Failure to use probability of success in deciding whether to pursue one goal or two.

Warren R. G. James (r01wj16@abdn.ac.uk)
School of Psychology, The Univerity of Aberdeen

Alasdair D. F. Clarke (a.clarke@essex.ac.uk)
Department of Psychology, The University of Essex

Amelia R. Hunt (a.hunt@abdn.ac.uk)
School of Pscyhology, The University of Aberdeen

Abstract

Difficult tasks should be attempted one at a time, while easy tasks can be undertaken in parallel. Reinforcing our previous conclusion that people are surprisingly poor at applying this logic, we find people fail to select standing positions that maximize their probability of success in throwing a beanbag into one of two possible hoops. We asked participants to explicitly report their odds of successfully throwing a beanbag into each hoop from the location they had chosen to stand, and estimates were highly accurate. Nonetheless, participants failed to use estimates of success appropriately to maximize success, suggesting a failure of insight, rather than limited or inaccurate information, can account for suboptimal decisions about standing position.

Keywords: Bounded Rationality; Optimal Behaviour; Awareness; Decision Making.

Introduction

Human skill is limited, and effective decisions must take these limitations into account. In Chess and Go, for example, it is impossible to select the optimal move by mentally simulating every possibility. An effective strategy must take into account the constraints of ones own memory capacity. Simon (1990) used the term bounded rationality to describe decisions that are rational given known constraints.

We recently reported a surprising failure to make effective decisions about whether to pursue one goal or two (Clarke and Hunt, 2016). In one experiment from that study, participants had to throw a beanbag into one of two hoops. The distance between the hoops varied and participants were told which one of the two hoops had been randomly selected to be the target only after they chose a place to stand. The optimal strategy when the hoops are relatively close together is to choose a standing position equidistant from both hoops, making an accurate throw possible irrespective of which hoop is the target. However, as the hoops move further apart, the probability of a successful throw from the center position drops below 50%. Now the best strategy is to stand close to one of the two hoops and hope it is the target on that trial. Despite the availability of this simple strategy, the distance between the hoops had no systematic effect on where people stood, demonstrating a profound failure to optimize throwing accuracy. The same failure to adjust strategy in response to difficulty was observed in deciding where to fixate to detect one of two targets (see also Morvan and Maloney, 2012) and in allocating attention when trying to memorize digit strings. These experiments took into account each individuals performance limitations by measuring throwing ability, visual acuity, and memory capacity in a separate session involving only single targets with no decisions. This allows for an individualized estimate of when a given participant should switch from attempting both goals to prioritising one. Nonetheless, bounded rationality could explain this failure as participants could be unaware of, or incorrect about, their own abilities (Schraw and Dennison, 1994).

In the current experiment, we repeated the beanbag-throwing task but explicitly asked participants to report on their expected accuracy before each throw. This allows direct comparison of estimated and actual throwing performance. Three outcomes are possible, all of which are informative about the cause of suboptimal decisions in this task. First, consistent with bounded rationality, participants may be unaware of their throwing ability and therefore fail to account for it when selecting standing positions. Second, participants may not adequately attend to their own abilities. Self-talk can improve throwing (Chang et al., 2014), suggesting explicitly reporting about ones own ability can improve task performance. Third, participants may have an accurate representation of their own skill, and drawing their attention to this information may not influence their standing position decisions. In this case, the results would suggest failure to maximize success in this paradigm is an example of a more fundamental limitation in decision competence.

Methods

Participants

Twenty-four participants (17 female) were recruited via the SONA systems at the University of Aberdeen and took part in return for course credit. The sample size of 24 (12 in each group) was based on our previous experiment (Clarke and Hunt, 2016). Average age was 19 years (SD = 1.5).

Procedure

We used a similar protocol to that of Experiment 2 in Clarke and Hunt (2016). The experiment was carried out over two sessions, with participants carrying out Session 2 a week after Session 1. Both sessions took place in the same sheltered paved area. The paving slabs were used as a convenient unit for measuring distances as they were approximately the same size as the hoops (the slabs measured 0.46 x 0.61m and the hoops had a diameter of 0.4m).

The aim of Session 1 was to measure how well each participant could throw a bean bag into hoops placed at seven evenly-spaced distances from three (1.38m) to fifteen (6.90m) slabs away, in two different directions (direction was counter balanced). A total of 84 throwing trials were completed in this session (12 beanbags for each of the 7 target distances). The data gathered from the first session were used to model how participants accuracy decreases as the distance to the target hoop increases (see Figure 1).

In Session 2, participants were again throwing beanbags into hoops, but there were now two hoops, and either one could be the target. Participants were asked to choose a place to stand before throwing, and were not told which hoop would be their target until after they had made this decision. To avoid having to re-position hoops from trial to trial, the area contained six hoops of three different colours, with blue hoops at the furthest distance, yellow at an intermediate distance, and red at the closest distance. The actual distance of each hoop colour depended on performance in the first session, with each colour corresponding to an estimate of the participants throwing accuracy from the centre: Blue hoops were placed at the slab distance where accuracy was expected to be closest to 10%, yellow hoops at 50%, and red hoops at 90% (see Figure 2). The beanbag on each trial was randomly drawn from a bag, which initially contained nine beanbags, three of each colour. The colour of the beanbag drawn from the bag determined which hoop pair was the target on that trial. Once all nine had been thrown they were replaced. This process was repeated 5 times for a total of 45 trials.

On each trial in Session 2, participants retrieved a beanbag and then chose somewhere to stand (the participants were told that they could stand anywhere on the paved area). The experimenter then informed them about which hoop (north or south) was their throwing target on that trial. The direction was randomised, with each direction equally likely on every trial. It was made clear to the participants that the direction of each throw was random and had been predetermined, with each direction being equally likely. They then performed the throw and their standing position and throwing accuracy were recorded.

Figure 1: This shows the set up for session 1 from the participant's point of view.

All participants followed the protocol described above, and half of the participants were also instructed to give an estimate (in percentages) of their expected throwing accuracy for both hoops from the location they had chosen to stand. They were prompted to provide this estimate after they chose a place to stand but before they actually threw the beanbag. This group will be referred to as the Online Estimation Group. This condition was included to test whether drawing the participants attention to this component of the problem they were faced with would help them to perform better.

Upon completion of the 45 trials in Session 2, all participants then performed a task similar to Session 1, but instead of throwing, they were required to give an estimate of their accuracy, in percentages, for each distance that had been tested in the first session. Participants stood in one spot, and one hoop was moved to different distances from them in either direction. The distances were split into two sets: 3, 7, 11, 15 and 5, 9, 13. One set was presented to the participants first, in ascending order (i.e. getting further from the participant); then the other set was presented to them in a descending order (i.e. getting closer). The order of sets was counterbalanced across participants so each set was presented first equally often. Results were analysed in R (R Core Team,

Figure 2: The setup for session 2

2016) and modelled using the lmer function from the lme4 package (Bates et al., 2014).

Results

Actual accuracy vs. Estimated accuracy

Participants throwing performance in Session 1 is shown in Figure 3. The relationship between accuracy and distance for each participant was modelled using logistic regression. Participants estimate of their own throwing ability is superimposed in blue. The majority of our participants were accurate in their ability to estimate their own throwing ability. This can be summarised by looking at the correlation between actual and estimated accuracy for each individual. This gives a median Pearsons correlation coefficient r of 0.89 (min=0.72, max=0.96).

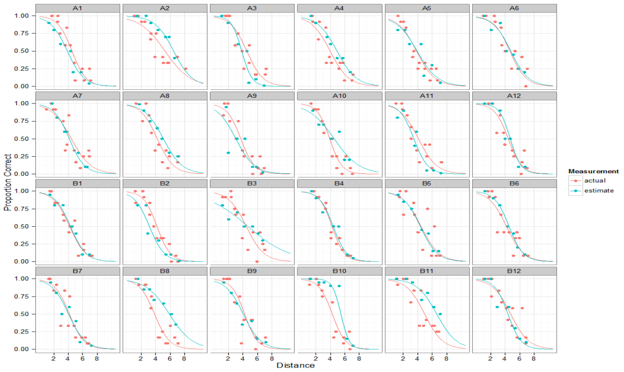

Figure 3: These graphs illustrate the accuracy (proportion correct) over the various distances for each participant (actual, in red), and their estimates of their own accuracy over the same distances (in blue).

Standing position

The optimal strategy for the closest hoop distance in Session 2 would be to stand in the middle, as the expected accuracy is 90% regardless of which hoop was selected as the target. For the farthest hoop distance, the optimal strategy would be to stand next to one of the two blue hoops, as this means that they would be approximately 100% accurate for that hoop and 0% accurate for the other hoop. Considering that each hoop was equally likely to be selected, this means that standing next to one blue hoop gives the participant a 50% chance of success, which is much greater than the 10% accuracy they would achieve by simply standing in the middle. The point where participants should switch between a centre and a side strategy is marked by a blue line in Figure 3 (the 50% point in their reported estimate of accuracy is shown by the red line). For the majority of the participants, it makes no difference whether we use their actual accuracy or their estimated accuracy to determine the ideal switch point, given the resolution of our experiment. The black dots illustrate the chosen standing positions on each trial: it is clear from these results that participants do not switch their strategy at either point; in fact, generally speaking, participants do not alter their standing position systematically with the distance of the hoops. These data are similar to those from the throwing task reported in Clarke and Hunt (2016). Interestingly, we can see that one participant (participant B11 in Figure 4) approaches the optimal strategy, particularly with respect to their estimated accuracy. Taken in aggregate, however, participants did not tend to stand closer to the hoop when they were further apart (a paired samples t-test comparing standing position for farthest to the closest hoop distance was non-significant; t(23)= -.49).

Standing positions across groups

To explore whether being asked to estimate the probability of successfully completing both possible throws had an effect on participants decisions, we compared the standing positions for the control group (Group A) to the online estimate group (Group B). For the closest hoops, participants in Group A stood on average 0.13 (standard deviation = 0.21) of the way from the central point to one of the two hoops, and participants in Group B stood at 0.14 (standard deviation = 0.27). Similarly, there was little difference between the two groups for the blue (far) hoops: participants chose to stand slightly further away from the central point (Group A: M= 0.128, SD= 0.17, Group B: M= 0.2, SD= 0.22). We conclude that being prompted to verbally report estimated accuracy for each trial had no effect on the strategies participants used to complete the task (close; t(22) = .09, far; t(22) = .895).

Analysis was also carried out to examine whether the accuracy of a participant's estimate of their own ability was correlated with how closely they followed the optimal strategy. To do this, the r value for each participant (representing the accuracy of their estimate) was correlated with a normalised value for the average distance of the standing position from

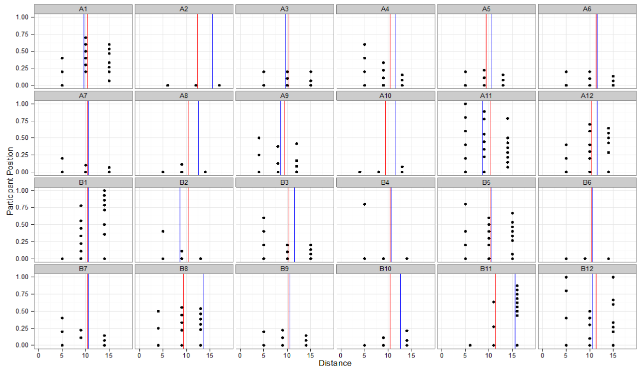

Figure 4: Black dots show the standing positions of each participant on each trial in Session 2 as a function of distance of the hoops from the centre (x-axis). Standing position was normalised so that 0 represents having stood at the centre and 1 being one of the side hoops . Red lines represent the distance at which participant should have switched between standing at 0 and standing at 1 based on their actual accuracy, and the blue lines is the switch point based on their own estimates of their accuracy.

the optimal position (in the 10% accuracy and 90% accuracy trials). The Pearson's correlation coefficients were very weak for both the 10% accuracy (r = -0.12) and the 90% accuracy conditions (r = -0.12). To clarify, had there been an effect of accuracy of estimates, there should have been a negative correlation for both accuracy conditions. The lack of a relationship suggests that participants with a greater awareness of their ability were not better at making using this information.

Discussion

Participants do not select standing positions that maximize their throwing performance, reinforcing the conclusion that participants fail to solve this task, as previously found by Clarke and Hunt (2016). This experiment was designed to test whether limitations in self-awareness of throwing ability could account for participants poor choices in standing positions. The results suggest participants were highly accurate in reporting on the expected outcome of their decision, but failed to make use of this information when deciding where to stand.

The second possibility we considered in the introduction is that participants have an accurate representation of their own skill, but fail to use this information in making their decision. If this is the case, asking participants to explicitly judge their expected accuracy before each throw should prompt them to make better decisions about where to stand. However, compared to the control group, the standing positions selected by participants who were explicitly asked about their accuracy before each throw were not more optimal. We asked participants only to state their expected accuracy for each hoop from the position they had selected; it may be that this was too indirect to cause participants to actually use this information in deciding where to stand. In future studies, it may be of interest to ask more probing questions. For example, asking them to explain why they choose to stand may lead to better decisions.

We conclude that participants fail to adopt an optimal strategy despite having highly accurate information about the expected outcome of their decisions. It is possible that the participants were not aware that this information was relevant to making a decision and therefore did not pay it due attention (Gegenfurtner et al., 2011) or were unable to use the information in an effective manner (Hardman and Cowan, 2016). It is also possible that participants were distracted by other, less relevant information. Gaissmaier and Schooler (2008) suggest that some people may engage in searching for a pattern that they may be able to exploit in order to form their decisions even when there is no pattern to the task. With this in mind, it may be of interest to investigate what information people deem to be relevant to a task they are performing. Finally, our sample did not widely vary in throwing ability and in self-awareness. A sample of participants with a wider range of throwing ability, and particularly including highly skilled throwers, may provide further insight into whether confidence in the relevant information can elicit optimal decisions for maximising accuracy.

We encounter situations with multiple goals and targets frequently in daily life, from deciding which locations to monitor while driving, to deciding how to invest time and resources in various projects. It is therefore surprising that our participants are so poor at making these decisions. In daily life these situations tend to be far more complex than the situation we have constructed here, and the expected outcomes of possible decisions would be similarly complex to calculate. Heuristics are simple rules of thumb to cope with decision making in complex environments. They can often lead to near-optimal behavior using less computation and information. For example, an effective rule for intercepting a high ball (e.g. in baseball) is to keep the ball at a fixed gaze position as you run towards it (Gigerenzer and Brighton, 2009). This simple heuristic allows a fielder to behave as if they had solved the differential equations that govern the balls movement. However, in our task the optimal solution can be fully described by a simple heuristic (i.e., *always stand in the center when the hoops are close, and switch to one hoop when they are far apart*). Our results suggest that participants are nonetheless unsystematic in their decisions.

Acknowledgements

This research was supported by a James S. McDonnell scholar award (A.R.H.).

References

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*.

Chang, Y.-K., Ho, L.-A., Lu, F. J.-H., Ou, C.-C., Song, T.-F., and Gill, D. L. (2014). Self-talk and softball performance: The role of self-talk nature, motor task characteristics, and self-efficacy in novice softball players. *Psychology of Sport and Exercise*, 15(1):139–145.

Clarke, A. D. F. and 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.

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

Gegenfurtner, A., Lehtinen, E., and Säljö, R. (2011). Expertise differences in the comprehension of visualizations: A meta-analysis of eye-tracking research in professional domains. *Educational Psychology Review*, 23(4):523–552.

Gigerenzer, G. and Brighton, H. (2009). Homo heuristicus: Why biased minds make better inferences. *Topics in Cognitive Science*, 1(1):107–143.

Hardman, K. O. and Cowan, N. (2016). Reasoning and memory: People make varied use of the information available in working memory. *Journal of experimental psychology. Learning, memory, and cognition*, 42(5):700–722.

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

R Core Team (2016). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Schraw, G. and Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary educational psychology*, 19(4):460–475.

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