participant had carried out the reaching task based on their time of arrival. The second experimenter carried out the throwing task (described below) with both groups of participants, and was blind to which condition they were in.

*Reaching Task*: Six PVC beanbags of three colours (two red, two yellow and two blue) were placed on a long table equidistant from each other (Figure 1). At the centre of the table were the red beanbags, halfway across each side of the table were the yellow beanbags (one on each side), and at the far end were the blue beanbags (one on each end). Three chairs were positioned alongside the table: one on each end (opposite the blue beanbags) and one in the centre (opposite the red beanbags). In order to make sure participants were aware of their own reaching span, they were first asked to sit on the chair positioned at the centre of the table. With their back always touching the chair they were asked whether they could reach and touch the red, green and blue beanbags individually (the answer was yes, yes and no respectively). They were then asked to stand up and move away from the table before starting the main part of the experiment. They were told which colour beanbag they would be asked to pick up, and that they had to choose a chair to sit on, at which point they would be told which of the two beanbags of that colour was the target. This was repeated for each colour. The order of the colour-location combination was the same for each participant starting from the middle beanbag and working towards the sides. Once all three colours had been tested, they were given the beanbags and instructed to go to the other experimenter for the second part of the experiment.



Figure 1. The three images down the left side show the task setups for all versions of the focus-divide dilemma presented in this series of experiments. In the reaching task, the participant selects a chair from which to pick up one of two beanbags of a specified colour on a long table. In the throwing task, two each of red, yellow, and blue hoops were taped to the ground and participants had to choose a place to stand to throw a beanbag into one of the two hoops of a specified colour. In the detection task, participants choose whether to fixate in the center or one of the side boxes to detect a target (the small white dot) which can appear in one of the two side boxes. In all three tasks, participants do not know which (left or right) is the target until they have chosen a chair, position, or box, respectively. The task sequence for all the experiments are shown to the right. The effect of the intervention (red boxes) on position choices in the throwing (E1) and detection (E2) tasks are compared to the control conditions (blue boxes).

*Throwing task*: The experiment was an abbreviated version of the throwing task used in Clarke and Hunt (2016). An outdoor, sheltered area of concrete slabs just outside the psychology building (Figure 1) was used because the slabs (measuring 0.46 x 0.61m) were useful for marking and recording hoop placement and standing positions, respectively. Six hoops with a diameter of 0.40m were taped down in a row with three slabs’ distance between them. The red hoops were 4.60m apart (10 slabs), the yellow hoops were 8.28m apart (18 slabs) and the blue hoops were 11.96m apart (26 slabs). These hoop separations were selected on the basis of throwing performance measured in several previous experiments; participants should stand in the center for the closest distance, and next to one of the hoops for the farthest distance, to achieve optimal performance. The colour of the beanbag the participant was handed on each trial specified which pair of hoops were potential targets on that trial. Participants took the beanbag and chose a place to stand. After choosing their standing position, they were told which hoop was the target (which, as they were informed, was determined based on a pre-generated random list), and they then attempted to get the beanbag into that target hoop. The experimenter stood on the grass to the side and handed beanbags to the participant in random order of colours (pre-specified and different for each participant. Standing position and throwing success or failure were recorded. There were 15 trials for each distance (45 total).

**Analysis**

All data processing, analysis and visualisation was carried out with R (vX.XX) with the tidyverse packages (x.xx) cite(). We normalised standing position by dividing it by the distance between the hoop and the center, so values reflect a proportion of this maximum distance. Statistical models where fit with brms and cmdstanr. As our normalised standing position measure is skewed, non-negative, and zero-inflated we will use a multi-level hurdle-lognormal model to analyse the differences in decision between the near and far hoop conditions, and for the primed and control group. We used weakly informative N(0,1) priors and maximal random effect structures for all models. Full code and model specifications can be found ….

**Results**

All participants who carried out the reaching task did so optimally. That is, when they were asked to reach for one of the two red or green beanbags, they sat in the central chair. When they were told they would need to pick up one of the two blue beanbags (too far away to reach from the central chair), they chose either the left or right side chair, resulting in a 50% success rate. No variance, either within, or between participants was observed. This replicated the behaviour observed by Clarke & Hunt (2016).

***Throwing task choice***

Standing position choices in the throwing task are shown in Figure 2. The optimal solution to this problem would give standing position = 0 for the closest hoop distance and standing position = 1 for the furthest (note that standing position has been normalized over hoop separations, with 1 being the distance from the center to the hoop). This benchmark is shown as the crosses in Figure 2. We were interested in whether participants would be able to learn from their optimal performance in the reaching task, to adopt a more optimal strategy in the throwing task, but it is clear from the results from the primed group (shown in blue) that participants fail to do so. We can also see that no participant managed to execute the optimal strategy in this version of the focus-divide dilemma. This allows us to conclude that experience of solving one form of the focus-divide dilemma (the reaching task) does not lead to optimal behavior in another.

We have shown every standing position choice for every participant as a separate point in Figure 2 to depict the large amount of variation in these choices, both within and between participants. Around a third of participants exhibit a strong central tendency, sticking with the mid point between the hoops irrespective of the distance between them. The rest of the participants vary their behavior from trial to trial. On a handful of trials (7/1440) participants even chose to stand outside of the range of the hoops, leading to a normalized standing position greater than 1. This is rare but noteworthy both because it reinforces the wide variance in human choice behaviour in this task, and because affects the y axis of the plots (which would otherwise be 0 to 1). This pattern of sub-optimal and highly variable choices replicates the striking findings of Clarke & Hunt (2016).



*Figure 2. Results for each individual in Experiment 1a. Each dot is a trial and each facet is a single participant. Delta is the distance between hoops. The participants shown in red are the control group, and those in blue were primed by making optimal decisions in the reaching task before completing the throwing task. The crosses show an “absolute” optimal standard, which is standing at 0 at the smallest value of delta at 1 at the largest. The Appendix presents a more moderate “optimal” standard, showing where participants choose to stand when they are explicitly aware of the optimal solution. By either standard, no individual, in either group, chose optimal standing positions.*

Although none of the individuals in either group could be described as optimal, there might be less obvious differences between groups. We fit a multi-level hurdle-lognormal model to the data to model behaviour in the focus-divide dilemma as an initial decision about whether to stand at the midpoint, and if not, how far away from it to stand. The results of this analysis are shown in Figure 3. We can see that the control group are equally likely stand in the central position irrespective of hoop separation distance (95% HPDI of [0.4, 0.79] for the near hoops and [0.25, 0.72] for the far hoops). The primed group, while not following the optimal strategy, do appear to be more likely stand in the centre when the hoops are close to one another compared to when they are far (95% HPDI of [0.64, 0.95] and [0.07, 0.47] respectively). Interestingly, the differences in second part of the model – how far from the centre to stand – are much more mixed. Both groups show little evidence for changing their choices about where to stand based on the hoop separation distance (for the control group, [0.35, 0.63] for near hoops, [0.16, 0.44] for the far hoops, compared with [0.30, 0.73] and [0.35, 0.55] for the primed group). Based on this, we conclude that our control group had a general tendency to stand closer to the center across all conditions. This is possibly a consequence of random sampling; some people just stick to the center more than others, and we happened to have more of that kind of person in our control group. Nonetheless, we included two follow-up experiments to be confident there was no transfer effect.



*Figure 3. Summary of the fixed effects from the Bayesian hurdle-lognormal model. The coloured bands indicate 50%, 80% and 95% highest posterior continuous intervals for (left) the probability of selecting a central standing position; (right) average distance from the centre, on trials in which participants chose not to stand at the centre.*

**Performance**

Accuracy for the three hoop separations in Experiment 1a are shown in Figure 4. We can assume that participants will be accurate on average 50% of the time if they stand at one hoop in the far condition. It is clear that both the primed and control groups fall short of this standard. Because we did not include a session to measure each individual’s throwing accuracy in this experiment, we cannot estimate what their optimal accuracy would have been for the other two distances. But the far distance is the condition in which there are the largest potential gains for executing an optimal strategy, so we can expect the difference for the middle and close conditions to be smaller.

**Discussion**

The results of Experiment 1A show that priming participants to optimally solve the focus-divide dilemma in the table task did not produce the large shift towards optimality in the subsequent throwing task, which shares the same solution. However, the intervention led (some) participants in the primed group to move away from a central position when the hoops were far apart, which is consistent with being closer to optimal. As this effect is smaller than predicted but in the expected direction, and there are large individual differences in both groups, we ran two conceptual replications of this experiment in Experiments 1B and 1C to rule out the possibility that there is a robust effect of the intervention that we missed in Experiment 1 due simply to chance.

**Experiment 1B & 1C: Replications**

Experiment 1B includes both the table intervention used in Experiment 1A and a second intervention in which participants completed a logic problem based on the focus-divide dilemma. The rationale is to confirm the results of Experiment 1A and also test whether framing the problem verbally might encourage teleological control (to use the language of Dickinson, 1985) that would more readily transfer to the throwing task. These two interventions were compared to a control group who completed arithmetic problems. Experiment 1C is similar to 1A, except we asked all participants to complete the throwing task twice, once before and once after the table-task intervention. We compared the before/after change to a control group, who completed the throwing task before and after completing a sodoku puzzle. This mitigated some of the problems with comparing groups of participants on a task that elicits highly variable choice behaviour. It also could increase the likelihood that participants connect the solution they execute in the table task to the throwing context they just experienced. Alternatively, these further experiments can increase confidence in the conclusion we have tentatively drawn from Experiment 1A, that the optimal solution to the focus-divide dilemma does not readily transfer from one situation to another.

**Methods**

**Participants**

*Experiment 1B*: 24 participants (17 female, mean age 21) from the University of Essex were randomly assigned to one of three groups: a reaching task group (nine participants), an arithmetic group (eight participants), and a logic puzzle group (seven participants).

*Experiment 1C:* 20 participants (11 female, mean age, 22.8) from the University of Aberdeen were randomly assigned to either complete the reaching task or the control condition (10 in each).

**Justification of sample size**

The aim of these experiments was to ensure the (largely) null effect of the intervention observed in Experiment 1A was robust. It would take a very large number of participants to be confident there is no effect. The effect we do expect is large, however, and we are building on the sample we have already collected in Experiment 1 to be confident that the effect of the intervention continues to be small-to-absent across several minor variations of the same conditions.

**Materials and Procedure**

The throwing and reaching task components of Experiments 1B were similar to the setup of Experiment 1A, although with some minor differences due to being carried out at different sites. For Experiment 1C, they were the same as Experiment 1A unless otherwise stated.

*Experiment 1B:* In this experiment we introduced a new control condition. Rather than engage in small talk as in Experiment 1A, here participants completed ten simple arithmetic questions (i.e., 27-12 = ). We also introduced a second experimental condition in which some participants completed five logic problems, the first two of which were inspired by the same decision-making problem used in the throwing task:

*You find out that you have an exam tomorrow. You have [10 HOURS] to revise and there will be two questions where you must only answer one. You know what the two topics that the questions will be on but you are not confident about either. Note down how you will divide your time in revision.*

A second version of the same question substituted “10 minutes” for “10 hours”. These were supplemented with three additional, unrelated logic puzzles. Participants in both the logic and arithmetic condition had ten minutes to complete the questions.

Experiment 1C: Participants carried out this experiment over two sessions, approximately one week apart. The first session had two parts: accuracy measurement, and decision trials.

Accuracy measurement. Participants stood in the middle of the area and threw 12 beanbags for each of four different hoop distances {1.38, 3.22, 5.06, 6.90m} in increasing order. The beanbags were cleared out of the way after each trial. They then threw to four new distances {2.30, 4.12, 5.98, 8.74} in the opposite direction, for a total of 96 trials.

Decision trials. Participants then carried out a block of the throwing task, as detailed in experiment one, but this time with four, rather than three separations {4.6, 8.28, 11.96, 15.64}. Participants carried out six trials for each distance, in a random order.

In the second session, participants were randomly assigned to two groups. Half the participants carried out the table task, which was the same as that used in Experiment 1A. The other (control) half were given a Sudoku to complete, which participants worked on for 5 minutes. Both groups then completed the decision trials again, and then a session of accuracy measurement, so that improvements in throwing ability over the course of the experiment could be accounted for in calculating optimal standing position in the second session. Finally, participants in the Sudoku group completed the table task at the end of the experiment, to confirm that they were indeed able to successfully execute the optimal strategy (which they were).

**Analysis**

Standing position choices in the throwing sessions of the two experiments were separately analyzed in the same manner as Experiment 1A. We present the results of the two models in the same figure to ease comparisons.

**Results**

These experiments were analyzed using the same Bayesian hurdle-lognormal model as E1A. The model fit is summarized in Figure X. We can see that there is a strong, consistent tendency to stand at the central position when the hoops are close to one another, particularly in Experiment 1C. As such, the estimates for the non-central standing positions have high uncertainty, as there are very few examples of this behavior. When the hoops are far apart, people still stand in the central position around a third of the time in experiment 1B (95% HPDIs of [0.13, 0.48], [0.23, 0.65] and [0.29, 0.71] for the three groups respectively), and a tenth of the time in experiment 1C (95% HPDIs of [0.08, 0.28] for the control (sudoku) group and [0.05, 0.16] for intervention (reaching) group). In trials in which participants chose a non-central position, they move around 0.25 to 0.75 units away from the central midpoint towards one of the hoops in the far condition.

The preference to stand in the midpoint of between the two hoops in the near condition is stronger than have observed before both experiment 1A and 1B. As such, we are prevented from observing an effect of our intervention here due to ceiling effects. When the hoops are far from one another, participants are less likely to stand in the centre in all conditions. However, in experiment 1b there is no clear difference between control (maths) condition and either the reaching (95% HPDI = [-0.07, 0.46]) or logic conditions ([-0.233, 0.37]). This behaviour is more pronounced in experiment 1c, although again, there is no difference between our conditions ([-0.4, 0.20]). Furthermore, we see no difference between condition in terms of the non-central positions that are chosen in the far condition: In experiment 1b we see differences of [-0.34, 0.17] and [-17, 0.35] between the control (maths) and either the reaching or logic puzzle condition. In Experiment 1c, we observe a difference of [-0.25, 0.30]. In summary, as all 95% HPDI intervals contain, there is little evidence that our intervention causes a robust change in behaviour in the focus-divide dilemma.



*Figure 4. Summary of the fixed effects from the Bayesian hurdle-lognormal model. The coloured bands indicate 50%, 80% and 95% highest posterior continuous intervals for (left) the probability of selecting a central standing position; (right) average distance from the centre, on trials in which participants chose not to stand at the centre.*

The results of both of these experiments indicate that none of these interventions were effective in shifting participants towards more optimal throwing decisions, in line with our general interpretation of Experiment 1A: making optimal decisions in a more obvious version of the task does not lead to systematic improvements in the subsequent throwing task.

**Experiment 2: Guiding optimal eye movements**

In Experiment 1, we did not find evidence that optimal choices during the reaching task transferred readily to influence standing position choices in the throwing task. People appear to have solved the reaching task choice problem but did not generalize this behaviour to new circumstances. One plausible explanation is the change in context. In everyday circumstances, one would not expect the same logic to apply to choosing a seat at a table as to choosing where to stand to make an accurate throw. Thus, although the group who were exposed to the reaching task understood and applied the correct logic, the change in circumstances (inside to outside, reaching to throwing) was too different for them to apply the same logic in the new situation, so they returned to solving the problem using variability.

In the next experiment, we therefore tested transfer of decision strategies without a change in the task. In Experiment 1, we used the throwing task version of the focus-divide dilemma. This was selected as we felt it shared several features with the reaching task (i.e., the version of the dilemma in which participants “solve” the problem). For Experiment 2, we will use the detection task version of the focus-divide dilemma. This is the original version of the paradigm as presented by Morvan & Maloney (2012). As this variant of the dilemma is a computer based task, each participant can easily complete hundreds of trials. It is also easier to cue the optimal behaviour on each trial in a consistent manner and minimizes the possibility of experimenter-driver effects.

We provided guidance that led to the execution of optimal decisions during the first session. We then removed that guidance in the second session, and let participants freely choose. Our question is whether consistently executing optimal responses in the first session will lead to improvements in performance in the same choice task once the guidance has been removed.

To test this, participants completed a training block where they were cued to fixate a particular box. The target would only appear after they had fixated that box. The cue directed the participants’ fixations to the location that would optimize their detection accuracy, given their own visual acuity as measured in the preceding sensitivity mapping phase. After completing this session, they completed a session where they freely choose which box to fixate, just as in the original experiment. The key question is whether guidance in the form of the cues directing the eyes to the optimal location will result in more subsequent optimal decisions in the free-choice task relative to a control group who was given an equivalent amount of practice, but no guidance.

Previous research leads to mixed predictions about whether this outcome is likely. On the one hand, there is evidence that solutions to insight problems can be provoked by guiding the eyes to relevant locations (Grant and Spivey, 2003) or in a pattern that is consistent with the solution (Thomas and Lleras, 2008). This suggests a pattern of behaviour can lead participants to restructure information in a way that is consistent with that behaviour. On the other hand, our results from Experiment 1 suggest people do not readily recognize and apply an optimal solution to this decision problem. Consistent with this, Weisberg, DiCamillo and Phillips (1978) attempted to prime participants to the correct solution for an insight problem, and found that participants could only use priming to benefit problem solving when they had been explicitly told the prime was relevant. Without instruction to use the primes, there was no benefit associated with them. In our experiment, being required to repeatedly perform optimal eye movements could cause participants to persist in these eye movements when given the opportunity to freely choose. If so, this trained optimal behaviour may, or may not, be accompanied by a more abstract insight about the optimal strategy. This experiment’s methods and analysis plan were pre-registered on the Open Science Framework <https://osf.io/yan5k/>.

**Methods**

*Participants.* Twenty-four participants (mean age 21, range 18-26, 19 female, 23 right-handed) were recruited from the University of Aberdeen community. All participants were unaware of the purpose of the experiment and had normal or corrected to-normal vision. Participants were offered either course credits (for undergraduate psychology students) or a £15 reimbursement for their participation. This study was reviewed and approved by the School of Psychology ethics committee. All participants provided informed consent.

*Setup.* The stimuli and equipment were similar to those used by Clarke and Hunt (2016, Experiment 1). All stimuli were presented against a uniform grey background (50% white). The square boxes had an edge length of 1.0˚ visual angle and were lighter than the background (75% white). The target was a small, light grey dot (80% white) that appeared for 500ms either in the top (dot-up configuration) or the bottom (dot-down configuration) of one of the boxes. The “page-up” and “page-down” keys of a standard keyboard were used as response keys, matching dot-up to the page-up key, and dot-down to the page-down key. If the participant blinked or if the eyes moved more than 1° while the target was on, the trial was terminated.

The experiment was executed in Matlab R2009b on a PowerMac running Apple Macintosh Pro OS X (version 10.6.4) software using Psychtoolbox (Brainard, 1997; Pelli, 1997; Kleiner et al, 2007) and EyelinkToolbox functions (Cornelissen, Peters & Palmer, 2002). Stimuli were presented on a 24–inch Sony Trimaster EL OLED monitor with a resolution of 1080p and a refresh rate of 60 Hz. A chin rest, with forehead bar, assured a viewing distance of 54cm. The right eye was tracked using a desktop-mounted EyeLink 1000 (version 4.594) (SR Research ltd, Mississauga, Ontario, Canada) to record eye position at 1000 Hz. A 9-point calibration sequence was used. The calibration and validation process was repeated every time the participant moved away from the chinrest. The process was also triggered if the participant broke fixation during the stimulus presentation five times in a row, or ten times cumulatively for each block of trials. The experiment took place in a dimly-lit, quiet room.

*Experimental Design and Procedure.* The experiment took place over two sessions. Session 1 lasted approximately 1 hour 30 minutes, and Session 2 lasted approximately 1 hour. There were three phases (see Figure 4): an Acuity Mapping Phase’, a ‘Directed Eye Movement Phase’ and a ‘Decision Phase’. Participants were randomly assigned to one of two groups: the control group completed the ‘Acuity Mapping Phase’ and the ‘Decision Phase’ in Session 1, and the primed group completed the ‘Acuity Mapping Phase’ and the ‘Directed Eye Movement Phase’ in Session 1. In Session 2, both groups completed the same ‘Decision Phase’. All participants were instructed to respond to the target as accurately as possible. No feedback was provided. To ensure that participants understood the task and that a psychometric curve could be fitted to the performance data successfully, participants needed to score 80% or above in the closest distance in the acuity mapping phase to take part in both sessions of the experiment and to be included in data analysis. Two participants were excluded on this basis and replaced (to retain a total of 12 in each group). The three phases are described below.

*Acuity mapping phase.* During four blocks of 96 trials each (384 trials total), boxes were presented in 8 different distances: 2.7°, 3.9°, 5.2°, 6.8°, 8.4°, 10.1°, 11.4° or 12.5° between the centrally presented fixation mark and the centre of each box. Each distance was presented 12 times per block. The order of distances was randomized. All trials per distance were presented in succession. Each trial started with the presentation of a black fixation cross in the centre of the screen. Participants were asked to initiate the trial sequence by key press while fixating the cross intersection (drift check). After a stable fixation of 700 ms (within a 1˚ radius circle around the cross) two equidistant, grey boxes were presented, one on the left and one on the right of the fixation cross, with the target present in one of them. After 500 ms of stimulus presentation a blank (grey) screen appeared, indicating that a response was expected. A red screen was displayed and the trial terminated whenever participants looked more than 1˚ from fixation during the trial interval (i.e. starting from the key press at the start of the trial, ending with the onset of the response screen).

*Decision phase.* The Decision phase consisted of 4 blocks. During each block, boxes were presented in 9 different distances (measured from the centre of the screen to centre of the box). Two of the distances were constant among participants (8°, 18°). The participants’ individual switch-point (E0) was the basis for the remaining 7 distances (E0 [-3°, -2°, -1°, ±0°, +1°, +2°, +3°]). The switch-point was estimated from the acuity mapping data as described by Morvan and Maloney (2012). Each distance was presented 10 times per block (90 trials per block, total 360 trials). Trial order was randomized. Participants initiated each trial with a key press while fixating a black fixation cross to initiate a drift check. The cross was positioned 4° above the horizontal meridian with an offset of half the (to-be-presented) box distance either to the left or to the right. After 700 ms of stable fixation, three boxes were presented along the horizontal meridian. The middle box always remained centred horizontally, and the left and right boxes were equidistant from the central box on either side. The fixation cross was positioned equally often between either the centre and left box or the centre and right box. As we were not interested in the choice between the right and the left box, but between the centre and side-boxes, this equates the distance from the fixation to each of these two locations. Participants were instructed to make a saccade towards the box of their choice. Once fixation was stable inside one of the boxes for 50ms, the target configuration was presented for 500ms in either the left or the right box (never the centre). Figure 4 illustrates the target presentation of a typical trial of the decision phase.



Figure 4. Left: Example illustration of a trial view from the decision phase of Experiment 2. The target could appear in the left or right box. Participants responded to indicate whether the dot appeared in the upper or lower portion of the box. This example illustrates the target in the upper portion of the left box. Right: A schematic to illustrate the organization of the experimental groups and sessions.

*Directed eye movement phase.* This phase was similar to the decision phase, with the difference that participants were not given a choice over which of the three presented boxes they would fixate. Instead, the three boxes were presented alongside the fixation cross at the beginning of the trial, and after 700ms of stable fixation of the fixation cross, one of the boxes ‘blinked’ (disappeared for 400 ms and then reappeared) and participants were instructed to execute a saccade to this box and remain fixated there until the target appeared. This phase consisted of 4 blocks of 10 trials for each of the 9 distances used (90 trials per block, total 360 trials). The same distances as in the decision phase were used. Which of the boxes blinked was dependent on the box distance and the participants’ individual switch-point. For distances equal or smaller than the switch-point distance, the central box blinked. For all other distances, randomly either the left or the right side-box blinked.

**Analysis**

The choice of which box to fixate was analysed using a similar approach to the one used in Experiment 1. The main difference is that whereas in Experiment 1 participants were free to stand at any location of their choosing, their behavior in Experiment 2 was constrained so that they could only fixate one of the three boxes. This allows us to simplify our analysis and use a binomial glm to describe participants’ rather than a hurdle-lognormal model. Another difference between experiments is that in the previous experiment we treated the separation distance between the two targets as categorical, here we have more distances and hence treat this variable as continuous. As before, the data is fit using a multi-level model. Note: we use less-conservative prior for the slopes as we expect a near vertical slope when participants are correctly following the ideal decision strategy.

**Results**

The Bayesian model’s posterior is summarised and compared against the empirical data in Figure 5 (*left*). For block 1 we can see that all participants managed to follow the instructions and there is a very large difference between the participants who were provided with *instruction* and those who weren’t. Furthermore, the estimates for the *no instruction* condition is extremely wide, reflecting the high degree of variance between participants when they are left to make their own decisions in our paradigm. If we look at the model fits for block 2, we can see that both groups of participants exhibit a high degree of variance. Crucially, there is no evidence that the participant who were presented with instructions in block 1 are still implementing the optimal strategy in block 2. We can examine the size of the difference between the two groups in block 2 by examine the model’s posterior density function as shown in Figure 5 (*middle*). The 95% HPDIs for instructed group is [-0.04, 0.10], while the no instruction group gives [-0.12, 0.03], and the posterior difference is [-0.03, 0.18]. So while there may be a small difference in participants’ sensitivity to the increasing separation, there is not enough… unsure how we want to phrase here. We can summarise these distributions in an alternative way and say that the probability that the instruction group had greater sensitivity is 0.93. Important, whether there is a difference or not is moot as we can firmly conclude that participants who were instructed in block one are unable to implement this strategy in block 2.

The main research question in Experiment 2 was whether optimal eye movement choices can be trained. To address this question, we calculated a difference score between each participant’s expected target discrimination accuracy (under an optimal model) and their actual observed accuracy, and compared the difference scores for the decision phase. This gives a measure of how close to optimal each participant is. During the first session, participants who were guided to make optimal choices were, of course, optimal (i.e. had a difference score around 0). This value was lower for the instructed participants in Session 1 than for the controls who received no guidance (). In other words, being guided to optimal choices about which box to fixate does indeed significantly improve detection accuracy relative to when these choices are freely made (see the Verification Phase at the end of the experiment reported by Morvan and Maloney, 2012, for similar evidence that optimal fixations significantly improve detection performance). For the control group, accuracy is nominally closer to optimal accuracy in Session 2 compared to Session 1, presumably due to practice, but this is not significant ().

To address the main question of whether training improves performance, we can look at Session 2 performance. The hypothesis was that participants who initially were guided to make eye movements that maximize their chance of correct target discrimination would continue to perform optimally when freely choosing where to fixate in Session 2, and should therefore differ from optimal less than the control group. However, a t-test comparing how much the two groups differed from optimal in Session 2 was not significant ). A full Bayesian analysis of the accuracy data is reported in the supplementary material.

The distribution of accuracy relative to optimal accuracy for each participant can be seen in the scatterplots in Figure 5. In Session 1, the primed group accuracy even slightly exceeds optimal accuracy due to chance, and possibly also because the estimate of optimal is based on the sensitivity mapping phase, which came first in Session 1; accuracy may have improved with practice, leading to better detection in the decision phase. The control group falls below the line. In the second session, where both groups freely choose which box to fixate, there is no difference between the groups. Figure 6 shows the choice behaviour of the two groups in Session 2 (see Supplementary Material for a more detailed figure, showing each participant relative to their own switch point for both sessions). There is no clear difference between the groups in terms of the distribution of choices between the center and side boxes. None of the 24 participants in this experiment could be described as optimal when allowed to freely choose which box to fixate, even after 360 trials of being guided to fixate using an optimal strategy.



Figure 5. (*left*): the coloured ribbons indicate the 90% HPDI for the fixed effect component of our generlalised linear model. The lines provide a summary of the data: each line represents a different participant and indicates the proportion of time they fixated the central square for different separation distances. Note that in block 1, the model estimate for the instruction condition is very narrow and is hard to see under the individual lines. (*middle*): The model’s posterior distribution of the slopes in block 2 of the experiment. (*right*): Boxplots illustrating the difference in overall accuracy between conditions and blocks.

**General Discussion**

Add in a longer version of this somewhere: “Suggesting the limitation comes from a failure to recognize the solution rather than a failure to execute it effectively, Appendix A presents data showing that non-naive participants can implement the optimal solution easily.”

The experiments clearly demonstrate that guiding naive participants to perform optimally did not help them to make better decisions when the guidance was removed. In these experiments, optimal performance is straightforward to attain when the participant has an explicit understanding of the domain-general principle that should guide the decision (see Appendix A). Even though this principle is intuitive once known, people do not appear to spontaneously construct or exploit this principle, nor do they apply it more broadly even when they seem to implement the principle in another context (in Experiment 1) or have experienced the benefits of using it (in Experiment 2). Instead, participants appear to default to relatively variable, sub-optimal patterns of responding.

The experiments were designed to reveal how readily participants recognize and adopt rational decision rules. Based on the results, we conclude a variable trial-and-error approach to making choices is a stable default. While the majority of participants had some degree of bias towards one choice option or another, there was also a large amount of variability from trial to trial under the same conditions. A small minority of participants responded the same way on nearly every trial; for example, 4/32 participants in Experiment 1A always stood in the center to throw, and 2 out of 24 participants in Experiment 2 always chose to fixate the center box. This could reflect a fixed decision rule, but not one which maximises success in the task. Effort minimization (Irons and Leber, 2016), stress (Shors and Dryver, 1992), and distraction (e.g. Wolford, Newman, Miller & Wig, 2004) could contribute to rigid or stereotyped behaviour. The rigid sub-optimal behaviour of a minority of participants in our sample is also consistent with the concept of *habits* as defined by Dickinson (1985): In rats, a response that has been consistently rewarded can, over time, lead to perseverative responses after the behaviour ceases to be reinforced. An intriguing study by DeWit, Kindt, Knot et al (2018), however, failed to induce habits in humans under similar protocols to those used previously in rats. In other words, people were highly sensitive to changes in the contingencies between responses and their reward or punishment, even after over-training with the original contingencies. This finding, together with the far more common tendency towards variability observed in most of our participants, suggests intriguing relationships between variability in behaviour and flexibility in responding to environmental change that remain to be explored in humans.

The results of Experiment 2 may initially seem to be at odds with previous research finding that eye guidance can guide thought (Grant and Spivey, 2003; Thomas and Lleras, 2009). Similar conclusions have been drawn from studies showing that gesturing can influence strategies in problem-solving (e.g. Alibali, Spencer, Knox and Kita, 2011). These results all suggest that motor behaviour can translate to abstract thought to promote solution-finding in difficult problems. Based on these results, one might have expected that performing responses consistent with the optimal solution would have prompted at least some of our participants to recognize the solution and apply it in future situations, but this did not occur. As noted in the introduction, contemporary research on problem-solving has tended to focus on relatively complex, abstract problems with a single correct solution that is easily recognized when achieved. These kinds of problems usually require reason to be solved; trial-and-error learning of the solution is generally not an efficient approach. Our results do not contradict the notion that action can guide people towards solving these more complex and abstract kinds of problems, because the problem we have presented in these experiments, like most of the problems we encounter in our daily lives, does not lead to an impasse if it is not “solved”, but instead leads to less efficient behaviour if it is not solved optimally. Indeed, the solutions to the routine dilemmas of a typical day (what to eat, what to wear, what to say) are highly context-dependent and difficult to predict, and most people would agree it is best not to “over-think” these. For example, we move our eyes around three times each second, and each of these movements can be thought of as a choice -- a resolution to a mini-problem of where the most useful and interesting visual information is coming from at the moment. Solving this problem optimally, although possible, requires complex computations even in a highly simplified and predictable environment (Najemnik and Geisler, 2008). In a complex and unpredictable environment, a more efficient approach to these kinds of “small” problems may be to solve them with variability (Krechevsky, 1937), allowing the constraints of the immediate environment to shape the set of viable choices and randomly varying within that set to allow for flexibility and learning to occur. Consistent with this notion, a stochastic model of fixation selection during visual search, which selects fixations at random from a population of common saccade vectors, describes human search behaviour reasonably well (Clarke, Green, Chantler & Hunt, 2016; Clarke, Stainer, Tatler & Hunt, 2017). A similar process of random selecting from a population of possible responses may guide other forms of decision, preventing stereotyped behaviour while avoiding over-thinking of minor choices.

Applying a consistent decision rule based on knowledge will restrict choice variability. Restricting this variability may stunt the potential for learning to shape behaviour in a way that flexibly adapts to dynamic and unpredictable environments. If applying reason comes at a cost to potential learning, a conservative use of reason might be warranted, especially in solving the simple, repeated problems of daily life. We are using Maier’s (1940) definitions of both variability and learning in this context, where variability is an approach to solving a problem (such as a rat choosing a direction in a maze) that allows the animal to define and refine the array of choice options and their consequences, and learning is the tendency to repeat choices that lead to positive outcomes. In our experiments, the variable but sub-optimal choices of our participants over the series of trials is consistent with a trial-and-error approach, and the failure to spontaneously transfer or maintain an optimal strategy suggests this trial-and-error approach is a strong and persistent default mode. Although this explanation for our results is speculative, a similar tradeoff between learning and logic has been previously proposed in the context of probability matching. When asked to repeatedly guess which of two events will occur on each trial, participants tend to be sensitive to the probability of these events. That is, an event that occurs with a probability of .8 will be guessed more often than one with a probability of .2, In fact, participants tend to match the ratio of their guesses to the probability of the events. Of course, this is a sub-optimal strategy for maximizing guessing accuracy, resulting in an overall success rate (in this example) of .68, compared to the .8 participants could attain if they guessed the event with the higher probability on every trial. This tendency towards probability matching has been interpreted as a cognitive limitation (e.g. West and Stanovich, 2003). However, Gaissmaier and Schooler (2008) noted that maximizing by only ever guessing the more likely event can become sub-optimal if the sequence is not actually random, that is, if there are patterns in the sequence that can be discovered and exploited. Indeed, participants who probability-matched (in a random sequence block) were more likely to detect and exploit a pattern in the sequence when it was introduced in a separate block. This is consistent with the notion that applying a constant rule (e.g., always guessing the more likely event) can optimize accuracy as long as the critical conditions persist (i.e. that the sequence stays random, and the likely event continues to be more likely). In a dynamic natural environment, random sequences are rare, and patterns can emerge and change. The cost of applying a consistent response rule is that the behaviours that could be reinforced by the current conditions will not be executed, and thus will never be reinforced and repeated.

The variable responses we observed in the choice behaviour of the participants in our experiments could in part be inherent to the participant; under conditions of choice uncertainty, there may be an element of random selection between options to avoid becoming fixed in a particular stimulus-response pattern and missing opportunities for learning, as described above. Some variability may also come from trial-to-trial variations in the immediate context. Summerfield and Tsetsos (2015) argue that inefficient economic decisions can be accounted for in the context of *efficient coding:* neural representations of different choice options emphasize the features that are most diagnostic in the immediate context. In other words, the relative appeal of different choices will vary according to the local context, as a result of a mechanism that facilitates their differentiation across a wide range of potential circumstances. Being sensitive to idiosyncratic changes in the environment can lead to behaviour that is sub-optimal in terms of maximizing potential gains and minimizing risk and energy expenditure in an environment that, over the long term, is stable. Many so-called “sub-optimal” choices, like probability matching, could result from over-sensitivity to short-term, idiosyncratic variations in local context. Similarly, opting for larger reward later over a smaller one now is only optimal if the environment is stable. However, environments are rarely stable, and incorrectly assuming they are could lead to disastrous outcomes, such as giving up a short-term sure reward in favour of a longer-term uncertain reward that never comes to fruition. Kolling, Wittman and Rushworth (2014), for example, showed that participants can strategically shift from being risk-averse to risk-prone depending on how large the payoffs are and how many chances they have remaining to achieve a goal. In this case, as in probability matching, being sensitive to changes in the local context yields better outcomes than a consistently applied principle that incorrectly assumes a stable world.

In conclusion, we have found a simple and intuitive decision rule is not recognized or adopted by most participants. Choices are governed by a complex set of individual and contextual factors and their interaction, even for seemingly simple decisions such as a rat deciding which way to turn at a junction point in a maze (e.g. Tolman, 1938). Nonetheless, based on the overall pattern of behaviour of the naive participants who completed the choice task in the current set of experiments, we can conclude that few, if any, settled on an optimal strategy even after being guided to make optimal choices in a similar, or the same, context. After training participants to make optimal choices, the majority immediately returned to making variable and idiosyncratic decisions. Given how strong and persistent this behaviour is, the benefits of this variability may, in many circumstances, outweigh the costs of finding and applying consistent rules.

### References

Alibali, M.W., Spencer, R.C., Knox, L. and Kita, S. (2011). Spontaneous gestures influence strategy choices in problem solving. *Psychological Science,* 22, 1138-1144.

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J,. Li, P., & Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of statistical software*, *76*(1).

Clarke, A.D.F., Greene, P., Chantler, M.J. & Hunt, A.R. (2016). Human search for a target on a textured background is consistent with a stochastic model. *Journal of Vision,* 16(7):4

Clarke, A.D.F., Stainer, M., Tatler, B. & Hunt, A.R. (2017). The saccadic flow baseline: Accounting for image-independent biases in saccadic behaviour. *Journal of Vision,* 17(11):12.

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

de Wit, S., Kindt, M., Knot, S. L., Verhoeven, A. A., Robbins, T. W., Gasull-Camos, J., ... & Gillan, C. M. (2018). Shifting the balance between goals and habits: Five failures in experimental habit induction. *Journal of Experimental Psychology: General*, *147*(7), 1043.

Dewsbury, D.A. (2000). Comparative cognition in the 1930’s. *Psychonomic Bulletin and Review,* 7, 267-283.

Dickinson, A. (1985). Actions and habits: the development of behavioural autonomy. *Philosophical Transactions of the Royal Society B,* 308, 67-78.

Gaissmeier, W. & Schooler, L.J. (2008). The smart potential behind probability matching. *Cognition,* 109, 416-422.

Grant, E.R. & Spivey, M.J. (2003). Eye movements and problem solving: guiding attention guides thought. *Psychological Science*, 14, 462-466.

Hesse, C., Kangur, K. & Hunt, A.R. (2020). Decision-making in slow and rapid reaching: Sacrificing success to minimize effort. *Cognition, 205,* 104226.

Irons, J. L., & Leber, A. B. (2016). Choosing attentional control settings in a dynamically changing environment. *Attention, Perception, & Psychophysics*, *78*, 2031-2048.

James, W., Clarke, A.D.F. & Hunt, AR. (2017). Failure to use probability of success when deciding whether to pursue one goal or two. in *Proceedings of the 39th annual meeting of the Cognitive Science Society.* Cogsci, pp. 2279-2283.

James, W.R.G., Hunt, A.R. & Clarke, A.D.F. (2023). Six of one, half dozen of the other: Suboptimal prioritizing for equal and unequal alternatives*.* *Memory and Cognition,* 51, 486-503.

James, W.R.G., Reuther, J., Angus, E., Clarke, A.D.F. & Hunt, A.R. (2019). Inefficient eye movements: Gamification improves task execution, but not fixation strategy. *Vision*, 3(3), 48.

Kolling, N., Wittman, M. and Rushworth, M.F.S. (2014). Multiple neural mechanisms of decision making and their competition under changing risk pressure. *Neuron,* 81, 1190-1202.

Krechevsky, I. (1937). Brain mechanisms and variability I: Variability within a means-end-readiness. *Journal of Comparative Psychology*, 23, 121-138.

Krechevsky, I. (1937). Brain mechanisms and variability II: Variability where no learning is involved. *Journal of Comparative Psychology*, 23, 139-163.

Maier, N.R.F. (1940). The behavior mechanisms concerned with problem solving. *Psychological Review,* 47, 43-58.

Metcalfe, J. and Wiebe, D. (1987). Intuition in insight and noninsight problem solving. *Memory and Cognition,* 15, 238-246.

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

Najemnik, J., & Geisler, W. S. (2008). Eye movement statistics in humans are consistent with an optimal search strategy. Journal of Vision, 8 (3):4.

Nowakowska, A., Clarke, A.D.F. & Hunt, A.R. (2017). Human visual search is far from ideal. *Proceedings of the Royal Society B: Biological Science,* 284, 2016.2676.

O’Doherty, J.P., Cockburn, J. & Pauli, W.M. (2017). Learning, reward, and decision making. *Annual Review of Psychology,* 68, 73-100.

Shors, T.J. & Dryver, E. (1992). Stress impedes exploration and the acquisition of spatial information in the right-arm radial maze. *Psychobiology*, 20, 247-253.

Stanovich, K. E., & West, R. F. (2000). Advancing the rationality debate. *Behavioral and brain sciences*, *23*(5), 701-717.

Summerfield, C. & Tsetsos, K. (2015). Do humans make good decisions? *Trends in Cognitive Sciences,* 19, 27-34.

Tatler, B.W. (2007). The central fixation bias in scene viewing: Selecting an optimal viewing position independently of motor biases and image feature distributions. *Journal of Vision,* 7(14):4.

Thomas, L. & Lleras, A. (2009). Swinging into thought: directed movement guides insight problem solving. *Psychonomic Bulletin and Review,* 16, 719-723.

Tolman, E.C. (1938). The determiners of behavior at a choice point. *Psychological Review,* 45, 1-41.

Weisberg, R., DiCamillo, M. & Phillips, D. (1978). Transferring old associations to new situations: a non-automatic process. *Journal of Verbal Learning and Verbal Behaviour* 17, 219-228.

Weisberg, R.W. (2006). *Creativity: Understanding innovation in problem solving, science, invention, and the arts.* John Wiley & Sons.

West, R.F. & Stanovich, K.E. (2003). Is probability matching smart? Associations between probabilistic choices and cognitive ability. *Memory & Cognition,* 31, 243-251.

Wolford, G., Newman, S., Miller, M. B., & Wig, G. (2004). Searching for patterns in random sequences. *Canadian Journal of Experimental Psychology,* 58, 221–228.

Acknowledgements

The authors would like to thank Jessie Grimsey & Michelle Tran for collecting data for Experiment 1b; Anca Popescu for collecting data for Experiment 1c; and Eleanor Mackay for collecting the data for Experiment 2. We would also like to thank Justin Claydon, Alex Anwyl-Irvine, Ellen Angus, and Anna Nowakowska for their contributions. This research was made possible through support from the James S. McDonnell Foundation (Scholar Award to ARH).

**Appendix: Optimal Choices in Non-Naive Participants**

Non-naive participants completed replications of both the throwing and detection versions of the experiments reported in Clarke and Hunt (2016). In the throwing task, participants made a choice about where to stand to throw a beanbag into one of two possible hoops. They are only told which hoop is their target after they have made their choice. The detection task is similar in structure, except participants make a choice about where to look to detect a probe that can appear in one of two boxes. In both tasks the distance between the possible target locations (hoops/boxes) is varied, and the logic is the same: stand/look in a central location, equidistant between the two potential target locations when they are close together, and stand/look close to one potential target location when they are too far apart to throw to/see reliably from a central position. The results demonstrate the optimal strategy can be easily and effectively implemented when it is explicitly known, as a basis for comparison with Experiments 1 and 2.

**Methods**

**Throwing Experiment**

*Participants*. Four participants completed the throwing task. Three are authors on this paper (AH, WJ, JR) and the final participant was a member of the lab, familiar with the paradigm. All participants were familiar with the optimal strategy but had not completed the experiment previously.

*Materials and Procedure*. The procedure was similar to that used in Clarke and Hunt (2016, Experiment 2), and the throwing task in Experiment 1, with the following modifications. A first phase was conducted, in which flat hoops with a diameter of 0.40m were placed at a range of distances away (1.38m, 3.22m, 4.14m, 5.06m, 6.9m & 9.2m). Participants threw 12 bean bags into a hoop at each distance in each of two directions and throwing performance (out of a maximum of 24) was used to determine the point at which each participant’s accuracy dropped below 50%.

In the second phase, participants were again asked to throw bean-bags into hoops, but this time there were two potential targets for each throw and participants needed to choose where to stand before finding out which of the hoops was the actual target. In the original experiment (Clarke and Hunt, 2016), the distance between the hoops in the second phase was based around the participants’ performance in the first phase. In the current version, six distances were randomly selected from a range of distances based on the performance of participants in previous experiments (0.46m to 11.5m). The random selection of distances within this range ensured that participants would be unable to rely on their prior knowledge of the structure of previous experiments in deciding whether they should stand in the middle or next to one hoop. Instead they had to base their decision on knowledge about their own ability. To begin the session, three pairs of hoops matched in color were placed at each of three separations (red was the closest separation, yellow the middle, blue the farthest). After 45 choice trials, the hoops were shifted to three new separations for a further 45 trials. The colour of the bean-bag that the participant drew at random from a bag before each throw determined which pair of hoops were potential targets on that trial (e.g., if a red beanbag was drawn, one of the two red hoops would be the target on that trial). Participants then took the beanbag and chose a place to stand. After they chose their standing position, they were told which hoop was the target, and they then attempted to get the beanbag into that target hoop. The bean bags would only be placed back into the bag once all nine had been thrown, to ensure that participants made an equal number of decisions for each distance. There were 15 trials for each distance (90 in total). Which of the two hoops would be designated as the target on each trial was determined by a pre-generated random sequence. On each trial, the experimenter recorded the color of the beanbag, the standing position (based on numbers chalked on the wall over each row of paving slabs), and throwing accuracy (0 or 1).

**Detection Experiment**

*Participants*. Five participants (4 female, all right-handed) completed the detection task. Four are authors on this paper (AH, WJ, EM, JR) and the fifth participant was a lab member, familiar with the paradigm. All participants were aware of the optimal strategy, had normal or corrected to normal vision, and like in the throwing task (Experiment 1a), were unaware of their personal switch-point. All participants provided informed consent.

*Materials and Procedure.* The participants completed a procedure that matched Session 1 of the control group as described in Experiment 2 above. In other words, they completed the Acuity Mapping Phase (384 trials) followed by the Decision Phase (360 trials). The distances between boxes, like in the throwing experiment, were randomly selected from the range of distances used in previous experiments to ensure the non-naiive participants would have to rely on their own visual acuity to decide where to fixate, rather than their knowledge of how the experiment had been set up for naive participants.

**Results**

***Choices.*** We first visually compared each participant’s choice behaviour to an individualized estimate of their optimal strategy (that is, choices that would have achieved optimal performance). The optimal choices for each participant in each experiment are determined based on performance in the first phase of the experiments (see supplementary information for performance curves). For the throwing experiment, the participant should choose to stand in the center for distances where accuracy from a central position is expected to be greater than 50%, and should stand near one hoop or the other for distances where accuracy from center is less than 50%. This is shown in Figure A1, with a dark blue line showing optimal performance and black circles showing actual standing position. It is clear from this figure that these expert participants made standing position choices that were close to optimal.

Similarly, for the detection experiment, the participant should choose to look at the center box when the separation between boxes is small enough that expected accuracy from the center is >75%. For expected accuracy from the center of less than 75%, participants should instead choose the left or right box, because if the target appears here they will be 100% correct and if it appears in the other box they will be 50% correct, giving an expected overall accuracy of 75%. This is shown as the blue line in Figure A1. Again, it is clear that expert participants’ choices of where to fixate in this task (the black dots) were close to optimal.





Figure A1. Results for each expert participant in the throwing experiment (1a) on the top row. Each dot is light grey and represents standing position on a single trial; the darker dots represent up to 15 overlaid trials. The dark blue line shows the optimal standing position for each participant. The standing position has been normalized to the distance from the center to the hoop. This means deviations from the line for small values of delta represent relatively small changes in standing position. The second row shows the results for each expert participant in the detection experiment (1b). Each dot represents the proportion of fixations made towards one of the side boxes at each value of delta. The dark blue line represents the optimal fixation strategy for each participant.

***Performance.*** For each participant in each experiment, we calculated the accuracy they would have achieved under an optimal strategy (i.e., had their choices perfectly followed the blue lines in Figure A1). The size of the difference between actual proportion correct and each participant’s optimal performance ranges from underperforming by 0.089 to out-performing by 0.081. Overall, the mean difference from optimal was -0.017 for the throwing task, and 0.036 for the detection task, suggesting participants overall were close to optimal in their performance.



Figure A2. Comparison of actual proportion correct to optimal (an estimate of the proportion correct the participant would have achieved given optimal choices) for the throwing and detection experiments. Each line represents an expert participant. Performance is close to optimal.

Although participants’ choices were broadly consistent with a near-optimal strategy, it is clear from the data that not all participants were accurate in switching at precisely the right point. These small inconsistencies had a trivial effect on accuracy, however, and all participants achieved a level of accuracy in the task that nearly equaled the accuracy expected under an optimal strategy. The results demonstrate that implementing the optimal strategy when it is known is straightforward.

1. Although the 1930’s is typically thought of as having been dominated by behaviourism, there existed a minority of researchers who were investigating and developing theories of animal problem-solving. Dewsbury (2000) suggests this research has been forgotten because it was marginalized by the behaviorists at the time, and subsequently swept aside by the cognitive revolution in the 1950’s, which tended to somewhat indiscriminately characterize all the research of this era as behaviourist. [↑](#footnote-ref-1)