Supplementary information

For: 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

Contains

1. Context
2. Experiment S1: Effect of table and solution of a related logic puzzle on throwing task decisions
3. Experiment S2: Repeated-measures effect of table task on throwing task decisions
4. Additional plots and analysis

### 1. Context

The paradigm used in this series of experiments took inspiration from Morvan and Maloney (2012), whose results suggested the eyes are not directed to the locations that maximize information gain during visual search, even when it is relatively simple to identify and select these locations (see also Nowakowska, Clarke & Hunt, 2017). Clarke and Hunt (2016) demonstrated that these idiosyncratic and sub-optimal decisions are not unique to eye movements, but are a more general feature of choice behaviour. Here we have shown that people default back to these sub-optimal and variable decisions even after being guided to carry out optimal choices. One of the interesting questions to tackle next is why this default persists. Our tentative explanation is that choice variability could be a reliable approach to solving problems under many natural circumstances. An interesting relationship between variability and learning was established in 1930’s based on research on rats in mazes, but the extent to which this approach to solving choice problems extends to humans is not known. Fascinating questions remain to be explored about both the benefits, and potential costs, of using reason or teleological control in making choices, relative to relying on trial-and-error learning. The paradigm we present in this paper provides a useful technique, and the results a clear basis, for exploring some of these questions.

The following two additional experiments provide further tests of the hypothesis that correctly performing the table task will increase the proportion of optimal choices in the throwing task. Both replicate Experiment 2 in the main paper in the sense that there is no clear evidence that participants who perform the table task are able to transfer those optimal choices to the throwing task. There was a small effect in E2 in the main paper of slightly more optimal choices in the group that performed the table task. This small effect is absent in these experiments. Both experiments were undergraduate thesis projects and have been attributed to their authors.

### 2. Experiment S1: Effect of table task and solution of a related logic puzzle on throwing task decisions

*Jessie Grimsey & Michelle Tran, University of Essex*

**Summary**

In this experiment, participants were randomly assigned to one of three groups. One group completed the table task, replicating the conditions of Experiment 2. A second group completed a set of five logic puzzles, two of which were abstract versions of the decision dilemma posed by asking participants to choose a place to stand. The third (control) group completed a set of arithmetic problems. All three groups then completed an abbreviated version of the throwing task. The logic question required participants to take a more abstract and hypothetical approach to the decision about how to allocate limited resources when faced with two possible tasks to complete. The hypothesis tested in this experiment is that an abstract approach to solving the problem would be more effective than the table task in priming participants to make better choices about standing position in the subsequent throwing task. Replicating the results from Experiment 2 (and Clarke and Hunt, 2016), standing position decisions were variable and sub-optimal. There was no discernible effect of group on either standing position choices or accuracy. Consistent with the conclusions from Experiment 2, the results suggest priming, either with the table task or an abstract version of the problem, does not lead participants to make better decisions.

**Methods**

24 participants (17 female, mean age 21) from the University of Essex were randomly assigned to one of three groups: a table task group (nine participants), a maths group (eight participants), and a logic puzzle group (seven participants).

Six hula hoops (0.5m in diameter) and nine bean bags were used. They were either green (+/- 1m from the centre) blue (+/-4.5m) or red (+/-7m). The target hoop and bean bag were presented in a randomized order. 15 crosses were marked -7m to +7m (including 0 to represent the centre of the hoop layout (similar to Figure 1)). These distances were based on the data from Clarke & Hunt (2016) and further pilot studies. There were three trials for each distance, giving a total of nine.

Participants in the maths condition complete ten simple arithmetic questions (i.e., 27-12 = ). The logic condition involved five questions, 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 maths condition had ten minutes to complete the questions.

The table task was conducted in a similar manner to Experiment 2 in the main paper, with bean bags positioned at separations of 0, 0.6 and 1.2m apart.

**Results**

Mean accuracy for the maths questions was 91.25%. All participants except one behaved optimally in the table condition. Performance on the logic questions was much lower (around 30%).

The standing positions chosen by each participant are shown in Figure S.1.1. They are broadly similar to those shown in Figure 2 in the main text, with variable choices in standing position that do not follow an optimal strategy, overall. There is no obvious difference in choice behaviour between the groups.

Calendar

Description automatically generated

Figure S.1.1. Normalized standing position for each of three distances between hoops. Each dot represents one trial and each panel is a participant. The throwing task was preceded by the reaching/table task, a relevant logic puzzle, or math questions. Which task the participant performed is denoted by the dot color.

Accuracy across the three groups is shown in Figure S.1.2. The table group is more variable, but the means are overall very similar across the three groups, suggesting the experience of doing the table task or the logic puzzles did not lead to better decisions in the throwing task.

### Chart, box and whisker chart Description automatically generated

Figure S.1.2. Boxplots showing throwing accuracy for each of the three groups of participants.

### 3. Experiment S2: Repeated-measures effect of the table task on throwing task decisions

*Anca Popescu, University of Aberdeen*

**Summary**

After failing to find a clear effect of priming in Experiments 1 and S1, this experiment provides a final test of the hypothesis that performing optimal decisions on the table task would facilitate optimal decisions in the throwing task. There were two sessions of the throwing decision task: one before, and one after, the table task. This allows for a within-subjects comparison before and after the intervention, as well as comparison to a control group (who did a sodoku puzzle instead of the table task). In addition, the full version of the throwing task from Clarke and Hunt (2016) was used, including two sessions to measure each participant’s throwing ability over a range of distances, before the first session and after the second. Participants tended towards a pattern of better decisions in this experiment and differed less from optimal accuracy, possibly because of practice and because the sample was not completely naive to the logic of the experiment. Nonetheless, they were still sub-optimal, and there was no difference between the primed and control groups in the second session in either standing position or accuracy, replicating the pattern observed in Experiment 2 and S1.

**Methods**

*Participants.* Twenty undergraduate students at the University of Aberdeen gave informed consent to participate (11 female, mean age, 22.8).

*Procedure.* Participants carried out this experiment over two sessions, approximately one week apart. The first session had two parts, described below: accuracy measurement, and decision trials. In the second session, participants were randomly assigned to two groups. Half the participants carried out the table task, while the other (control) half were given a Sudoku to complete. 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.

*Accuracy measurement.* In the first session, 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.

*Table and control tasks.* The table task was the same as that used in Experiment 1. The control task was a Sodoku puzzle, which participants worked on for 5 minutes.

**Results**

All participants except one successfully managed to complete the table task. Standing positions for the *throwing* task, for each participant, are shown in Figure S.2.1. More participants modified their standing position appropriately with distances than in the previous experiments in this paper, and also compared to the results from Clarke and Hunt (2016), James, Clarke and Hunt (2017), and James, Hunt & Clarke (2023). This could partly be due to practice effects, but given that all of these studies sampled mostly from the same population (psychology undergraduates at the University of Aberdeen), it is also likely that participants in the current sample were not completely naive. We have been careful not to recruit the same participants to related studies, but over the preceding three years, we ran many versions of the study with members of the student population, who have consequently been exposed to this decision dilemma and were told the solution during post-experiment debriefing. The studies we have run vary in their surface features (e.g. throwing, table task, visual detection, memorization, pointing at lights), but pose the same dilemma, and because of the novel methods and general failure to implement a solution that seems simple in retrospect, discussion of the experiment amongst the students is likely to occur more than is typically the case. In other words, our local population of undergraduates, as a group, is increasingly non-naive. Given that several members of our control group already appear quite close to optimal, without any priming or training, the results of this experiment need to be interpreted with caution. Nonetheless, it is clear from the data presented in the Figure that they are consistent with the results from Experiment 2 in the main paper, and in Experiment S1 above, in showing that priming participants with the table task did not increase optimal choices in the throwing task.

Chart, scatter chart

Description automatically generated

Figure S.2.1. Top row scatterplots show the actual accuracy relative to an estimate of optimal accuracy in the first and second session of supplementary Experiment 2. The primed group did the table task between sessions. The bottom row line plot shows the average standing position across hoop separations during the second session. There is no difference between groups on either measure.

### 4. Additional plots and analyses

Chart, line chart

Description automatically generated

Figure S.4.1. Performance curves from the sensitivity mapping phase for the 5 non-naive participants in Experiment A1. Chance is 0.50. The value of Delta where accuracy is 0.75 is the switch point calculated for each individual for the choice task. For separations below this value, they should fixate the center box, and for separations above this value, they should fixate one of the two side boxes.

Chart, line chart

Description automatically generated

Figure S.4.2. Performance curves from the throwing accuracy session for the 4 non-naive participants in Experiment A2. The value of Delta where accuracy is 0.5 is the switch point calculated for each individual for the choice task. For separations below this value, they should stand midway between the two hoops, and for separations above this value, they should stand near one of the hoops.

Diagram, engineering drawing

Description automatically generated

Figure S.4.3. Performance curves from the acuity mapping phase for participants in Experiment 2 in the main text. Chance is 0.50. The value of Delta where accuracy is 0.75 is the switch point calculated for each individual for the choice task. For separations below this value, they should fixate the center box, and for separations above this value, they should fixate one of the two side boxes. These switch points are plotted in Figure S.4.2 below.

A picture containing shoji, crossword puzzle, sitting, building

Description automatically generated

Figure S.4.4. Proportion of fixations to the side boxes for each participant in Experiment 2. The red crosses show Session 1 performance and the blue show Session 2. The top two rows are participants in the Primed group, and the bottom two rows are control. The blue lines depict optimal performance based on each participant’s sensitivity shown in Figure S.4.1 above.

**Bayesian model of the results of Experiment 2.**

We used a Bayesian generalised linear mixed effect model to compare the effects of practise and instruction on accuracy between groups and over values of delta. The advantage of using such an approach is that it allows us to estimate the effect size, and the associated uncertainty. Modelling was done using rStan v2.17.3 (Carpenter et al., 2017) and the rethinking package v1.59 (McElreath, 2016) for R 3.4.0. A logit link was used and we accounted for differences between individuals by allowing intercepts to vary. We included three fixed effects: group (control/instructed); block (in the first block, one group was instructed where to fixate, while in the second block, neither group were given instruction); and delta (the separation between the centre and side boxes). The model that allowed for all two-way interactions and a three-way interaction between these predictors outperformed the other models we considered (using WAIC) so we report results from that model.

To visualise the model, we computed the predicted accuracy for the average participant, and recoded our variables to give: baseline (the control group in the first block), instructed (the instructed group in the first block), practice (the control group in the second block) and transfer (the instructed group in the second block). The posterior distributions are shown in Figure S.4.3. We can see there is some evidence for an effect of learning for the small separations (the transfer condition outperforms the effect of practise), but this does not hold over for the large separations. A plausible interpretation of this result is that instructions biased people to fixate the center more often than the sides; this bias promotes near-optimal performance for the close separations, to the detriment of performance at the far separations. Alternatively, there may have simply been more participants with a central bias in our primed group to begin with. Regardless, the results are consistent with Experiment 1 in suggesting people do not readily recognize and apply an optimal solution to this decision problem. There is very little evidence to suggest experience with implementing the optimal solution led participants to continue to implement it when making subsequent choices in the same decision context.

Chart, line chart

Description automatically generated

Figure S.4.5. Posterior distributions showing the predicted accuracy for a hypothetical average participant. Instruction improves accuracy across all separations. The “transfer” distribution estimates performance after this instruction is taken away. It is slightly better than practice at the smaller separations, but this difference is not observed at the farthest separations.

References

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. & 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*.*

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. (in press). Six of one, half dozen of the other: Suboptimal prioritizing for equal and unequal alternatives*.* *Memory and Cognition.*

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.

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.