Attraction Effect Absolute EVD

In Experiment 1 (Absolute EVD) we devised a perceptual representation of lotteries (using grids), and manipulated the difference in expected value between target and competitor prospects.

Participants

Forty-seven students and staff (14 male) from the University of Manchester volunteered to take part. Participants were between 21 and 53 years old (M=27). Informed consent was collected and participants were paid £7.00. The experiment took approximately 45 minutes to complete.

Design

We used a within-subjects design to test the effect of expected value difference on the dependent variable of preference reversal rate. For each participant preference reversal was measured as the rate that an alternative was chosen when it was the target minus the rate that it was chosen when it was the competitor. This measure is based on that used by Wedell (1991). Note that a rate could be calculated since each unique choice set was presented eight times.

The expected value difference independent variable had four levels (zero, three, six and nine). In order to achieve four levels of expected value difference, eight prospects were chosen. Four (called the V prospects) were given a fixed probability of 0.2 and varying values to achieve the expected values 8, 11, 14 and 17. The other four prospects (the P prospects) were given a fixed value of 25 and varying probabilities to achieve the same expected values as the V prospects, see Table 1. Each V prospect had a higher value, but lower probability than each P prospect. Each of the V prospects was paired with each of the P prospects to create 16 choice sets. Prospect pairs with the same expected value had an expected value difference of zero, while, for instance, a prospect with expected value eight paired with a prospect with expected value 17 had an expected value difference of nine.

Table 1: Expected value difference between prospect pairs.

	Low probability prospects (V)			
High Prob prospects (P)	.2(40)	.2(55)	.2(70)	.2(85)
.68(25)	9	6	3	0
.56(25)	6	3	0	3
.44(25)	3	0	3	6
.32(25)	0	3	6	9

Each of the choice sets was presented once with prospect P as the target and once with prospect V as the target resulting in 32 unique choice sets. Each of these unique choice sets was repeated eight times creating a total of 256 trials per participant. The experiment was divided into four blocks of 64 trials. The presentation of the trials was completely randomised with the exception that two trials could not be consecutive if the only difference between them was the position of the decoy. This control was added to ensure that the experimental manipulation was not too obvious to the participants. The decoy position defined whether prospect V or prospect P was the target. The decoys were always of the same value as their target but with 10% lower probability.

Stimuli

Participants were asked to choose between three prospects. Each prospect had a probability p of winning a value v in the form p(v) or (1-p)(0). The probability of each prospect was presented to participants using a grid consisting of 100 squares (see Figure 1). For a prospect with a success probability of 0.6, 60 of the squares were shaded green. For a success probability of 0.4, 40 of the squares were shaded green, and so on. The position of the green squares within the grid was randomised.

The value of each prospect was presented in an identical 10 by 10 grid immediately below the probability information. For the value grid, the display was not randomised, and was shaded red. Forty red squares indicated a value of 40, 70 red squares a value of 70, and so on. There were six grids in total, one probability and one value grid for each of the three prospects.

Procedure

At the start of each trial all of the displays were blank. Participants revealed the probability and value information for each prospect by holding the mouse button down in a blue shaded area between the probability and value displays. Lifting off the mouse button cleared the display. Participants could only view one prospect at a time, but could choose any order to view them in. This prevented participants from assessing alternative prospects by making a perceptual comparison of the visual density of the different displays, encouraging them instead to encode the information.

In each trial, participants were given five seconds of viewing time that they could distribute between each of the prospects as they saw fit. A five-second countdown timer was displayed on the left of the interface. The timer only counted down while the participants had the mouse button depressed, and the probability and value information were visible. Participants could only make their choice once the timer had reached zero. An enforced assessment duration allowed us to control the amount of effort participants put into each trial.

Participants were asked to choose the prospect they preferred by clicking a button marked `Select' that was positioned below the value grid for each of the prospects. Participants did not receive any feedback on their decisions.

Figure 1. The stimuli used in Experiment 1. The density of green squares (top row) represents the probability of the prospect. The number of red squares (bottom row) represents the value of the prospect. Each column represents an alternative prospect. The probability and value of only a single prospect are displayed at any one time. The participant chooses a prospect by pressing the corresponding 'select' button.

Analysis

As noted, in order to determine the preference reversal rate, we subtracted the rate that P prospects were chosen when they were the target, from the rate that they were chosen when they were the competitor. This metric was averaged across the prospect pairs for each level of expected value difference. A positive difference indicated that participants preferred a prospect more often when it was the target than when it was the competitor; this is the outcome that is expected when measuring the attraction effect. A negative difference indicated that participants preferred a prospect more often when it was the competitor than when it was the target. No difference indicated that participants were consistent, and chose a prospect the same number of times regardless of whether it was the target or competitor. As was the case in the Wedell (1991) design we coded each decoy selection 0.5 in favour of the target and 0.5 in favour of the competitor.

Attraction Effect Relative EVD

Methods

We designed Experiment 2 (Relative EVD) to be a replication of Experiment 1 using a relative, rather than absolute measure of expected value difference. We also extended the design to include two further stimuli presentation types. This allowed us to replicate our findings and show that they extend to other paradigms that have been used to elicit the attraction effect. See stimuli section for more details.

Participants

One hundred and forty-three undergraduate Psychology students from the University of Manchester volunteered to take part, fifty in Experiment 2a, 52 in Experiment 2b and 41 in Experiment 2c. Participants received course credits for participating.

Design

In each trial participants chose between a target, competitor and decoy prospect. Our independent variable was the difference in expected value between the target and competitor prospects. Expected value difference had four levels 0, 20, 100 and 300%, reflecting the percentage increase in expected value from the smaller to the larger expected value prospect. Sixteen prospect pairs (See table 2) were created spanning probability and value space, four for each of the IV levels.

Each prospect pair was presented to the participants eight times with one prospect as target, and eight times with the other prospect as target. The decoys had either 20% fewer value points, or 20% fewer probability points than the target. The experiment consisted of 256 trials (4 IV levels x 4 prospect pairs x 2 decoy positions x 8 repetitions) which were presented in random order. The dependent variable of preference reversal rate was calculated in the same way as in Experiment 1. See the Experiment 1 analysis section for details.

Table 2: Stimuli values used in Experiment 2

	Alternative Prospect			
Reference Prospect	0%	20%	100%	300%
.12(83)	.24(42)	.29(42)	.48(42)	.96(42)
.17(59)	.24(42)	.29(42)	.48(42)	.96(42)
.59(17)	.42(24)	.42(29)	.42(48)	.42(96)
.83(12)	.42(24)	.42(29)	.42(48)	.42(96)

Stimuli

Experiment 2a, was very similar to Experiment 1, conveying the probabilities and values via the same grid display. The stimuli differed in that the probability grid was not randomised, there was no time limit and participants could view all three prospects simultaneously.

Experiment 2b presented the stimuli in a manner more similar to the original Wedell (1991) design, sentences were displayed on screen to convey the objective probabilities and values of each prospect (see the bottom panel in Figure 2. Prospects were presented simultaneously in sentence form `p probability of v points'.

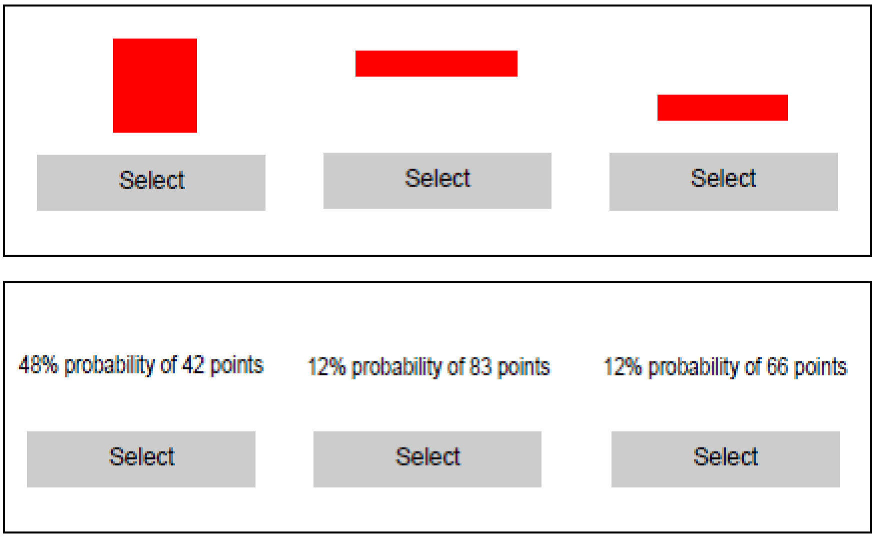

Figure 2. Stimuli used in Experiment 2. The bottom panel shows the descriptive stimuli from Experiment 2b, while the top panel shows the rectangle stimuli in Experiment 2c.

In Experiment 2c we modelled our stimuli on those used by Trueblood et al (2013) using rectangle area instead of expected value (see the top panel in Figure 2). Participants were presented with three rectangles and asked to choose the rectangle with the largest area. The height and width of the rectangles in pixels were the same as probabilities and values used for the prospects in 2a and 2b. Since both the area of a rectangle and expected value of our prospects are given by the product of their attributes, it was simple to substitute expected value difference for area difference.

Procedure

For all three experiments, participants were presented with three prospects or rectangles simultaneously on a computer monitor. Below each option a button marked `select' allowed participants to indicate which prospect they preferred or rectangle they perceived to have the largest area. The on-screen order of the three stimuli in each trial was randomised. 256 trials were presented in four blocks of 64 trials with an enforced one-minute break between blocks.

Attraction Effect Pointing

Method

Our method of creating motor lotteries that are mathematically equivalent to explicitly described lotteries, is adapted from Wu et al. (2009); Wu, Delgado, and Maloney (2011). This involves measuring a subject's variance over repeated attempts to hit a target. This variance is used to determine the width of target necessary to achieve a chosen probability that it will be hit. Subjects can then be offered choices between targets of differing widths equivalent to offering them choices between lotteries of varying probability.

Subjects

Sixty (eight male) undergraduate subjects with a mean age of 20 (SD = 2) were recruited from the University of Manchester. Subjects received course credit for taking part in the experiment, and attended for one session of approximately 40 minutes.

Materials

A 19 inch touch-screen at a resolution of 1280 by 1024 pixels was used throughout the experiment. In the choice phase subjects responded by pressing 1, 2 or 3 on the numeric keypad of a standard Windows keyboard. The experiment was created in the Python programming language and run on a Microsoft Windows 7 PC.

Design

We used a 3 context (target, neutral competitor) x 2 paradigm (motor, traditional), entirely within subjects design. The attraction effect predicts that the same alternative will be chosen more often when it is a target than when it is neutral or a competitor. To test this, we constructed a safe lottery (high probability) which had a value of £20 and a success probability of 70%. This was offered alongside a risky lottery (low probability) which had a value of £75 and a success probability of 30% (see Figure 3). This pair of lotteries was presented nine times with each as the target and another nine times with no decoy present (neutral). This resulted in 27 trials per paradigm. This design allowed us to calculate a choice rate for each lottery in each context i.e., the proportion of times it was chosen for each placement of the decoy. The two paradigm levels (motor and traditional) differed only in the way that the probability of the lottery was displayed. In the traditional paradigm, the value and probability of a lottery were displayed on-screen in numerical format (see Figure 3). In the motor paradigm the probability of the lottery corresponded to varying widths of targets (see Figure 4). In both paradigms the hypothetical amount to be won was displayed in the form '£20'.

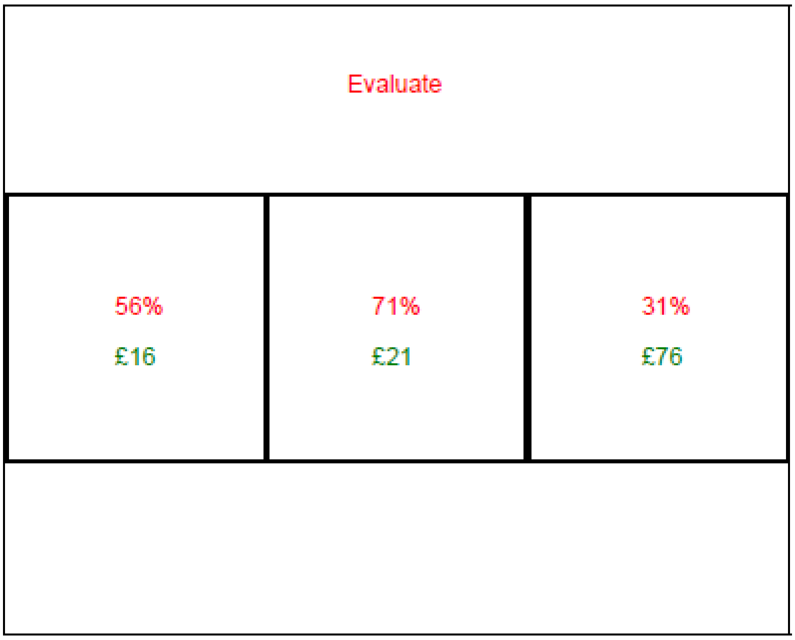

Figure 3: Example stimulus in the traditional condition. Subjects indicated which lottery they would prefer to play.

As in previous studies with rapid pointing tasks (review in Trommershauser et al., 2008) a time limit meant that subjects could not be sure of hitting the target. Our subjects completed a training phase prior to the choice phase in which we recorded their hit points to recover a distribution of their accuracy around a target center. This distribution of subjects' hit points was then used to create individualized stimuli in the choice phase. This allowed us to control the probability of success for each subject by varying the width of the targets they were presented with.

A pilot study was conducted to determine a value that would make the risky lottery subjectively equivalent to the safe lottery. Decoys were always .15 less likely to win and of £5 less value than their dominating lottery.

Procedure

Training phase

In the training phase subjects learned their own motor noise in a rapid pointing task. Subjects were instructed to touch a green bar on the left hand side of the screen with the index finger of their dominant hand, after which they had 500 msec to touch a yellow target bar on the right hand side of the screen with the same finger. The target zone was 20 pixels in width and 1025 pixels to right of the start bar. The start bar was 50 pixels wide and covered the full height of the display (1024 pixels), as did the target zone.

Figure 4: Example stimulus in the motor condition. Subjects were told to indicate which of the three lotteries they would prefer. The width of the targets (black bars) was manipulated to achieve probabilities identical to those used in the traditional paradigm.

Subjects completed 100 training trials as described above. Subjects were not informed that they would be making decisions between different target widths in the subsequent choice phase. If subjects successfully hit the target within the time limit the word ‘Hit’ would appear in green. If they were within the time limit but missed, the word ‘Miss’ was displayed in orange. If they exceeded the time limit, the message ‘Too slow’ was displayed in red and a warning sound was played.

Choice phase

In the choice phase subjects saw three lotteries presented on the screen and a message stating ‘Evaluate’ above them (see figure 4). Subjects were instructed to press the space bar when they were ready to indicate their choice. The message would then change to ‘Choose’, and one of three lotteries would disappear. Subjects had to choose between the remaining two lotteries by pressing the appropriate number on the keyboard. In the majority of trials the decoy was removed when the space bar was pressed. However, in some trials the decoy remained and either the target or competitor lotteries were removed. These trials were excluded from the analysis but were included in the stimuli to encourage subjects to evaluate all the lotteries. Subjects indicated their preferred lottery using the numeric keypad, 1 for the left most lottery in the display, 2 for the middle lottery and 3 for the right-most lottery. The 54 trials (27 motor and 27 traditional) were presented in random order such that the motor and traditional trials were mixed together. The position of each alternative on the screen was also randomized.

Subjects only had two seconds to make their response after pressing the space bar. If they exceeded this time limit, ‘Too slow’ was displayed and the trial was discarded from the analysis. The process of evaluating, removing an option, and choosing under a time constraint was adapted from Soltani et al. (2012) with the purpose of forcing subjects to take all three alternatives into consideration when making their choice.

Rather than present identical choices nine times, the values and probabilities were jittered such that values were either £19, £20 or, £21 and probabilities were either .69, .70 or .71. This resulted in nine safe lotteries with a mean value of .7(£20). The same procedure was applied to the risky lottery and the decoy lotteries. In the choice phase subjects were asked to indicate which lottery they would prefer, but they did not go on to play the lottery, nor receive any other type of feedback.

Attraction Effect Loss Domain MTurk

In Experiment 1 (MTurk) we adapted Wedell's (1991) demonstration of the attraction effect in choices between prospects. The original study involved choices between prospects exclusively in the gain domain where participants chose between alternatives that all had an expected value of approximately $10. The alternatives were described in terms of probability of success and monetary value of success. These two attributes (value and probability) were traded off to produce attraction effect choice sets.

Methods

Participants

Sixty participants (42 male) were recruited using the Amazon mechanical turk platform. Participants had a mean age of 33 (SD = 11). Informed consent was collected and participants were paid $0.50. The experiment took approximately 10 minutes to complete.

Design

Within subjects, our independent variable of domain had two levels, loss and gain. Our dependent variable was preference reversal rate. This rate is a measure of the attraction effect based on that used by Wedell (1991). The rate an alternative is chosen when it is the competitor is subtracted from the rate the same option is chosen when it is the target. The attraction effect is present when an alternative is chosen more often when it is in the target configuration than in the competitor configuration. The same alternative is presented several times in each configuration, hence it was possible to calculate a rate for each.

In each of the gain and loss domains we presented participants with 10 pairs of prospects, one high probability (safe) prospect with a value ranging from $12 to $25, and one low probability prospect (risky) with values ranging from $15 to $33 (see Figure 5). All prospect pairs had an expected value of approximately $10. In the loss domain the gambles were identical except that the value amounts were negative.

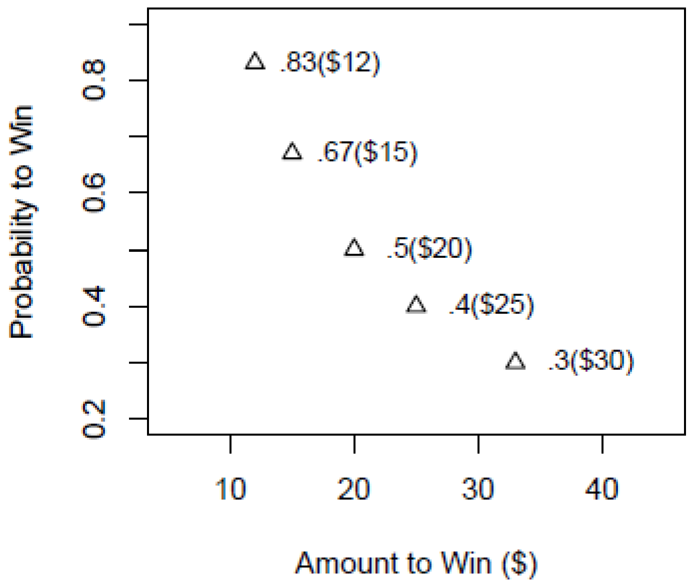

Figure 5. Prospect values used in Wedell (1991) and in Experiment 1. Each prospect was paired with all others creating a total of 10 choice pairs.

In both domains we used range decoys which was one Wedell's (1991) between subjects conditions. Range decoys have previously been shown to elicit the largest attraction effect size Trueblood et al (2013) and consequently provided the strictest test of our hypothesis that the effect would be reduced in the loss domain. In the gain domain this meant that the safe prospect decoys had a $2 smaller value, and the risky prospects had a 5\% smaller probability. In the loss domain, the safe prospect decoys had a 5\% greater probability and the risky prospect decoy had a $2 greater value.

In each domain the 10 safe and risky prospect pairs were presented twice. Once with the safe prospect as the target and once with the risky prospect as the target. The decoy position defined whether a prospect was the target or competitor.

Stimuli

In the gain domain, three prospects were presented simultaneously (target, competitor and decoy) in sentence form: `83\% probability of winning $12'. Each choice set was prefixed with the question: `Which option do you prefer?'. In the loss domain the sentences took the form `83% probability of losing $12'. The word `losing' was coloured red to indicate that domain had changed.

Procedure

Participants answered 40 questions in total. Within each domain the order that the choice sets appeared in was randomised. The loss and gain domains were presented as separate blocks. Participants experienced each choice set as a separate web page. On each page they indicated their choice by pressing the appropriate radio button. They then pressed a continue button to advance to the next page. A page counter kept participants informed of their progress through the experiment.

Attraction Effect Loss Domain Lab

In Experiment 2 (Lab) we extended a more recent demonstration of the attraction effect in choices between prospects Soltani et al (2012). Whilst the Wedell (1991) experiment equated the expected value of the alternatives on offer, it should be noted the risk preferences are likely to mean that participants did not in fact find these alternative subjectively equivalent. The Soltani et al (2012) study resolved this issue by having a preliminary session in which each participant's point of subjective equality was found for a fixed high probability prospect.

Methods

Participants

Thirty undergraduate students (3 male) from the University of Manchester volunteered to take part. Participants had a mean age of 21 (SD = 1.8). Informed consent was collected and participants received course credit for taking part. The experiment took approximately 45 minutes to complete

Materials

The experiment stimuli were presented on a 17 inch display at a resolution of 1280 x 1024 pixels. The experiment was programmed in Python and run on Windows 7. Participants responded using the numeric keypad of a standard Windows keyboard.

Design

As per Experiment 1, we used a within subjects design, testing an independent variable of domain (loss and gain). Our dependent variable was preference reversal rate. The dependent variable was measured in exactly the same way as in Experiment 1.

In each of the gain and loss domains we presented participants with two prospects, one high probability (safe) prospect with a value of either £20 or -£20, and one low probability prospect (risky) that had a 30% probability of winning an amount that each participant found subjectively equivalent to the safe prospect.

We determined a point of subjective equality for each participant using a two-alternative forced-choice paradigm. We paired a fixed safe prospect with a risky prospect and systematically varied the value of the risky prospect. This method is adapted from Soltani et al (2012). The value varied from £20 to £120 in ten steps of £10. Each step was repeated eight times. This allowed us to fit a logistic curve to the data and recover the indifference point such that the participant would be equally likely to choose either prospect. This approach is commonly used in studies of perception to recover perceptual points of subjective equality (PSE) of stimuli. The recovered PSE was then used for the risky prospect in the main study. This process was carried out for both the gain and loss domains. In each domain the safe and risky prospects were presented nine times with the safe prospect as the target and nine times with the risky prospect as the target. The decoy position defined whether a prospect was the target or competitor.

Stimuli

Participants were presented with prospects on a computer monitor. The screen was divided vertically in to three panels. Each panel showed a probability expressed as a percentage and a value prefixed by a £ sign. In the gain domain the values were coloured green, whereas in the loss domain the values were coloured red and prefixed by a minus sign. During the evaluation phase of each trial the word Evaluate appeared in white at the top of the screen. During the choice phase of each trial the word `Choose' appeared in yellow.

Procedure

The experiment started with a session to determine each participant's PSE. Participants were presented with the safe prospect and a risky prospect with a random value until all the risky prospects had been presented eight times. Participants could observe the prospects for as long as they liked. When they were ready to indicate their choice they pressed the space bar and then had two seconds to respond by pressing the appropriate number on the numeric keypad (1 for left panel, 2 for middle and 3 for right panel). As per Soltani et al (2012) a logistic curve was then fitted to the data in order to calculate the PSE.

In the choice phase participants were instructed that after evaluating the prospects they would press the space bar and one of the prospects would be removed at random. They then had two seconds to choose from the remaining prospects. In the majority of trials, the decoy was removed and participants chose between the target and competitor. However, in some dummy trials the decoy remained. These trials were discarded but forced the participants to consider all three alternatives on each and every trial since they could not know which prospect would become unavailable.

In both phases of the experiment the risky and safe prospects were slightly jittered such that participants were not presented with same choice twice. For each prospect the value could vary by £ 1 and the probability by 1%. There were 36 trials of interest in the choice phase (2 domains x 2 decoy position x 9 repetitions). These were supplemented with a further 18 dummy trials. The trials were presented in a random order but were blocked by domain.

Random Generation Experiment 1

Participants first observed blocks of binary outcome random sequences following an unbiased Bernoulli process (p = 0.5) and were then instructed to generate random outputs to match the properties of the observed process.

Method

Participants.

Twelve undergraduate students from the University of Manchester participated on a voluntary basis and gave informed consent. Participants received course credit as payment. There were no exclusion criteria.

Materials.

Participants were seated in front of a 19-inch LCD display. The experimental stimuli were presented using the Python programming language on a PC running Windows 7. Participants responded using a standard Windows keyboard.

Design.

We compared the statistical properties of sequences generated by a truly random Bernoulli process (p = 0.5) and those generated by our participants (N = 12) using four methods contrasting:

i) the expected frequency of sub-sequence occurrences per block of length 20

ii) the proportion of blocks of length 20 on which there was at least one sub-sequence occurrence. We call this the *occurrence rate* which is the complement of the *non-occurrence probability* described by Hahn & Warren (2009)

iii) occurrence frequency histograms for three subsequences of interest – perfect runs (e.g. 0000), perfect alternations (e.g. 0101) and sequences such as (0001) which when compared to a perfect run of the same length has implications for the gambler's fallacy phenomenon

iv) boxplots illustrating medians and IQRs of occurrence frequency distributions for the three sub-sequences outlined in iii)

Procedure.

Participants were told they would first observe the output of a machine generating a random sequence of 1's and 0's, and that they should attend to it (Presentation Phase) before going on to generate a sequence (Generation Phase).

Presentation Phase: Each digit (a 1 or 0) appeared on the screen for 250 msec before being replaced by the next digit in the sequence. The display of each digit was accompanied by a corresponding tone. The display was full screen with a black background. The digits were displayed in white in 80 point Arial font in the centre of the screen. 1's were accompanied by a 1200 hertz tone, and 0's by an 800 hertz tone. After every 20 digits the sequence paused and participants were required to complete a distractor task. The distractor task consisted of counting the number of vowels in a list of 10 words. In total participants observed 600 digits over 30 blocks of length 20.

Generation Phase: Participants were asked to generate a new sequence representative of the one they had just observed in the Presentation Phase. They used the keyboard to press either 1 with their left hand, or 0 with their right hand. For each key press participants saw the appropriate digit on screen and heard the corresponding tone, exactly as in the presentation phase. As in the Presentation Phase, participants generated 600 digits in 30 blocks of 20 and the same distractor task was used in between each block.

Analysis

We counted sub-sequences using sliding windows of lengths *k* = 3 to *k* = 9 and for global sequence length *n* = 20. For illustration we describe the analysis and present results for *k* = 4. For each participant, and each of the 30 blocks of data collected, we slid a window of length *k* = 4 through the *n* = 20 outcomes generated. We then undertook 4 analyses of these sequences:

Analysis 1 - Over the 360 (12 observers x 30 blocks) length 20 sequences, we calculated the expected value of the participant-generated frequency distribution for each of the 16 possible sub-sequences (0000, 0001,…,1111). For an unbiased random process the expected frequency of each sub-sequence should be 1.0625 per sequence of length 20.

Analysis 2 - Over the 360 (12 observers x 30 blocks) length 20 sequences, we calculated the occurrence rate – i.e. the proportion that contained at least one occurrence for each of the 16 possible sub-sequences (0000, 0001, …, 1111). Even for a random process this metric will not be the same for all subsequences since non-occurrence probabilities vary due to the sliding window analysis (see Hahn & Warren, 2009).

Analysis 3 - Over the 360 (12 observers x 30 blocks) length 20 sequences, we generated histograms illustrating the proportion of the 360 sequences containing 0, 1, 2, etc… occurrences of the three sub-sequences 0000, 0101, 0001.

Analysis 4 - Over the 360 (12 observers x 30 blocks) length 20 sequences, we generated boxplots illustrating the median and IQR of the distributions obtained in Analysis 3. We generated the same amount of simulated data as that obtained from human participants using an unbiased Bernoulli process (*p* = 0.5). We refer to these simulated sequences as theoretical participant-generated and their properties are analyzed in an identical manner to the human data. By repeatedly generating (N = 1000) theoretical participant data sets we were able to place confidence bounds on the metrics described in Analysis 1 and 2 for the theoretical participant.

Random Generation and Judgement - Chunking

Methods:

Participants

One hundred and eighteen people volunteered to take part from the University of Manchester student and staff population. Participants gave informed consent, and received course credit or £7.50 as reimbursement for their time. Participants' mean age was 21, *SD* = 4.8, 77 % of participants were female. There were no exclusion criteria.

Materials:

Participants were seated in front of a 19" monitor at a 1280 x 1024 resolution. Participants made responses using a standard Windows keyboard.

Design

We tested for the effect of recent exposure to a genuine .5 Bernoulli process on a person's ability to both generate and judge a random sequence. Within subjects, participants generated and judged a sequence both before and after being exposed to the Bernoulli process. Between subjects we manipulated the nature of the experience by chunking it into different sized blocks. There were three levels of this chunk size IV (100,10,5). Overall this resulted in a two within (experience: pre,post) x three between (chunk size: 100,10,5) design.

All participants were presented with the same global sequence of length 200. The sequence was generated by a Bernoulli process, but was checked to ensure that it had a representative alternation rate of approximately .5. The nature of the experience differed only in terms of the size of chunk that the global sequence was divided into, 2 x100, 20 x 10 or 40 x 5.

Procedure

The experiment consisted of generating and then judging random sequences, this was conducted both before and after participants observed a genuine random sequence.

Generation: Participants were asked to generate a series of coin tosses by pressing 1 for heads and 0 for tails on the computer keyboard. They were instructed to produce a sequence that they thought would be representative of flipping a fair coin. Participants were encouraged to type at a speed of roughly 1 press per second, and they could see an 'H' or 'T' appear on the screen. Each display of H or T replaced the previous display so participants could not see the history of their sequence. Participants were instructed to generate a sequence of length 200 in 2 blocks of 100. In between each block the screen would display the message ' sequence 2 of 2:' and required the participant to acknowledge it by pressing 'c' to continue.

Judgement: The judgement task was adapted from Falk & Konold (1997) and consisted of 10 sequences of heads and tails. Each sequence was of length 21 and varied in alternation rate from 0.1 to 1. The participants' task was to rate each sequence according to how likely it was to have been produced by a fair coin. Participants were required to start by giving scores of 0 and 10 to the least and most likely sequences respectively. They were then free to score the other sequences relative to these and they could use the same score more than once.

Observation: Participants were told that some people find the production of random sequences difficult and that they would now be presented with a genuine random sequence generated by the computer. They were instructed simply to attend to the presentation. An 'H' or 'T' was presented with a SOA of 700msec. Each presentation replaced the previous one so that the history of the sequence was not visible. All participants experienced the same sequence except that in different conditions the number of breaks changed (1,19 or 39). After each break the participant was required simply to press the letter 'c' to continue the presentation.

Participants then repeated the generation and judgement tasks mimicking a fair coin. The post judgement task used the same alternation rates as the pre judgement task, but used different instantiations of the sequences.

At the end of the experiment participants completed a questionnaire designed to elicit their beliefs about the gambler's fallacy. This included asking people how they would bet after a sequence of five heads in a row, and whether after five heads: heads was most likely, tails was most likely, or both were equally likely.

Analysis

In this section we detail how our dependent variables were derived from the data.

Generation Analysis: Participants were asked to generate a sequence representative of a fair coin by pressing 1 for heads and 0 for tails on a computer keyboard. In all conditions they produced a sequence of length 200 in two blocks of 100. These data were then measured for alternation rate: the number of switches between H and T as a proportion of 198 possible. This generation data was also measured for the ratio of sequences containing a run of length 4 to those containing a run of length 3 with an alternation (e.g. HHHH : HHHT).

Judgement Analysis: Based on a task devised by Falk & Konold (1997) participants judged sequences of varying alternation rates according to how likely they found it that the sequence was produced by a fair coin. We fit a quadratic curve through each participant's data to recover the alternation rate they judged most random.

References

Falk, R., & Konold, C. (1997). Making sense of randomness: Implicit encoding as a basis for judgment. *Psychological Review*, *104*(2), 301.

Hahn, U., &Warren, P. a. (2009). Perceptions of randomness: why three heads are better than four. *Psychological Review*, 116, 454–61.

Soltani, A., De Martino, B., & Camerer, C. (2012). A range-normalization model of context-dependent choice: A new model and evidence. *PLoS Computational Biology*, 8, e1002607.

Trommershauser, J., Maloney, L. T., & Landy, M. S. (2008). Decision making, movement planning and statistical decision theory. *Trends in cognitive sciences*, 12, 291–297.

Trueblood, J. S., Brown, S. D., Heathcote, A., & Busemeyer, J. R. (2013). Not just for consumers: Context effects are fundamental to decision making. *Psychological Science*, 24, 901–908.

Wedell, D. H. (1991). Distinguishing among models of contextually induced preference reversals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17, 767–778.

Wu, S.-W., Delgado, M. R., & Maloney, L. T. (2009). Economic decision-making compared with an equivalent motor task. *Proceedings of the National Academy of Sciences*, 106, 6088–6093.

Wu, S.-W., Delgado, M. R., & Maloney, L. T. (2011). The neural correlates of subjective utility of monetary outcome and probability weight in economic and in motor decision under risk. *The Journal of Neuroscience*, 31, 8822–8831.