Six of one, half dozen of the other: Suboptimal prioritizing for equal and unequal alternatives

Warren James, Amelia R. Hunt, and Alasdair D.F. Clarke

Abstract

When presented with two difficult tasks and limited resources, it is better to focus on one task and complete it successfully than to divide your efforts and fail on both. Although this logic seems obvious, people demonstrate a surprising failure to apply it when faced with prioritizing dilemmas. In previous research, the choice about which task to prioritise was arbitrary, because both tasks were equally difficult and had the same reward for success. In a series of three experiments, we investigated whether the equivalence of two tasks contributes to suboptimal decisions about how to prioritize them. First, we made one task more difficult than the other. Second, we compared conditions in which both tasks had to be attempted to conditions in which participants had to select one. Third, participants chose whether to place an equal or unequal reward value onto the two tasks. Each of these experiments removed or manipulated the arbitrary nature of the decision between options, with the goal of facilitating optimal decisions about whether to focus effort on one goal or divide effort over two. None of these manipulations caused participants to uniformly adopt a more optimal strategy, with the exception of trials where participants voluntarily placed more reward on one task over the other. In these, choices were modified more effectively with task difficulty than in previous experiments. However, participants were more likely to choose to distribute rewards equally than unequally. The results demonstrate that equal rewards across two tasks are preferred over unequal, even though this reward equivalence leads to poorer task strategies and smaller gains.

A hungry donkey placed equidistant from two bales of hay will starve to death; at least, this happens in the "Buridan's Ass" paradox (as described in Lamport 2012[1]). Previous work (Clarke & Hunt, 2016) explored a variant of this decision problem, which can be illustrated as follows:

Imagine a hungry donkey in a herd of other hungry donkeys, and two empty troughs. The donkey does not know which trough the farmer will deposit the hay into, but once it has been dropped off, she will want to reach the trough as quickly as possible before the hay is devoured by the others. Where should the donkey wait for the farmer? The best choice depends on the distance between the troughs; if they are close together, she can stand midway between them, and reach either trough fast enough to get at least some hay. If the troughs are far apart, though, if she waits in the middle then most of the hay will be gone before she gets there. So a clever donkey would stand close to one trough and hope the farmer chooses to fill that one; at least now she has a 50% chance of getting lots of hay.

This *focus-or-divide* dilemma is an example of a resource-allocation problem, and is a simplified version of a problem we face routinely in daily life: for example, in deciding which projects to try and accomplish in a given timeframe, or deciding where to wait for a person when our rendezvous point was vaguely defined. In experiments that present focus-or-divide dilemmas, people consistently demonstrate an inability to apply the simple solution described above (Clarke & Hunt, 2016; James, Clarke, & Hunt, 2017; James, Reuther, Angus, Clarke & Hunt, 2019; Morvan & Maloney, 2012). In fact, the vast majority of people, across a range of task contexts, do not alter their decisions about whether to pursue one goal or two with the difficulty of the goals. By not following this relatively simple logic, the rate of success can fall far below that which would be expected had the optimal strategy been implemented.

We can formalise this dilemma with mathematics rather than donkeys. Let E represent the participant's expected accuracy under some behaviour ϕ. In the scenario outlined above, we have:

$$E(\phi) = P_c(A)P_s(A|\phi) + P_c(B)P_s(B\,|-\phi) \qquad (1)$$

where A and B are two possible tasks, one of which will be selected with probabilities $P_c(A) = 1 - P_c(B)$. We do not know ahead of time which task will be required, but we can choose whether to prioritize one goal over the other (i.e., $\phi = +/- 1$) or equally prepare for both possibilities ($\phi = 0$). Once we have set ϕ, the goal (A or B) is selected, and our chance of succeeding is given by

[1] The origin of the Buridan's Ass paradox is unclear, but this paper offers a modern explanation of "Buridan's Principle".

$P_s(A|\phi)$ or $P_s(B|\phi)$. In some versions of this task, ϕ = -1, 0 or 1 (Morvan & Maloney, 2012), while in others, intermediate levels of prioritization are allowed (Clarke & Hunt 2016).

Previous work on this choice paradigm has, to our knowledge, been restricted to the case where A and B are equal and symmetric goals: both *A* and *B* are equally likely to be selected (i.e., $P_c(A) = P_c(B) = 0.5$), while the difficulty of the two tasks, $P_s(A) = P_s(B)$, has been systematically varied with respect to a parameter Δ:

$$E(\phi, \Delta) = P_c(A) \times P_s(A|\phi, \Delta) + P_c(B) \times P_s(B\,|-\phi, \Delta) \qquad (2)$$

In our donkey example, Δ represents the distance between the two troughs. When Δ is small, such that $P_s(A|\phi=0, \Delta) > 0.5$, setting $\phi = 0$ maximises our expected accuracy[2]. If we increase the difficulty so that $P_s(A|\phi=0, \Delta) < 0.5$, preparing equally for both potential tasks is no longer optimal, and instead we should opt to gamble on either task A or task B being selected, i.e, ϕ = 1 or ϕ = -1. Figure 1 shows a schematic representation of the task. In more general terms, the solution to this decision dilemma is to focus on a single goal when the demands of achieving multiple goals exceed the available resources. When both goals are achievable given the constraints, one can focus on achieving both. Throughout this manuscript, we will refer to the success rate that could be expected under the optimal strategy as “optimal accuracy.”

---

[2] This requires a few assumptions about the nature of $f(\Delta) = P_s(A|\phi=0, \Delta)$. For example, when Δ is small, the centre strategy is optimal if $f(\Delta - \phi) + f(\Delta + \phi) < 2f(\Delta)$. When Δ is large, we require $f(0) + f(2\Delta) < f(\phi) + f(2\Delta - \phi)$. A formal description of the family of functions for which this holds is outside the scope of this paper. We have verified that these criteria hold for the vast majority of empirical psychometric curves collected during our experiments.

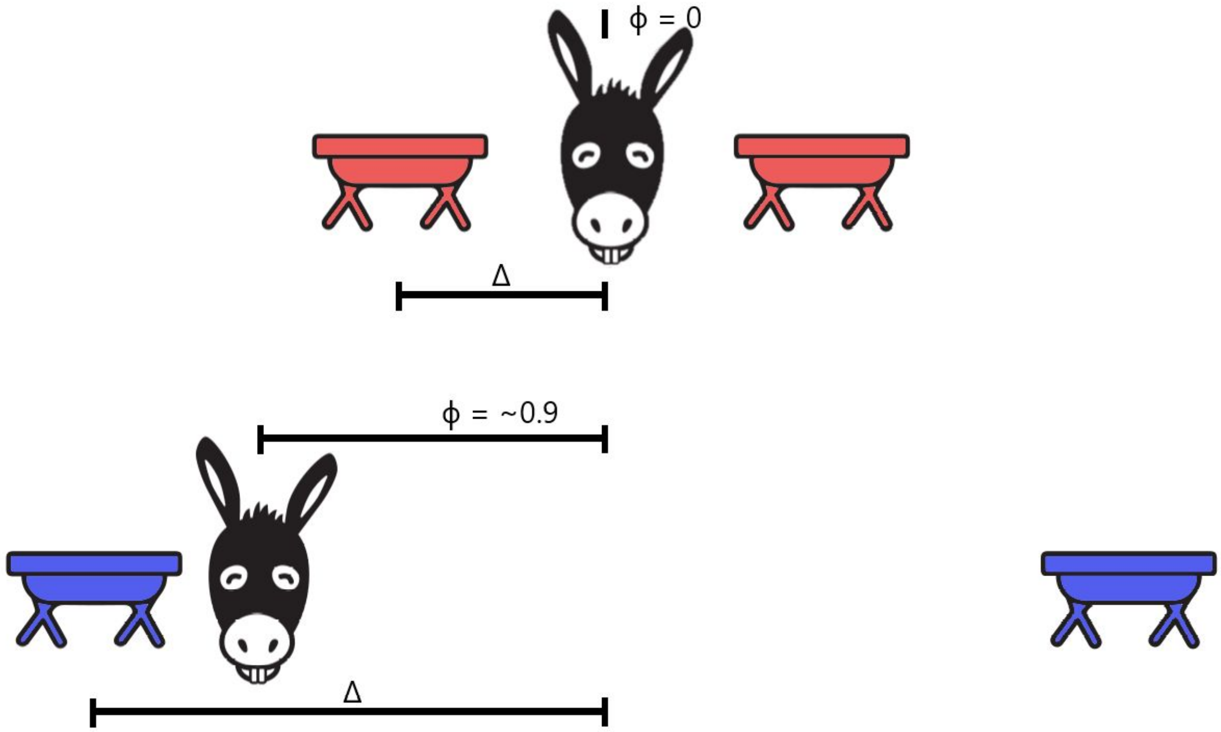

Figure 1: *A schematic example of the new Donkey dilemma. Δ measures how difficult the two goals A and B are to complete. In our case, Δ measures the distance from the mid point to either goal. ϕ represents the position the donkey chose to prepare for the two goals. When Δ is small, both A and B are easy to complete, and so we should position ourselves equally between them, (ϕ=0).*

Expected Utility Theory (EUT) posits that a decision maker will select between various courses of action by comparing the expected utility, that is, the value of an outcome and its respective probability (Mongin, 1997). There are obvious limits to this strategy; the decision problem becomes intractable as the number of possible outcomes to compare grows (see Bossaerts & Murawski, 2017, for a discussion of this). Therefore, when determining optimal decisions, one must take into account the limitations of the individual. The term Bounded Rationality (Simon, 1990) refers to the idea that people make decisions that are optimal given their own limitations. Bounded rationality asserts that when people make decisions that appear to be irrational, they can usually be explained by limitations on cognitive systems like memory and attention. However, the focus-or-divide dilemma is unlikely to push the limits of our cognitive systems. The solution is not difficult to compute, and implementing even an approximation of the optimal solution leads to accuracy that matches optimal accuracy. Participants have been shown to have all the information they need to implement the correct solution (James et al, 2017). Moreover, participants who explicitly understand the optimal solution can implement it easily (Hunt et al., 2019). In other words, this decision problem has an obvious-in-retrospect solution, but something about the way the problem has been presented to participants prevents them from discovering or implementing it.

One possible barrier to making optimal decisions is interference from task-irrelevant factors, such as framing effects. In a now-famous example, Kahneman & Tversky (1979) presented people with two choices: Option A offered participants a 50% chance to win $1000 and a 50% to win nothing, and Option B would give them $450 with 100% certainty. In terms of expected utility, Option A is the better choice as this would on average reward the participant $500. However, this was not the observed pattern of behaviour for most participants, who tended to prefer Option B. At first glance this may seem consistent with risk aversion, a long-established (and arguably rational) tendency to prefer a sure outcome over an uncertain one (e.g. Holt and Laury, 2002). But Kahneman and Tversky (1979) found that framing problems as a loss instead of a gain led participants to reverse their bias. That is, people choose a 50% chance to lose $1000 over a sure loss of $450 even though now the expected utility of Option B is the higher of the two. By showing that the choices people make can be altered by simply changing the way a problem is framed, Kahneman & Tversky (1984) demonstrate that we are not simply comparing options and choosing the one that offers the highest utility. Kahneman and Tversky's insights would seem to argue against bounded rationality, and instead suggest human decisions depend on cognitive shortcuts, or heuristics, which can be biased or flawed in systematic ways. The changes in decisions that can be observed when the same problem is presented in different ways can reveal the underlying basis for decision heuristics.

The choice problems used to evaluate human decisions are often abstract and hypothetical, with participants being expected to imagine they are in a situation where choosing either option would result in one of the two stated outcomes. These decisions are often unusual circumstances for most people, such as preventing deaths or winning/losing sums of money. As such, questions have been raised about how well behaviour in these hypothetical decisions, with hypothetical large consequences, represent the behaviour of people in real life situations (Camerer and Hobbs, 2017). In the current study, we rely on concrete decisions with easily-observable outcomes. Our key question is whether re-framing the same decision problem but with nonequivalent options could facilitate decisions. That is, does a Buridan's Ass dilemma interfere with the process of weighing up alternatives and selecting actions that maximize utility, and can this explain the sub-optimal decisions observed in other studies that involve choices between equal options (e.g. Clarke and Hunt, 2019; James, Clarke and Hunt, 2017; Morvan and Maloney, 2012)?

In the case of the focus-or-divide dilemma presented here, participants are required to throw a beanbag into one of two hoops. Importantly, they are only told which hoop is the target after they have chosen a standing position. According to the logic described above, participants should stand between the hoops when they are close together, and next to one hoop when they are far apart. In previous versions of this experiment, the majority of participants did not modify their standing position with the distance between hoops. This may be because participants are focused on other aspects of the task that may have caused them to ignore the relevant information. The choice between attempting both targets or focussing on one is not arbitrary; however, choosing to focus on one target entails an arbitrary decision about which target to focus on. That is, expected utility is the same regardless of whether they chose to prioritise *A* or *B*. This decision presents a Buridan's Ass dilemma. Although participants in previous research do not (like a starving donkey) take an infinite amount of time to decide upon a course of action, they also do not perform the task in a rational way. One way to avoid having to make this arbitrary choice is to always pursue both goals, regardless of their difficulty (a behaviour adopted by a small but consistent subset of participants in Clarke & Hunt's (2016) experiments, who stood in the centre of the two hoops on every trial). Another possible reaction could be to over-generalize, that is, even though only a component of the overall choice is arbitrary, participants may make choices that are entirely arbitrary, and their choices are thereby variable, but insensitive to the difficulty of the tasks. This describes the behaviour of the majority of participants in the experiments of Clarke and Hunt (2016).

The hypothesis we explore throughout this manuscript is that breaking the symmetry between *A* and *B* will lead participants to behave more rationally. We test this in three different ways. In Experiment 1, we introduce a rationale for choosing one of the two options by making one easier than the other (i.e., $P_s(A|\phi) \neq P_s(B|\phi)$). In Experiment 2, we compare the standard condition, in which only one of the potential goals becomes the target, to a condition where both potential goals are known to be the target, which removes the pressure on the participant to try and predict this aspect of the trial. Experiment 3 gives participants the opportunity to introduce their own asymmetry to the problem by letting them decide how to split a monetary reward between the two potential targets. Across all three experiments, the basic decision dilemma and the solution remains the same: we present trials where the two targets are close enough that both can be reached from a central position, and trials where they are far enough apart that participants would achieve better accuracy by committing to one or the other. Breaking the symmetry between the two targets has little bearing on the first choice (whether to hedge or commit) but makes the second choice (committing) no longer arbitrary. If this equality between the two goals was the reason for the poor decisions, we should see choices that are closer to optimal when we break the symmetry between goals.

Experiment 1: Manipulation of Difficulty

We created an asymmetry between the two potential target locations by making one of the options more difficult, i.e, $P_s(A|\phi, \Delta) \neq P_s(B|\phi, \Delta)$. Similar to Clarke and Hunt (2016), we asked participants to decide where to position themselves in order to throw a beanbag into one of two hoops, but one of the hoops was smaller than the other. To maximise success, participants should stand closer to the smaller hoop, in proportion to the size difference. Such a strategy would demonstrate that participants are sensitive to the relative size manipulation and can use expected performance to modify their standing position choices. This would be consistent with results from James et al. (2017), who demonstrated that participants have reasonably accurate information about their own throwing ability in this task. Taking the probability of success for each of the targets into account in choosing a standing position would also be consistent with *spatial averaging* (Chapman et al., 2010), a behaviour observed in visually-guided reaching. In these experiments, participants need to begin a reach before a target location is known, and reaching trajectories tend to be spatially weighted to reflect the probability of different locations becoming targets.

If having a reason to choose one hoop over the other facilitates optimal decisions, we should see optimal changes in standing position as the distance between the hoops varies. That is, when the hoops are close, accuracy greater than 50% can be achieved by standing at a location between the hoops that maximises the accuracy for both (i.e., centrally, but slightly closer to the small one). As in Clarke and Hunt (2016), we also included distances where accuracy from a central position would be less than 50%; at these distances, standing next to the smaller hoop will ensure accuracy of at least 50%.

Methods

Participants

There were 21 (4 male) participants in Experiment 1. All participants were recruited from the University of Aberdeen community via word of mouth. The protocol for this and all other experiments in the paper were reviewed and approved by the Aberdeen Psychology Ethics Committee. None of the participants had taken part in related studies run by our lab. Power analysis was carried out using bootstrapping methods and previously collected datasets. Full details are available in the supplementary material.

Equipment

The experiment was conducted in a sheltered, outside, paved area. Participants were required to throw bean bags into hoops placed on the ground. The paving slabs (each measuring 0.46cm × 0.61cm) acted as a convenient unit by which to record the placement of hoops and where participants chose to stand. Figure 2 shows an image of the testing area and setup.

Procedure

This experiment was based on the Throwing Task described by Clarke & Hunt (2016) and was conducted over two sessions, conducted on different days at least one week apart[3]. The first session allowed us to measure $P_s(A|\phi)$ for each participant, while session two involved the focus-or-divide decision paradigm.

[3] The week's delay was for logistical reasons; a group of seven undergraduate students helped with data collection as part of a research methods class, and the week delay between sessions simplified coordination of testing.

<u>Session 1.</u> The goal of this session was to obtain a throwing performance curve over distance for each participant for the two hoop sizes. The key difference between this study and the throwing task carried out by Clarke & Hunt (2016) is that there were two hoop sizes (diameters of 63.5cm and 35.5cm) The large hoops were tested at a set of seven distances between 7 and 25 slabs (3.22m - 11.5m). The small hoops were tested at seven distances from 3 to 19 slabs (1.38m - 8.74m). Participants threw 12 bean bags into each hoop size at each of seven distances, for a total of 168 throws. Two different directions were used (hereby referred to as North and South) with the starting direction counterbalanced across participants. The results of Session 1 (presented in full in the supplementary information) were used to model the relationship between *accuracy*, *distance* and *hoop diameter* for each participant using a generalized linear model.

Figure 2: *The top figure shows a schematic outline of the step for the second session. The bottom figure shows the area in which the experiment took place.*

<u>Session 2.</u> Session 2 presented the focus-or-divide dilemma, by asking participants to choose where to stand before the target hoop had been specified. Six hoops, of three different colours, were placed on the paved area (see Figure 2). The red hoops were always closest together, with the yellow hoops being further out, and the blue hoops being the furthest apart. For each pair of hoops, there was a small hoop and a large hoop. The hoop positions in Session 2 were determined by each individual's throwing ability, to equate overall expected accuracy across participants. To do this, we calculated the distances at which a participant would be 10%, 50%, and 90% accurate for both hoop sizes based on the model of their throwing performance in Session 1. The midpoint of these values was then taken so that there would be a common central point for both sizes of hoops. For example, if a given participant was 50% accurate when the large hoop was ten slabs away, and 50% accuracy when the small hoop was eight slabs away, the small and large hoops would both be placed 9 slabs from the centre point in Session 2, to approximate an expected overall accuracy from the center point of 50%. Each colour pair corresponded to expected throwing accuracy (Red = 90%, Yellow = 50%, Blue = 10%) as measured from an unmarked central position, equidistant from both hoops. Hoop size was alternated.

To sample across a range of separations between hoops, a second block of decisions was included. These were set up the same way as the first set (with a hoop pair, defined by color, at each of the three separations). The red pair (one large and one small) was positioned one slab closer than their 50% slab (called 50%-1), the yellow at the 50%+1 slabs, and the blue at the 50%+2 slabs. The two sets of three hoop separations were tested in two blocks of 45 trials each, for a total of 90 standing position decisions per participant in Session 2.

On each trial, participants would draw one beanbag at random from a sack, with the colour of the beanbag indicating which pair of hoops would be the target for that trial. The sack contained nine beanbags (three of each colour: red, yellow, and blue). Once thrown, each beanbag was removed from the paved area. After all nine had been removed from the bag and thrown, the bag was refilled. The bag was set off to the side of the paved area so all participants had to return to this location before each trial. Participants were told that they were allowed to stand anywhere they wanted on the paved area. They were also informed that each hoop was equally likely to be the target, and that the order of target hoops had been predetermined in a random fashion. The data recording sheet used by the experimenter included a printed sequence of 90 targets to follow, on which the target on each trial was independent and randomly selected between north and south hoops. Once participants had stood in their chosen position and informed the experimenter they were ready, their standing position was recorded (in slab units) and they were told which hoop to aim for (as either the "North" or "South" hoop). The experimenter then recorded throwing accuracy, collected the beanbag, and instructed the participant to draw a new beanbag for the next trial.

Analysis

All analyses for this and subsequent experiments were carried out using R (v3.4.3, R Core Team, 2016) with the tidyverse collection of packages (v1.3.0,Wickham et al. (2019). Our main goal in the current experiment was to assess whether or not standing position decisions improved with unequal hoops sizes, and this question can be addressed through a simple presentation of the data. To keep the results section simple and focused on this question, we present a descriptive analysis of the Session 2 results below. A full reporting of all the data from both sessions, and a formal Bayesian analysis, can be found in the supplementary materials. This analysis made use of the brms package (v2.8.0, Burkner, 2017, 2018). Because the descriptive analysis of results addresses the questions the experiments are designed to answer, and the models of the data are complex and lengthy to describe, a similar approach to the results was taken in all the experiments presented in this report.

Results

Standing position. The results, summarised in Figure 3a, show the effect of the manipulation of hoop size on standing position choices. Participants tended to stand closer to the smaller, harder-to-hit, hoop than the larger one. This demonstrates that participants are motivated and capable of responding rationally to changes in the task structure in order to improve their accuracy. However, using hoops of unequal size did not help participants to solve the focus-or-divide task optimally, as there is no systematic tendency to shift from centre to side positions as Δ increases. This replicates previous versions of this experiment (Clarke & Hunt, 2016; James et al., 2017; James et al., 2019).

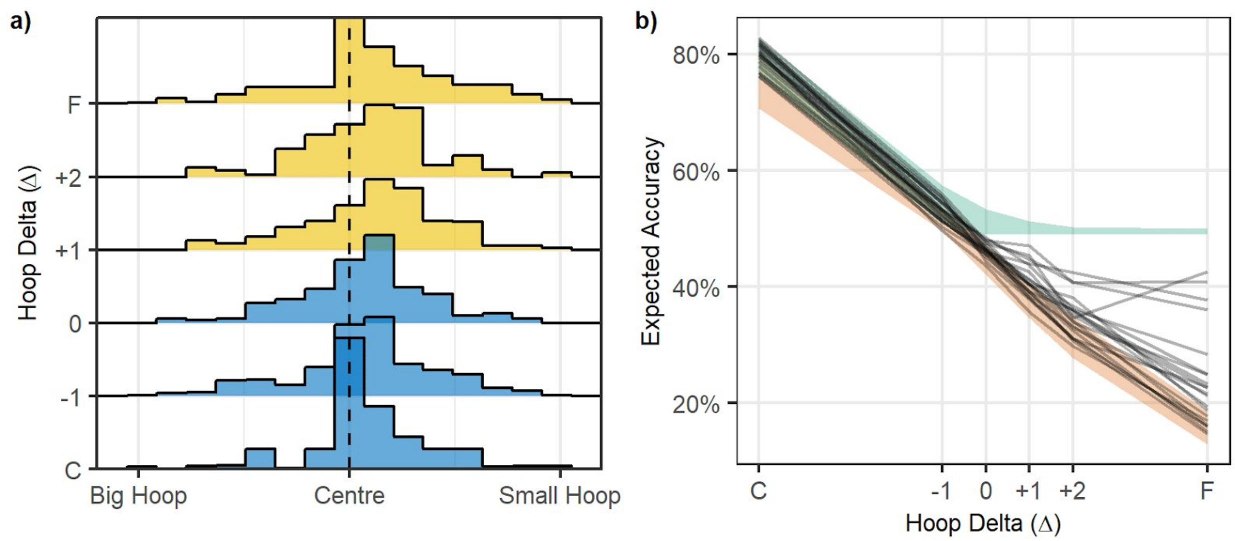

Figure 3. *a)The histograms show the distribution of standing positions for each value of Δ. The increments of Δ increase on the y axis from C (close) to F (far). The colour of the histogram indicates the optimal strategy, with blue representing cases when the participants should have stood near the centre, and yellow the cases where standing next to the small hoop was the best strategy. b) The black lines show Expected Accuracy for each participant for each of the six hoop distances (with close (C) to far (F) distances on the x axis). The green shaded area shows the range of Optimal Accuracy for this group of participants, and the red shows their minimum accuracy (the range of accuracy that would be expected if participants chose the least optimal standing position).*

Accuracy. We use the performance data collected in Session 1 to calculate each participant's optimal accuracy, by estimating what their throwing accuracy would have been had they chosen the optimal place to stand. To obtain a lower bound, we estimate each participant's minimum expected accuracy based on a counter-optimal standing position. In this experiment, the counter-optimal strategy would be to stand next to one of the hoops when they are close together, and in the centre when they are far apart as this would produce the lowest expected chance of success. Both of these measures vary across different participants, and the ranges are shown in Figure 3b. We can compare these to the accuracy that would be expected given a participant's actual standing position choices in Session 2, using their performance in Session 1. We present this measure instead of actual throwing accuracy in order to remove variability due to chance and variation in throwing accuracy from one trial to the next, and to provide an estimate that is more directly comparable to the estimates for optimal and counter-optimal accuracy.

From Figure 3b, it is clear that none of the participants made use of the optimal strategy. Expected accuracy fell far short of optimal accuracy. This is expected from the standing position results, as the majority of participants did not modify their standing position with the distance between hoops. The decrease in accuracy as distance increases is consistent with previous studies and reflects the fact that the decisions have less effect on accuracy when the two tasks are easy.

Discussion

The aim of this experiment was to provide participants with a concrete and intuitive difference between the two targets to use in guiding their decisions. We can conclude from this experiment that participants are sensitive to the differences in hoop size, and adjust their behaviour to increase their expected accuracy. However, creating an asymmetrical decision did not help participants solve the focus-or-divide problem. Participants stood in a central location just as often when the two targets were close together as when they were far apart. This led to accuracy rates that fell far short of what they could have been had an optimal strategy been adopted.

We confirmed and extended this conclusion in an experiment presented in the supplementary material, using a different manipulation of the symmetry between the two possible targets and a different task context. Briefly, we used an eye movement task and presented a small, brief target inside either a left or right square (see also Clarke and Hunt, 2016; James et al., 2019; Morvan and Maloney, 2012). Participants had to choose a place to fixate in anticipation of the target appearing. The distance between the squares determined whether the best location to fixate was between the two squares (when they were closer), or inside one of the two squares (when they were too far apart to be visible from the center). We manipulated the probability of the target appearing in one square over the other (with an 80/20 share across the two locations). The results are in line with the results we observed above: participants adjust their fixation decisions to match the probability manipulation, but do not make more optimal fixation decisions with respect to the distance between the squares.

Experiment 2: Two Throws

Experiment 1 provided participants with a tangible reason to select one location over another. Even though there was a logical reason to favour one position over the other, participants would still experience positive feedback (i.e. they would achieve the goal on that trial) when selecting the "right" location, and negative feedback (i.e. they would miss) when selecting the "wrong" location. The sequence of targets was pre-determined and unpredictable, and participants were informed of this. Nonetheless, it is possible that participants suspected there were predictable patterns to exploit in this sequence, which may have distracted them from relevant information. Searching for patterns is a cognitively demanding task in itself (Wolford et al., 2004), which may have distracted participants from making better focus-or-divide decisions.

In Experiment 2, we added a condition where, instead of only one of the hoops becoming the target, participants were instructed to throw two beanbags - one at each hoop - on every trial. This should remove any reason for participants to try and "discover" an underlying pattern in the task, because there are no random variables except which beanbag color is drawn from the bag (and the participant does the drawing). Importantly, to maximize accuracy to achieve both goals, the optimal strategy remains the same. Standing in the center when the hoops are close together will increase the likelihood that participants will hit both, and standing next to one when they are far apart will ensure that at least one of the goals will be achieved. In the nomenclature presented in the introduction: $P_c(A) = P_c(B) = 1$.

Methods

Participants

Eighteen participants (8 male) took part in this experiment, with an average age of 22 (between 19 and 30). Participants were recruited via word of mouth. None had previously participated in any related experiments. Please see supplementary materials for the power analysis.

Procedure

This experiment followed the same protocol as in Experiment 1, with the following exceptions. First, each hoop was the same size (0.4m in diameter) so standing equidistant from both would give participants an equal chance at each target. The supplementary materials present individual throwing performance from Session 1, which was used to determine the hoop positions for each participant in Session 2, using the same methods described in Experiment 1. Session 2 was split into two conditions: the One-Throw and the Two-Throw condition. The order of these was counterbalanced across participants. The One-Throw condition followed the same procedure as in Experiment 1. In the Two-Throw condition, participants still selected one bean bag at a time from a sack containing nine, with three of each colour. They were then handed a second bean bag of the same colour from a separate pile. Participants were then, as in Experiment 1, instructed to choose somewhere to stand, at which point they would notify the experimenter. They would then throw each bean bag to each of the two hoops of the same colour, in whichever order they preferred. The stated goal for both conditions was to get as many bean bags into the hoops as possible. As before, the experimenter would record the standing position and throwing accuracy on each trial, and clear the beanbag from the paved area after each throw.

Results

Standing Position

From Figure 4a, it is clear that there was little change in standing position with increasing distance between hoops in either the one-throw or the two-throw condition. In the Two-Throw condition, participants opted more often to stand in the centre overall, and variation in standing position increased moderately as the distance grew, though this did not result in participants (as a whole) performing the task in a more optimal way (see accuracy results below). A more detailed Bayesian analysis of these results reinforces this general pattern and interpretation, and is presented in the Supplementary materials.

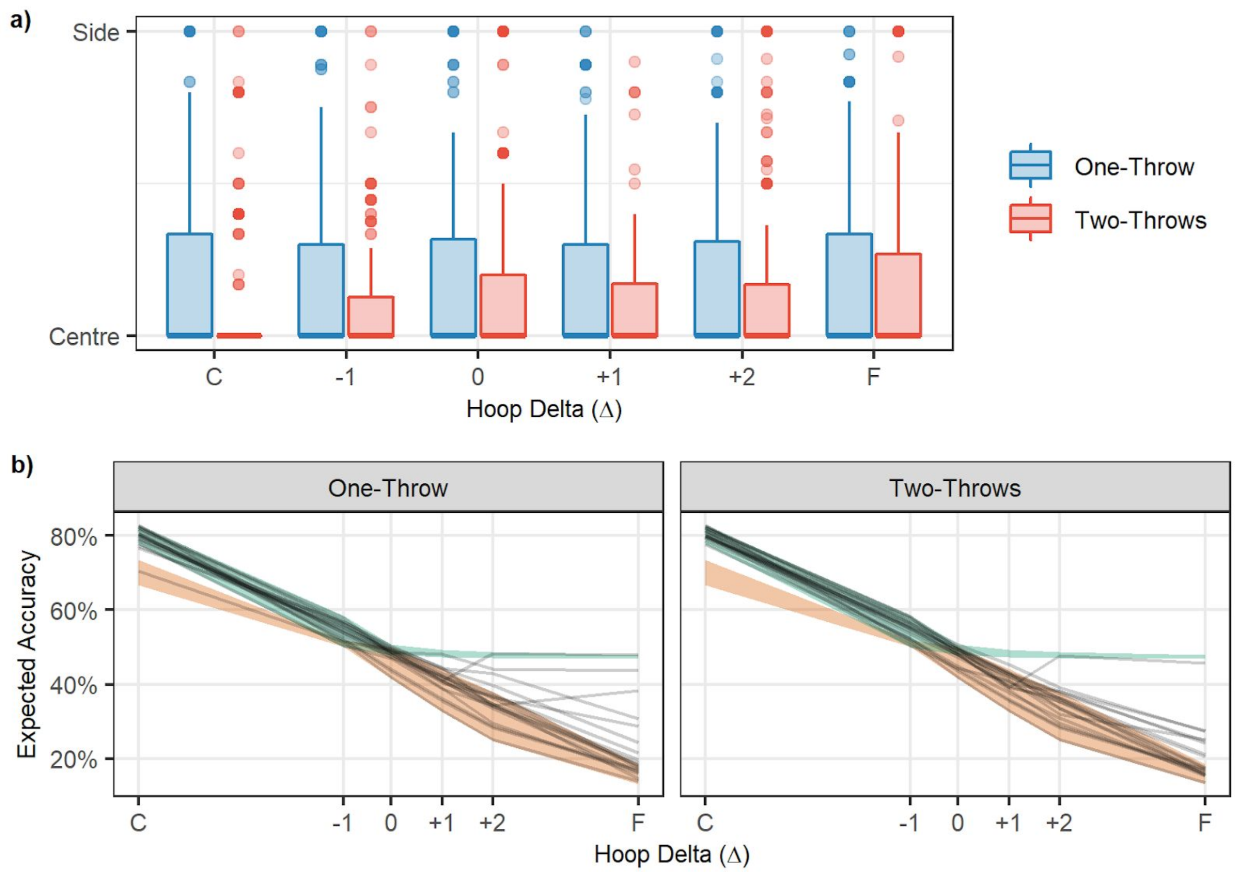

Figure 4. a) *The boxplots in the top panel show the distribution of standing positions across the distances tested in this experiment. Please note that these boxplots are highly skewed in all cases; the median standing position was the center. b) The black lines in the bottom panels show the expected accuracy for each participant based on their standing positions across the distances tested. The shaded areas represent boundaries on accuracy, with green showing expected accuracy for optimal decisions and orange for counter-optimal decisions.*

Accuracy

Figure 4 shows that most participants' accuracy (black lines) dropped below that which would be expected had they employed the optimal strategy (green). Note that, like for Experiment 1, we are calculating an estimate of accuracy based on the standing position choices and throwing performance of each individual, to remove variability and ease comparison to optimal and sub-optimal baselines. One of the 18 participants was close to achieving optimal expected accuracy in both conditions. This is consistent with other results showing the occasional participant approaching an optimal strategy, but the majority performing far below this (e.g. in Clarke and Hunt (2016) Experiment 2, 1 of the 12 participants approached optimal).

Discussion

Most participants failed to perform optimally in the task, whether they had to throw a beanbag into one hoop, or both. The results rule out a search for patterns as the reason for suboptimal decisions, because the decisions are similarly sub-optimal in the absence of any uncertainty about which hoop would be the target. Removing this uncertainty seems to have pushed participants to stand closer to the centre. This suggests that, in general, participants were not willing to give up on achieving one goal in order to achieve a higher chance of success at another. By moving to stand next to one hoop, in the Two-Throw condition participants would be ensuring at least one loss. Perhaps the prospect of a certain loss outweighed the certain success that would have been achieved at their chosen target. In any case, the results rule out an explanation for poor decisions based on the distraction of not knowing which hoop would be the target.

Experiment 3: Unequal Reward

The expected utility of a particular course of action is the sum of the expected gain or value of each possible outcome multiplied by the probability of each of these outcomes occurring. Thus far, the "gain" associated with each target was still symmetric: leaving aside the intrinsic reward of getting the beanbag in the hoop, in terms of monetary value, the two goals were equally null. In Experiment 3, we introduced monetary rewards for accurate performance, resulting in the new expectation:

$$E(\phi, \Delta) = R_A P_c(A) P_s(A|\phi, \Delta) + R_B P_c(B) P_s(B\,|-\phi, \Delta) \tag{3}$$

where R_A *and* R_B *are the rewards for achieving goals A and B*.

In the previous experiments using this paradigm, $R_A = R_B$. In the current experiment, we gave participants the opportunity to decide on the relative value of the two targets. Before each trial, we asked them to choose either an equal split ($R_A = R_B$), or to assign 80% of the reward to one target and 20% to the other (i.e., $R_A = 0.8$, $R_B = 0.2$). With this design, participants are offered the chance to avoid the Buridan's Ass dilemma altogether, by ensuring that they are no longer equidistant from two equally rewarding options. This allows us to evaluate participants' preference for symmetrical options, as well as the effect of asymmetrical values on their decisions (presuming they choose these).

Rewards have been shown to improve decisions in some contexts (e.g. Goodnow, 1995; Phillips and Edwards, 1966), although there are limits (for a review, see Camerer & Hogarthm, 1999). In a gamified version of the focus-or-divide dilemma (James et al, 2019), in which a penguin character could earn fish rewards for accurate performance, participants improved their performance on the task relative to participants who were not given this additional motivation.

However, a close inspection of the reason for the improvement revealed that it was not a result of participants making better prioritizing decisions, rather, participants performed better in other aspects of the task (such as making fewer key-press errors, and monitoring locations more vigilantly). Decisions with or without motivation remained equivalently sub-optimal. Similarly, offering financial incentives for accuracy (Morvan & Maloney, 2012) did not improve decisions relative to not doing so (Clarke & Hunt, 2016). Overall, this suggests a lack of extrinsic reward is unlikely to explain the suboptimal decisions observed in the throwing experiments.

The current experiment goes beyond simply assessing the effect of rewards, by providing insight into how the “Buridan’s Ass” dilemma is regarded by participants. Participants were given 50p on each trial and were given the choice between splitting the money equally (25p/25p) or unequally (40p/10p) across the potential targets. That is, before they made a decision about where to stand, they were given an opportunity to designate one of the targets as more valuable than the other. Irrespective of the distance, participants should opt for an unequal split and then stand next to the hoop with the greater value. If equally-rewarding and equally difficult hoops creates an unpleasant dilemma that distracts participants from the optimal solution, at least some of our participants may recognize this and choose to make one choice more valuable than the other, and in so doing, maximize their gain in the experiment.

A related question is whether consistent relationships will emerge between hoop distance, people’s choices of how to split the reward, and their choices in where to stand. The requirement to judge how to divide the reward might nudge people to think about the consequence of the hoop separation more carefully. As the hoop separation increases, they may consider splitting the reward unevenly, because they recognize that they are likely to fail from using a central strategy at far hoop separations. Thus, there may be a tendency, at least among some participants, to divide the reward unevenly at far separations. Among these participants, they may also commit to standing closer to the hoop they have made more valuable. Indirectly, this could lead participants who split the reward unevenly to approach an optimal strategy.

Methods

Participants

20 participants took part in this experiment (15 female) with an average age of 22.6 (between 20 and 30). All participants were recruited via word of mouth at the University of Aberdeen. Please see supplementary materials for the power analysis.

Procedure

Participants signed a consent form which contained details about the reward schedule and how much they could expect to earn on average. All participants were given £4 as a baseline and were told that they could expect to earn an additional amount ranging from £0 to £4.80

depending on their performance. They were also told that we expected them to earn between £1.50 and £2.50 on average.

This experiment followed a similar procedure to that of Experiments 1 and 2. However, in this experiment, both the measuring and decision sessions took place in one session. First, participants were taken to the paved area in order to measure their throwing ability across the same eight distances used for the small hoops in Experiment 1 (slabs 3 to 19 slabs or 1.38m to 8.74m), with participants throwing 12 bean bags for each distance (96 trials total). After this, they performed a brief computer-based task in the lab. The task was a brief pilot of an unrelated experiment, which involved detecting shapes among cluttered and uncluttered scenes. This was done only to make efficient use of participants' time while the experimenter calculated their performance curves and set up the hoops for the second session, much in the same way as before. For the final task, participants were taken back to the paved area to complete the decision session, as follows.

There were two main changes to the paradigm. First, four hoop distances were used: the distances at which participants were 90%, 75%, 25%, and 10% accurate, based on their individual Session 1 performance. Each of these distances was tested 3 times for a total of 12 trials. Second, participants were told they had 50p to split between the two target items in one of two ways. They could either split it equally across both potential targets (25p/25p) or make it an 80/20 split (40p/10p). Participants were asked how they would like to split the money before they made a choice about where to stand. If they opted to make an unequal split, they were asked which hoop they would like to be worth 40p and which 10p. Participants were informed that the target hoop had been randomly predetermined so that each hoop was equally likely to be the target on each trial. It was reaffirmed that any money they earned by successfully throwing the bean bag into the target hoop would be given to them upon completing the experiment. Participants would then pick a place to stand, at which point they would be told which hoop was the target for that trial. The experimenter recorded standing position and throwing accuracy and cleared each beanbag from the area after each throw.

Results

We can see from Figure 5a that participants made use of both the 50/50 and 80/20 reward splits. Overall they favoured the equal split, although this tendency decreased as Δ increased. Figure 5b shows individual choices of how to split the reward, and it is clear from this figure that participants are inconsistent in their choices about whether to split evenly or not. While five participants opted for an unequal split most of them time (mean proportion of 1.26/3), no participants followed the optimal rule of always splitting the value unequally and standing next to the more valuable target, which would have given participants a larger rate of return.

Standing position choices are shown in Figure 5c. When compared to the results of Experiments 1 and 2 (Figures 3a and 4a), standing position choices in this experiment appeared to be more sensitive to Δ. This sensitivity was more pronounced when participants had opted for an unequal split, shown in the red box plots. A more detailed Bayesian analysis of these results reinforces this general pattern and interpretation, and is presented in the Supplementary materials.

Figure 5d presents the expected earnings per trial. This value represents the expected earnings based on Session 1 throwing performance and the chosen standing positions, in order to ease comparison with the *optimal expected earnings:* this is what participants would have earned if they had adopted the optimal strategy of splitting the money unequally and then standing close to the more valuable hoop. Most participants fall far short of this. The lower bound is the counter-optimal earnings: this is what participants would have earned if they had made the poorest possible choices. The poorest possible choice would be to split the monetary reward unequally and then standing next to the lower value option for all distances except the furthest, in which case, standing in the centre was the worst choice for participants.

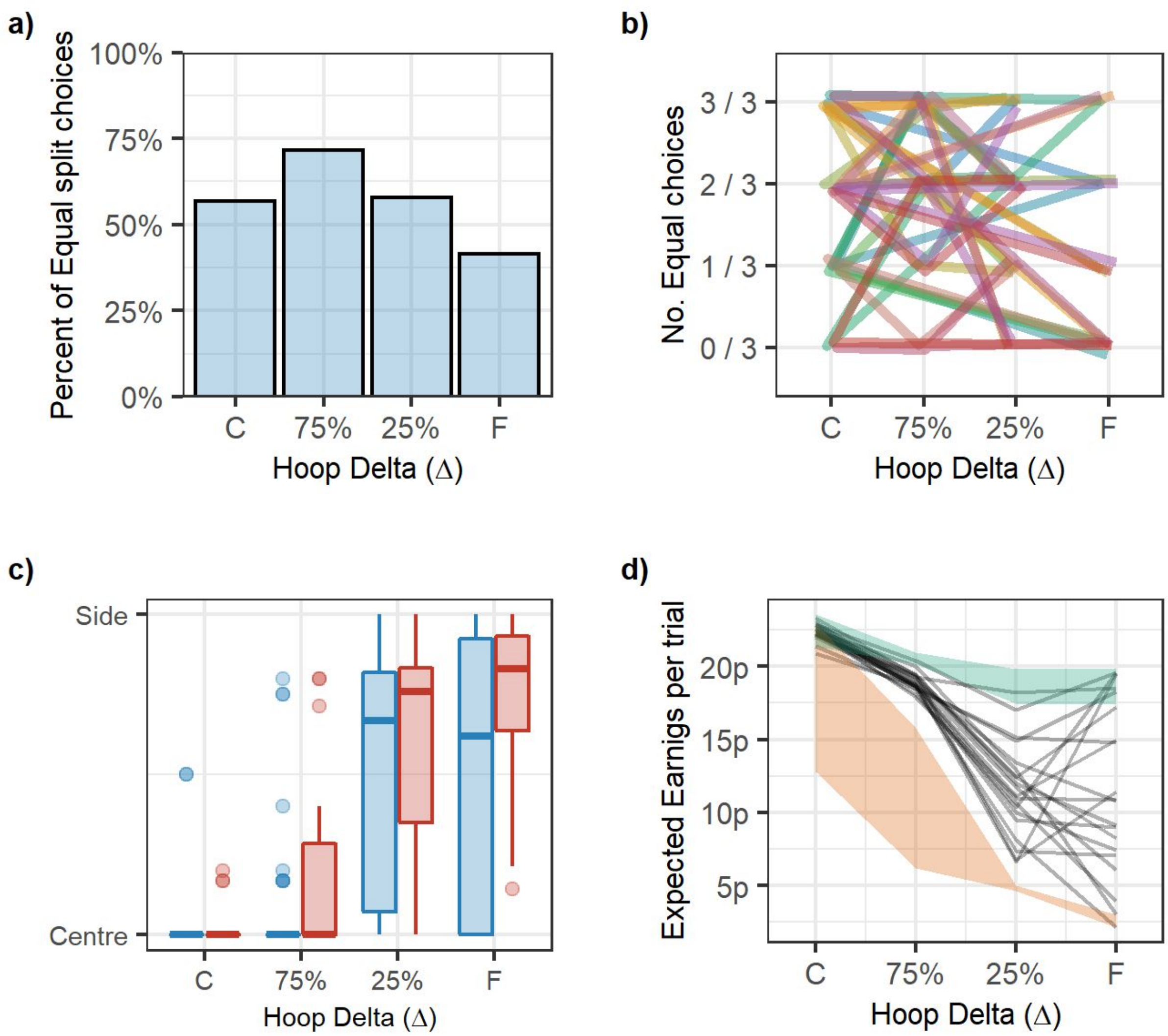

Figure 5: *a) The proportion of times that participants opted for an equal split across the different distances tested. b) This figure shows the same data as 5a, but each participant has a unique line to trace their decisions across each distance. c) boxplots indicating the standing positions for the four distances tested with the colours (red for unequal and blue for equal) showing whether participants had opted for an Equal or Unequal split. d) This figure shows the average expected earnings across the different distances tested. The shaded areas represent the ranges for the greatest expected outcome (green) and the worst (orange).*

Discussion

There is some evidence from this experiment to suggest that reward does facilitate more optimal decisions in a focus-or-divide dilemma, given that standing position was modified by Δ more in this experiment than had been observed in previous ones. However, reward for successful performance is only part of the story. On trials in which participants chose to split the reward unequally, they were more likely to choose standing positions that were closer to one of the potential target hoops than had they opted for an equal split. As such, participants needed not only a reward, but also the option to split the money unequally before they were willing to move away from one target when the two targets were too far away to achieve success from the center. We can therefore conclude that the introduction of a monetary incentive to prioritise one target over the other was indeed effective in causing some participants to vary their strategy with distance more consistently. Given this, it is striking that participants persist in splitting the reward equally on the majority of trials. Even when the two targets were at the furthest separation, participants chose an equal split on over a third of the trials, leading to losses in potential monetary gains. Not only did participants stand to gain less money if they were successful when they had split the money equally, but participants were also less likely to be successful when they split the money equally, because they were less likely to adopt the optimal strategy of standing next to one hoop.

General Discussion

In line with previously published work on this paradigm (Morvan & Maloney, 2012; Clarke & Hunt, 2016; James et al., 2017; James et al., 2019), the results showed that our participants fail to use the optimal strategy in a focus-or-divide dilemma. We had speculated that participants' ability to adjust behaviour ϕ to maximize expected accuracy $E(\phi, \Delta)$ might depend on having clearly different expected accuracy levels for tasks A and B for a given ϕ. That is, for this decision problem:

$$E(\phi, \Delta) = P_c(A) \times P_s(A|\phi, \Delta) + P_c(B) \times P_s(B|-\phi, \Delta)$$

The hypothesis was that participants may be able to adjust ϕ to maximize E only when $P_s(A|\phi) \neq P_s(B|\phi)$. In exploring this question, we extended the previous findings of suboptimal decisions to include conditions where the decision between the two goals is no longer arbitrary. The current study thereby rules out a Buridan's Ass dilemma as a plausible explanation for the poor decisions shown in previous studies.

In all three experiments, participants had to choose a place to stand to throw a beanbag into one of two hoops. When the hoops are close together, and the participant does not know which hoop will be the target, standing in a central position ensures equal accuracy for both possible targets. As the distance from the center to the hoops increases, beanbag throwing accuracy drops below 50%. For hoops beyond this distance, accuracy of at least 50% can be achieved by standing close to one of the two hoops instead of staying near the center. In previous studies, choosing which hoop to stand near could be considered a Buridan's Ass dilemma, because the two options were equal in terms of likely success and reward. In Experiment 1 of this series, one of the hoops was smaller than the other. Participants clearly made an attempt to maximise their rate of success by standing closer to the smaller hoop, suggesting they are sufficiently motivated and capable of doing so. Nonetheless, they were not able to maximize accuracy when it came to modifying their decisions as the distance between the hoops changed. This resulted in accuracy that was far worse than could have been achieved, particularly when the hoops were far apart and participants persistently stood close to the center, causing their throwing accuracy to fall far short of 50%. These results show that giving participants a reason to prioritise one target over another did not lead to more optimal focus-or-divide decisions. A similar experiment in the supplementary material manipulated probability and showed a similar pattern, with participants usually, but not always, prioritizing the more likely target. These results support the notion that people are sensitive to information about the expected utility of different options and will make use of it when making decisions (Gao & Corter, 2015; Wolford et al., 2004; Yellott, 1969). However, participants were no more likely to solve the focus-or-divide dilemma optimally.

Experiment 2 included a condition where participants had to throw two beanbags, one at each hoop. In the two-throw conditions, participants no longer had to guess which hoop was likely to be the target; they knew both were targets. If uncertainty about which hoop would be the target was distracting people from the optimal strategy, we reasoned that the two-throw condition would improve accuracy relative to the one-throw version of the experiment. Inconsistent with this prediction, the mean throwing accuracy for the two hoops that would be expected based on where participants chose to stand in the farthest hoop separation condition was 22% and 21% in the one-throw and two-throw conditions respectively. Not only are these values similar, they are far lower than the 50% accuracy that would be expected if participants had chosen to stand close to one hoop when they were far apart.

In Experiment 3, participants were given the choice to split a monetary award equally or unequally across the two hoops. Participants tended to stand closer to one of the targets when they had opted for an unequal split in the reward (Figure 5c). This suggests that the interaction between asymmetry and financial reward facilitates the use of a more optimal strategy. Interestingly, however, none of our participants managed to consistently follow the optimal strategy of always opting for an unequal split, then standing next to the most valuable target. Even more surprising, they did not prefer to split the money unequally, choosing to do so on less than half of trials overall. Unlike previous results from this paradigm, participants did tend to vary their strategy more appropriately with task difficulty, particularly when they had opted for unequal rewards. This result suggests that choosing unequal rewards can lead to better focus-or-divide decisions. Further research could explore the effect of equal and unequal rewards under conditions where they have been chosen by the participants versus by the experimenter, but this question is outside the current focus of this study. For the present purposes, the results clearly demonstrate that people do not avoid Buridan's ass dilemmas when given the opportunity. In fact, they seem to prefer them over unequal options.

One caveat about the final experiment is that it is possible we sampled a group of participants who were more likely to make better focus-or-divide decisions in the first place. In previous experiments, a handful of individuals do approach optimal strategies in this task, so having more of these participants in the sample could lead to the impression that the conditions improved their performance, rather than having been better in the first place, which is why we have made within-group comparisons wherever possible in these experiments. More generally, individual differences in decision strategies are common, and present an important challenge for explaining the biases and heuristics people tend to use (e.g. Clarke, Nowakowska and Hunt, 2019; Jasper, Bhattacharya, & Corser, 2017; Zhang, Morvan, Etezad-Heydari, & Maloney, 2012). In our particular decision problem, the solution is trivially easy to implement when it is known (Hunt et al., 2019), so it is important to ensure the participants have not been exposed to the decision problem before. Our participants were naive insofar as they had not participated in any of our previous decision experiments, but we are sampling from a population of undergraduates who all complete psychology experiments regularly, and may communicate with one another about different experiments they have participated in. An experiment which involves throwing beanbags at hoops may stand out among the others, which tend to involve computer-based tasks and questionnaires, so the population we sampled from may be less naive than the general population. Given that this was the last in a series of similar experiments, this population may have had more knowledge of the optimal solution than was the case earlier for the previous experiments.

The current results have some elements in common with previous findings suggesting that people avoid situations that have the potential to incur a sure loss (Chapman et al., 2010; Kahneman & Tversky, 1979; Hudson et al., 2007). Opting for the unequal reward conditions, for example, may be partly due to a reluctance to "give up" on the less rewarded target. A similar drive to avoid sure loss could explain why some participants continue to stand in the center between two distant targets even though they know they cannot throw accurately to either one (James et al., 2017). A preference for equal chances across both targets cannot be the entire explanation for poor decisions in the focus-or-divide dilemma, however, because if this were the case, participants would consistently "divide" instead of focus -- in the context of the bean-bag throwing version of the dilemma, this would lead participants to consistently stand at, or at least near, the center. This is not the pattern that is observed in studies so far, however, including the experiments in the present series -- some participants do stand in the center consistently, but just as many shift towards one target or another (see supplementary material to see the full range of individual decisions), and most make a wide range of different decisions about where to stand. The main consistent finding across experiments is that participants do not adjust their strategy as the distance between the hoops increases, but the behaviours they adopt instead of this adjustment are widely varied and cannot be accounted for solely by a bias to equate the odds of success across both targets.

The series of experiments conducted here converge on the conclusion that poor decisions in a focus-or-divide dilemma are not a consequence of the Buridan's Ass situation entailed by previous investigations, in which there are two tasks that are equally difficult and equally rewarded. The results also demonstrate a tendency to prefer these balanced odds over unbalanced ones. In that sense, the Buridan's Ass situation, at least in this context, seems to be a preferable state for most participants.

References

Bossaerts, P., & Murawski, C. (2017). Computational complexity and human decision-making. Trends in Cognitive Sciences, 21(12), 917 - 929.

Brainard, D. H., & Vision, S. (1997). The psychophysics toolbox. Spatial vision, 10, 433–436.

Burkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. Journal of Statistical Software, 80(1), 1–28. doi: 10.18637/jss.v080.i01

Burkner, P.-C. (2018). Advanced Bayesian multilevel modeling with the R package brms. The R Journal, 10(1), 395–411. doi: 10.32614/RJ-2018-017

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

Camerer, C., & Mobbs, D. (2017). Differences in behavior and brain activity during hypothetical and real choices. *Trends in cognitive sciences*, *21*(1), 46-56.

Chapman, C. S., Gallivan, J. P., Wood, D. K., Milne, J. L., Culham, J. C., & Goodale, M. A. (2010). Reaching for the unknown: Multiple target encoding and real-time decision-making in a rapid reach task. Cognition, 116(2), 168 - 176.

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

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

Frederick, S. (2005). Cognitive reflection and decision making. The Journal of Economic Perspectives, 19(4), 25-42.

Gaissmaier, W., & Schooler, L. J. (2008). The smart potential behind probability matching. Cognition, 109(3), 416–422.

Gallivan, J. P., & Chapman, C. S. (2014). Three-dimensional reach trajectories as a probe of real-time decision-making between multiple competing targets. *Frontiers in neuroscience*, *8*, 215.

Gao, J., & Corter, J. E. (2015). Striving for perfection and falling short: The influence of goals on probability matching. Memory & Cognition, 43(5), 748–759.

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

Holt, Charles, A. and Susan K. Laury. 2002. Risk Aversion and Incentive Effects . *American Economic Review*, 92(5):1644-1655.

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

James, W. R. G., Clarke, A. D. F., & Hunt, A. R. (2017). Failure to use probability of success in deciding whether to pursue one goal or two. Proceedings of the 39th Annual Conference of the Cognitive Science Society, 2279 - 2283. (Austin, TX: Cognitive Science Society)

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.

Jasper, J. D., Bhattacharya, C., & Corser, R. (2017). Numeracy predicts more effortful and elaborative search strategies in a complex risky choice context: A process-tracing approach. Journal of Behavioral Decision Making, 30(2), 224–235.

Kahneman, D., Slovic, P., & Tversky, A. (1982). Judgement under uncertainty: Heuristics and biases. Cambridge University Press, Cambridge.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. Econometrica, 47(2), 263-291.

Kahneman, D., & Tversky, A. (1982). On the study of statistical intuitions. Cognition, 11(2), 123 - 141.

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

Kiani, R., & Shadlen, M. N. (2009). Representation of confidence associated with a decision by neurons in the parietal cortex. Science, 324(5928), 759–764. doi: 10.1126/science.1169405

Koehler, D., & James, G. (2010). Probability matching and strategy availability. Memory & Cognition, 38(6), 667–676.

Koehler, D., & James, G. (2014). Probability matching, fast and slow. The psychology of learning and motivation, 61, 103–131.

Mongin, P. (1997). Expected utility theory. Handbook of economic methodology, 342350.

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

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

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

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

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

Reimers, S., Donkin, C., & Le Pelley, M. E. (2018). Perceptions of randomness in binary sequences: Normative, heuristic, or both?. *Cognition*, *172*, 11-25.

Simon, H. A. (1990). Invariants of human behavior. Annual review of psychology, 41(1), 1–20.

Sloan, L. L., Rowland, W. M., & Altman, A. (1952). Comparison of three types of test target for the measurement of visual acuity. Q Rev Ophthalmol, 8(1), 4–16.

Stanovich, K. E., & West, R. F. (2008). On the relative independence of thinking biases and cognitive ability. Journal of personality and social psychology, 94(4), 672.

Szollosi, A., Bago, B., Szaszi, B., & Aczel, B. (2017). Exploring the determinants of confidence in the bat-and-ball problem. Acta Psychologica, 180, 1 - 7.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2011, May 04). The cognitive reflection test as a predictor of performance on heuristics-and-biases tasks. Memory & Cognition, 39(7), 1275.

Wickham et al., (2019). Welcome to the tidyverse. Journal of Open Source Software, 4(43), 1686, https://doi.org/10.21105/joss.01686

Wolford, G., Newman, S. E., Miller, M. B., & Wig, G. S. (2004). Searching for patterns in random sequences. Canadian Journal of Experimental Psychology/Revue canadienne de psychologie exp´erimentale, 58(4), 221.

Yellott, J. I. (1969). Probability learning with noncontingent success. Journal of mathematical psychology, 6(3), 541–575.

Zhang, H., Morvan, C., Etezad-Heydari, L.-A., & Maloney, L. T. (2012, 10). Very slow search and reach: Failure to maximize expected gain in an eye-hand coordination task. PLOS Computational Biology, 8(10), 1-12.