ARE YOU SURE IT IS RANDOM? –

MEASURING PEOPLE'S BELIEFS OF

GENERATED RANDOMNESS

By

András Lajos Sári

Submitted to

Central European University

Department of Economics

In partial fulfillment of the requirements for the degree of Economics MA

Supervisor: Marc Kaufmann

Abstract

In this thesis, my aim is to measure people's beliefs about how easy it is to identify their made up random series as human made. I achieve this by having them participate in a game, where they generate 8-long coin toss series of heads and tails. Their main goal is to make up sequences that my algorithm classifies as a true series of coin tosses upon comparing it with a true Bernoulli series. After generating the series the participants are offered a deal to exchange their game for a lottery, thus revealing their beliefs. I ran the experiment on 30 people whom I randomly assigned into 5 treatment groups representing 5 different offers. In the experiment I identify people generated series with over 60% accuracy. I find that people display a wide range of beliefs from seriously underestimating to seriously overestimating their chance of success, overestimation being the more common effect. In conclusion people are particularly bad at predicting how easy it is to identify them.

Acknowledgements

I would like to thank my supervisor Marc Kaufmann for the consultations and advices he provided me throughout the thesis writing process. I also want to thank Annie Liang, Jon Kleinberg and Sendhil Mullainathan for sharing their dataset with me, as this thesis could not have been completed without it.

Table of contents

Abstract 1

Acknowledgements 2

Table of contents 3

List of Figures, Tables or Illustrations 4

Introduction 1

Chapter 1 – Experimental design 4

Chapter 2 – Data on subject-created random sequences 8

Chapter 3 – Participants' confidence in their ability for generating random sequences 11

Chapter 4 – Conclusion 17

References 19

Appendices 20

List of Figures, Tables or Illustrations

Figure 1: Number of people who accepted the offer in each treatment group 11

Figure 2: Estimated number of people accepting the offer if faced with a given winning chance, based on survey 12

Figure 3: Introduction page 20

Figure 4: Generation page 21

Figure 5: Offer page 22

Figure 6: Accepted screen 22

Figure 7: Declined page 22

Figure 8: End survey 23

Figure 9: Winning page 24

Figure 10: Losing page 24

Figure 11: Cognitive test 25

Introduction

In today's world we are surrounded by decisions that we need to base on probabilistic outcomes. Choices from simple recreational activities, ranging from card or dice games to betting on the outcome of a football match to complicated decisions in fields of investments and insurance that can have a huge impact on our lives. Though choices with uncertain outcomes have always surrounded us, it is a relatively new phenomenon that we can make informed decisions based on calculated risk, rather than on our gut feeling. As we have to make more and more decisions based on risks, I believe it is essential to understand the main driving forces that influence our behavior regarding those decisions.

My main goal in this thesis is to uncover people's beliefs about how well they generate randomness. In order to identify their beliefs I conduct and run an experiment. I ask participants to create several 8-long pseudo-random coin-toss series, without using an actual coin. Then they face a choice: if they choose it, they can let an algorithm compare their series to an actual random one. In this case they get the payoff only if the computer deemed their series less likely to be made by a human than the random one. The other choice is simple: they can choose to have a lottery instead which has the same payoff options as the previous choice, just with a fixed probability, which depends on the treatment group. From this choice I can have a simple approximation of the person's sophistication. I built the evaluating algorithm using the dataset of Kleinberg, Liang and Mullainathan (2017) which contained 21975 people generated 8-long coin toss series.

It is a long established fact that people tend to make mistakes when it comes to assessing probabilities (Bar-Hillel & Wagenaar, 1991; Kahneman & Tversky, 1973). These mistakes can come in many forms: one can believe that after a series of the same outcome the chance of the next outcome to be different is increasing, (Gambler's fallacy) or lower than it actually is. (Hot hand fallacy) Xu and Harvey (2014) found evidence that these two fallacies may go hand-in-hand, one causing the other. People also commit mistakes regarding unlikely events: they tend to overestimate their probability. Benjamin, Rabin and Raymond (2016) explore this error further by observing the biased connection between sample size and the probability of tail events. (Non-Belief in the Law of Large Numbers)

Bar-Hillel et al. (1991) categorizes the behavioral experiments regarding randomness into two categories: perception and generation tasks. In perception tasks the main focus is on the people's perception of the outcomes of random variables, while in the generation tasks researchers concentrate on the pseudo-random variables participants generate. While both of these categories are well-researched, there is relatively little research going in the direction of exploring people's awareness of their mistakes. My aim in this thesis is to find evidence pointing towards a degree of sophistication in people's biases concerning generation tasks.

As people are prone to generation biases, these biases influence their behavior in a multitude of situations, one example being how they execute mixed strategies. Wang, Xu and Zhou (2014) show that rather than playing the classic mixed strategy Nash equilibrium in rock-paper-scissors games people tend to play cyclic strategies. Though it is widely accepted that these strategies have a behavioral root, I believe that players' belief could be a key factor determining these strategies. A better understanding could not just improve our strategies in a multitude of games, but also help game developers to gain a deeper insight into how their games' metagame will evolve.

Chen, Moskowitz and Shue (2016) find evidence that decision makers are more likely to make an incorrect decision, if they encounter a long series of the same decisions, due to gambler's fallacy. Though their research shows that these errors tend to decrease with more experience among other factors, it is still relatively unknown how deeply decision makers are aware of their biases. Understanding sophistication in this situation could help in developing better policies, as so far there is relatively little known how people change if their degree of sophistication changes.

Though my results are prone to experimenter bias, the difference in behavior between how people conduct themselves in an experiment in contrast with real life, I believe that this difference should not be a major issue in my case, as there is neither a moral dilemma, nor is it easy for the participants to infer which choices are the ones they “should” take.

Due to the small sample size my results should be interpreted with caution. Overall I find evidence, that participants’ beliefs range from underestimation to serious overestimation of their capability of producing random series, overestimation being the more pronounced effect among most of them. I also encountered some behavioral effects distorting my measurements ex post; therefore I propose some possible modifications for the experiment.

In the following section I will describe the experimental design, detailing the different parameters of the experiment. Then I will introduce the data I collected and the results I have drown from them, and finally I will conclude.

Chapter 1 – Experimental design

In this section I describe the details of my experimental setting.

I ran my experiment using o-tree (D.L. Chen, Schonger, & Wickens, 2016) in order to run it smoothly and make all the processes involved automated. I ran six sessions each consisting of five people. Though one session would have been more ideal, due to technological difficulties [1] I had to change my experiment slightly in order to accommodate these developments. I could bypass these difficulties by creating my own network using my mobile phone as a WI-FI hotspot and making participants join on the network on their smartphones and participate in the experiment using their devices.

Running the experiment in multiple sessions has some drawbacks. Firstly had I run my experiment in one session I could be sure that information does not propagate between the sessions (as there is only one). In order minimize information diffusion I asked participants not to mention any information regarding my experiment to each other during the duration of my experiment and I made sure that participants do not leave the experiment with the knowledge of which of their series won and which of their series lost, therefore participants in the following sessions may not get armed with information regarding winning strategies from previous participants.

Overall one session's duration was about 15-20 mins. All participants who took part in the experiment received a 500 HUF worth of general purpose voucher redeemable at any Aldi shop. To motivate the participants for conscious choice, all participants who won their game

[1] I got a soft MAC-based ban from the CEU network upon running my o-tree server

are eligible to win a 5000 HUF worth of general purpose Spar voucher. The winner for the voucher was drawn randomly from the IDs of the winners. [2]

I ran the experiment in two versions: English and Hungarian in order to lower the barrier of eligibility and to diversify the participants, even though there is a non-negligible selection-bias as most participants are university students.

All screens of the experiment can be found in the appendix. The experiment starts with an introduction page detailing the information players need to know regarding the experiment. Their main task is to make up eight-long series of coin toss results. Each of their series is then compared to a proper random series by an algorithm. The algorithm's task is to decide which series was made by a human. A player wins a round if the algorithm determines their coin tosses as the randomly generated and the randomly generated as the human-generated. There are overall eight rounds and a player wins, if he manages to win four or more rounds. The result of the computer analysis is only revealed at the end of the experiment, so there is no learning involved between the rounds.

After they proceed they see a page where they can generate their 8-long pseudo-random coin toss series. There are eight pages of this, so participants generate a total of 64 coin tosses. There is no default option and, due to the fact that the experiment was conducted on mobile devices, there is no fast fill option available; participants may not go through the boxes using the tabulator button as they could if there was a keyboard available. I believe generally this does not matter, but in my case this could be crucial as filling the same pages multiple times could lead to participants developing a muscle-memory-like effect and fill similar patterns not because of conscious choice rather than because of ease of input.

[2] Due to legal reasons, money cannot be awarded for experiments in Hungary, hence most experiments award vouchers

After they are finished with the coin-tosses they get an offer to a lottery to which they can exchange their initial deal, the analysis of their coin tosses. This lottery has x% chance to win and 100-x% chance to lose instantly. Each individual in a session gets a different x. This is my key variable as accepting (or rejecting) the offer reveals the chance a participant thinks she has to win the computer analysis is lower (or greater) than x. I ran my experiment the following values for x: 25%, 33%, 41%, 49%, and 57%. I chose these numbers after testing the game with a 5-person group and calculating the theoretical chance to win. In every session I assigned one person to each treatment group randomly.

I make sure that accepting or rejecting the offer has no effect on the information sets the participants end up in at the end of the game, that is they know the outcome of both the computer analysis and the lottery regardless their choice. (I reveal them this information during the page where they can decide about the offer.) I chose this in order to avoid participants rejecting the offer out of curiosity of the outcome of the computer analysis. After participants accept or reject their offer they end up on a page that confirms their choice and then they proceed to the page that reveals the outcome of the game.

I used the dataset on human-generated coin tosses which Kleinberg, Liang and Mullainathan (2017) used in their research.[3] This dataset contains 21975 eight-long pseudo-random coin toss series. I ranked all the series based on their occurrence. The algorithm chooses the less represented series as the randomly generated and the more represented as the human-generated one. I also calculated the empirical probability a person loses a round using the following formula:

$$\sum_{i=0}^{255}(\frac{\#occurence}{21975} * (255.5 - i)) \ ,$$

[3] Special thanks to the authors for sharing the dataset with me.

where i is the index of the 8-bit code in the decreasing order based on occurrence and *#occurence* is the number of times the series appeared in the 21975 sample. The formula is easy to understand if one views it as the chance a person writes a code times the probability he loses with that code, summed up for all possible eight-long codes, where I approximate the true probability with the relative frequency. Based on this calculation my estimate for a person losing a round is 0.605. Using the binomial distribution, my estimate for a person losing the game is 0.609[4]. In this calculation I assume that the people who produced the dataset are not fundamentally different from the people participating in my experiment regarding coin toss generation. There is some evidence pointing towards these distributions being more general that one would generally assume (Kleinberg et al., 2017).

Participants also need to fill out two pages worth of forms toward the end of the experiment. The first with general demographic information; they submit their age, gender, years of education, highest degree earned and the answers to some questions regarding their interaction with probabilities in their everyday life. The second page imposes a series of cognitive tests that aim to measure their understanding of probabilities and their cognitive capabilities.

[4] Using the following formula to calculate the chance to lose: $\sum_{i=5}^{8}\binom{8}{i} * 0.605^{i} * (1-0.605)^{8-i} = 0.609$

Chapter 2 – Data on subject-created random sequences

In this section I will describe the data collected during my experiment.

I recruited my participants mostly by spreading the news of my experiment through my social network. Unfortunately, my 30 participants do not form a representative sample of the society. Even if I could reach everyone, one could also argue that the reward offered and time the experiment required already introduce a selection bias, as a student is far more likely to participate with the above mentioned conditions than an executive at a large firm. Although this could limit the scope of my results, I believe one could still gain a lot of insight from this slightly distorted sample.

Table 1: Summary statistics

	Summary statistics
Average of rounds won	3.13
Hungarians	12
International students	18
Male	15
Female	15
Average cognitive test score	3.33
Number of offers accepted overall	8
Average chance to win a round[5]	39.05%
Number of rounds won	94
People winning 4 or more	12
Number of winners of the game	11

Table 1 displays the summary statistics about my sample. Out of the 30 people who participated in my experiment, 26 were still students. On one hand students form a fairly monolithic subgroup of society, but on the other hand my group of students is fairly diverse, as there are both M.A., B.A., M.Sc. and B.Sc. students from four different universities in my

[5] This was calculated by averaging the chance of winning of each series

sample studying a wide array of fields ranging from ones that are fairly mathematics oriented like electrical/mechanical engineering and economics to ones that are less involved with stochastics like law. There were 15 males and 15 females in my sample. I tried to have balanced groups regarding to the two language versions I ran the experiment with; in the end I managed to have 12 Hungarians and 18 international students who used the English version.

While my ex ante prediction of an individual losing one round was 0.605, out of the overall 240 rounds played 94 was won by a player, which implies that the empirical estimate of a player losing a round is 0.608. This leads to the conclusion that the Kleinberg Liang and Mullainathan data is a fairly good approximation to how people generally randomize, regardless of the input method, or the demographic difference between their sample group and mine.

Overall 11 people won the game out of the 30. If there was not offer made 12 people would have won. Out of the 8 people who accepted the offer, it changed the outcome of the game on 3 occasions: two people lost instead of winning, and one person won instead of losing.

Table 2: Covariate balance data

Treatment	Average of rounds won	Hungarian	Male	Female	Average cognitive score (0-5)	Offer accepted	Average score in a round (0-255)
25%	2.83	3	2	4	3.16	1	98.45
33%	2.66	3	2	4	3	2	95.16
41%	4.83	2	3	3	3.5	2	119.64
49%	2.5	2	4	2	3.33	2	84.625
57%	2.83	2	4	2	3.67	1	99.52

Table 2 shows the covariate balance data among the treatment groups. Due to the small sample size and the number of treatment groups, it is a natural consequence that the treatment groups are not entirely balanced, but the randomization distributed most attributes fairly evenly among the treatment groups.

My main variable of interest is whether the participant accepted the offer according to his treatment. An offer of 41% means, that he could have exchanged the initial game to a game that wins in 41 times out of 100 on average. Out of the 30 people 8 participants accepted their respective offer.

I also collected data on the time each individual spent generating an 8-long series. The average time spent on a generating page is 26.6 seconds. Collecting this data opens up the possibility to analyze if thinking time has an effect on the outcome of the round.

The chance of winning the game if one could perfectly make up random coin toss series is approximately 0.64, given that one never accepts the offer, as the chance of winning one round is 0.5.[6] When asking the participants the reasoning why they accepted or rejected the offer, many people argued that as they needed at least 4 out of 8 rounds to win they have an equal chance of winning and losing, if they generated randomly. I believe they reached this conclusion by simply observing that they need to win at least half of the rounds, so it must be a 50-50 game, even though it isn't as winning has a higher chance if one used a true randomization device. One could argue though, that if they underestimated their chance to win, they should have been more likely to accept offers, especially the 57% offer, as they thought their chance to win is lower than their actual chance to win based on their assumptions, given that they indeed generate randomly. Nevertheless, running the experiment with an odd number of rounds would reduce the chance of participants committing similar mistakes.

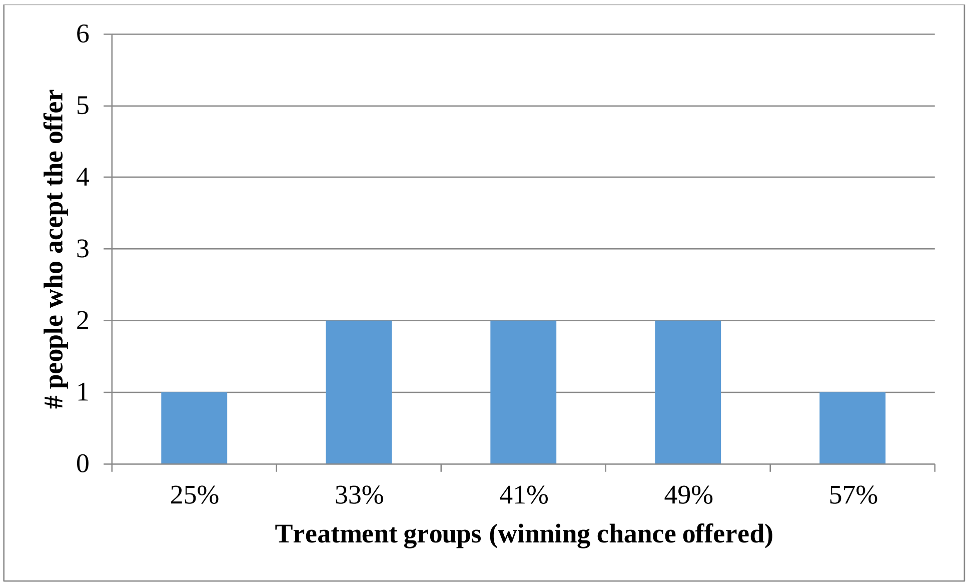
[6] The exact calculation is : $\sum_{i=0}^{5}\binom{8}{i} * 0.5^8 = 0.637$

Chapter 3 – Participants' confidence in their ability for generating random sequences

In this chapter I will detail my results and findings.

Firstly I want to note, that one has to take all my results with a grain of salt, as my sample size is small. I believe my findings should be interpreted as clues that point towards interesting phenomena in human behavior, rather than universal facts. It may be worthwhile to check whether my findings hold for larger sample sizes.

Figure 1: Number of people who accepted the offer in each treatment group

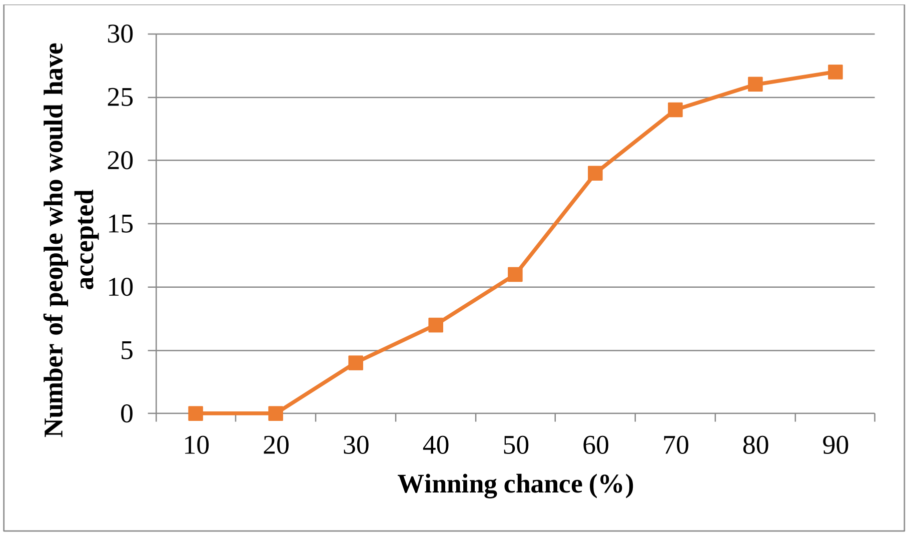


My first result is on the number of people who accepted the offer in each treatment group, which is shown in Figure 1. Before running the experiment I expected that fewer people should accept the offer the lower the winning chance is on the offer. Ex post my results show that the winning chance offered seems to have no effect on the person accepting or rejecting the offer, if one accepts that the treatment groups are balanced. In my opinion the right conclusion here is not that it has no effect, rather the effect it has is small enough in order to be unidentifiable on this sample size.

Based on both the ex ante and the ex post estimation of the chance of winning the game, people are better off accepting any offer has a greater chance of winning than 40%. Out of the 18 participants who faced an offer worth taking, only 5 took the offer, thus 13 people displayed overconfidence in their success. Out of the 6 people who got offered the best chance of winning (57%), only one accepted, so this shows there are at least 5 people in my sample who are seriously overconfident. Out of the 12 participants who faced an offer that was worse than the initial game 3 people accepted, therefore 3 people underestimated their chance of success in the game. It can be seen from these results, that there is both evidence for underestimating and overestimating the chance of winning the game, most participants tending towards overconfidence.

I believe rather than people being sophisticated about their generation bias, the main reason people accept or reject seems to lie in them being pessimistic or optimistic about themselves. I base this assumption on the answers I got in the end survey, as I asked participants to write down the reasoning behind their choices. While I got various answers, a recurring theme in many of the answers of those who rejected is that they believed in their generated numbers to be fairly well-generated.

Figure 2: Estimated number of people accepting the offer if faced with a given winning chance, based on survey



In the end survey, after finishing the game and accepting or rejecting the offer, I asked the participants to estimate the minimum winning chance of the offer they would have accepted. Figure 2 displays their answers. It can be interpreted as, if all the people who answered the survey question were offered a lottery which had x% chance to win, y number of people would accept, if they reported truthfully. The average value of people's estimation of their winning chance is 53%, while the median is 51%, both of which display overconfidence, as the true threshold assuming full sophistication should be around 40%.

Out of the 30 people 27 gave answers to this question. Out of the 27 people, 17 said they would have accepted the best (57%) offer. If only the people who rejected their offer and answered the question are counted, out of 20 people 10 would have accepted the best offer, even though out of 6 people only 1 accepted. I believe there might be an endowment type effect at work here, as a possible explanation is that participants view their generated series as their "own", while look at the offer as something foreign, and thus valuing their own higher than the foreign thing. Not facing the best offer most seem to report that they would have given up for a better than 50% deal, while when facing it people are reluctant to trade it away.

As the results of Figure 2 shows, one could conclude that the belief about the winning chance of the game shows a large amount of heterogeneity among the participants. While there are many people who are incredibly overconfident in their abilities of generating random series estimating their chance to win over 70%, there are also others who seriously underestimate their chances of winning, given that we assume that people generally perform similarly in this task. This result falls in line with the conclusion I drew from their responses to the offer.

If one wants to model people's beliefs about their generation biases, my results suggest that using a continuum of beliefs is a better approach than classifying people into a binary (for example sophisticated-naïve) sets in cases where this distinction may matter.

If the winning chance has little effect on the probability a person accepts the offer the natural question to ask is: what is the main determinant of a person accepting or rejecting an offer? Running an ordinary least squares regression shows, that the main determinant is the participant's gender, as shown in Table 3. If we control for the gender of the player we can see in Model 1 a slight increasing, though not significant, pattern in the chance of people accepting the offer the better the offer is, though the 57% treatment groups still seems to be an outlier.

In Model 2 I added the average score of player's generated series, which ranges between 0 and 255. Due to the design of the evaluation algorithm, a point increase in the score of a series increases the chance that the series is a winning one by 1/256. The sign of the average score is pointing towards the right direction, as scoring more on average makes people less likely to accept, but the effect is significant neither economically, nor statistically.

Table 3: Main determinants of accepting an offer

VARIABLES	(1) Model 1	(2) Model 2	(3) Model 3
33.playeroffer	0.167	0.162	0.161
	(0.246)	(0.250)	(0.251)
41.playeroffer	0.240	0.269	0.244
	(0.247)	(0.257)	(0.252)
49.playeroffer	0.313	0.286	0.305
	(0.252)	(0.261)	(0.257)
57.playeroffer	0.146	0.142	0.150
	(0.252)	(0.256)	(0.256)
Average_score		-0.00151	
		(0.00285)	
player_male	-0.439**	-0.421**	-0.399*
	(0.163)	(0.169)	(0.193)
Cognitive			-0.0336
			(0.0830)
Constant	0.313*	0.455	0.406
	(0.182)	(0.327)	(0.295)
Observations	30	30	30
R-squared	0.259	0.267	0.264

Lastly, I use the data I recorded for the cognitive score. Participants were asked five mathematics and probability related questions at the end of the experiment. They got one point for each correct answer; therefore the cognitive score is scaled from 0 to 5. In Model 3 I measure the effects of the Cognitive score on the chance to accept. While the result is not statistically significant it points towards the direction that more well-versed people in mathematics and probability tend to accept with a lower chance.

As I collected how much time each participant spent generating each 8-long series, I have the opportunity to look into how thinking more or less has an effect on the generated series. My independent variable is the time spent on a generating page in seconds, while my dependent variable is the score of the series, which still ranges between 0 and 255.

Table 4: thinking time's effect on performance

	(1)	(2)
VARIABLES	Model 4	Model 5
time	-0.0252	-0.0371
	(0.163)	(0.171)
Constant	100.2***	100.5***
	(6.297)	(7.324)
Observations	240	240
R-squared	0.000	
Number of person		30

Table 4 shows the results of my regression. In Model 4 one can see the simple OLS results, with only time included as the independent variable, while in Model 5 I controlled for participant specific fixed effects. Both results show that increasing time spent on a generating page has no effect on the quality of the series produced. This hints that the difference between the binomial distribution and the human-generated distribution is neither caused by overthinking nor caused by not thinking enough on the generated series. It seems more to be an intrinsic characteristic to our behavior.

Because I found that thinking time has no significant effect on the quality of the generated series, the only method I can suggest to those who are facing a real-life generation problem is to use a randomization device, for example toss a real coin.

Chapter 4 – Conclusion

Seeing the reasoning behind the participants' choices, I can conclude that running the experiment with odd number of rounds would probably decrease the chance of them miscalculating their probabilities to win the game, though I am not sure it would completely eliminate this problem.

While my method of revealing participants' beliefs through choice has the advantages of being easy to understand for the participants and easy to implement, ex post the experiment revealed some drawbacks as well. One of these drawbacks, as seen from my analysis, is that it requires a quite large sample size for proper measurement. Though this can be easily addressed, there is another issue of other behavioral effects distorting the measurement. It may be beneficial to try to run the experiment with the mechanism for eliciting probabilities from Karni (2009), though if one does so, I think one should be extra careful about the participants fully understanding how the experiment works. Though introducing a test round may be beneficial in this case, the participants learning the outcome of one round may be problematic, as it can cause them to update their beliefs, and thus it may bias the results of the measurement.

My sample of 30 people is quite small therefore all my results should be interpreted accordingly. It may be worthwhile to run the experiment on a larger scale if one wants a good measure of the proper distribution of beliefs among people.

I can also conclude that based on my data the main determinant of accepting or rejecting the offer seems to be the gender of the participant, as my results seem to show that women are significantly more likely to accept the offer than men regardless what they are offered, even though I have to add that this effect might be driven by an attribute that is correlated with gender not the person's gender itself.

Overall I believe all my analysis shows, that while people are fairly similar in their biases of generating random series, as they did generate series exactly as the Kleinberg, Liang and Mullainathan (2017) data predicted, they have fairly various beliefs about how good they are at this task. Their beliefs range from underestimation to massive overestimation of their skill in producing random series, most people tending towards being overconfident. Because of the high range of beliefs, I suggest modeling people regarding their beliefs about generated randomness with a continuum of beliefs rather than binary ones.

References

Bar-Hillel, M., & Wagenaar, W. A. (1991). The perception of randomness. *Advances in applied mathematics, 12*(4), 428-454.

Benjamin, D. J., Rabin, M., & Raymond, C. (2016). A model of nonbelief in the law of large numbers. *Journal of the European Economic Association, 14*(2), 515-544.

Chen, D. L., Moskowitz, T. J., & Shue, K. (2016). Decision making under the gambler's fallacy: Evidence from asylum judges, loan officers, and baseball umpires. *The Quarterly Journal of Economics, 131*(3), 1181-1242.

Chen, D. L., Schonger, M., & Wickens, C. (2016). oTree - An open-source platform for laboratory, online and field experiments. *Journal of Behavioral and Experimental Finance*.

Kahneman, D., & Tversky, A. (1973). The belief in the law of small numbers. *Psychological Bulletin*.

Karni, E. (2009). A mechanism for eliciting probabilities. *Econometrica*, *77*(2), 603-606.

Kleinberg, J., Liang, A., & Mullainathan, S. (2017). *The Theory is Predictive, but is it Complete?: An Application to Human Perception of Randomness.* Paper presented at the Proceedings of the 2017 ACM Conference on Economics and Computation.

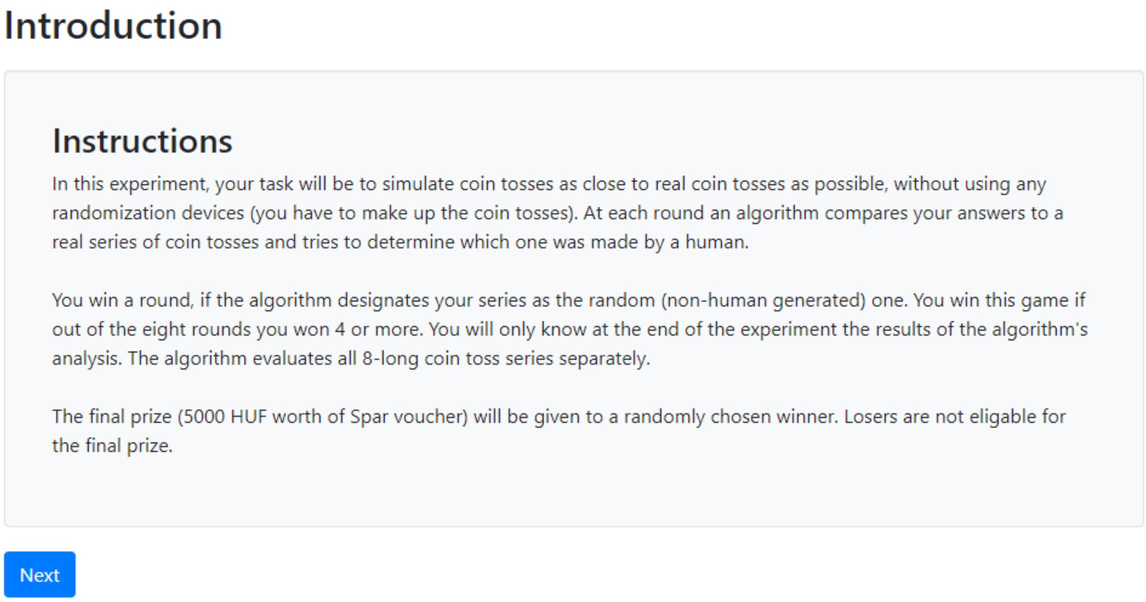
Wang, Z., Xu, B., & Zhou, H.-J. (2014). Social cycling and conditional responses in the Rock-Paper-Scissors game. *Scientific reports, 4*, 5830.

Xu, J., & Harvey, N. (2014). Carry on winning: The gamblers' fallacy creates hot hand effects in online gambling. *Cognition*, *131*(2), 173-180.

Appendices

In the following section I post the screens the participants go through during the experiment.

Figure 3: Introduction page



The introduction page is the first one the participants see. It describes the general information players need to know during the experiment. Figure 3 displays the introduction page. Figure 4 shows the first one of the generating pages; there were overall 8 of these. One can see at Figure 5 the offer page for the 25% treatment group. Clicking yes will direct the participant to the accepted page displayed at Figure 6. If the participant clicked No, she will be directed to the page showcased at Figure 7. After clicking next on their respective page, participants will reach the screen shown in Figure 8. After filling out the forms and clicking next, the participant will end up in one of the following screens shown in Figure 9 and Figure 10 based on the outcome of the game. Figure 11 shows the questions asked at the end as a cognitive test.

Figure 4: Generation page

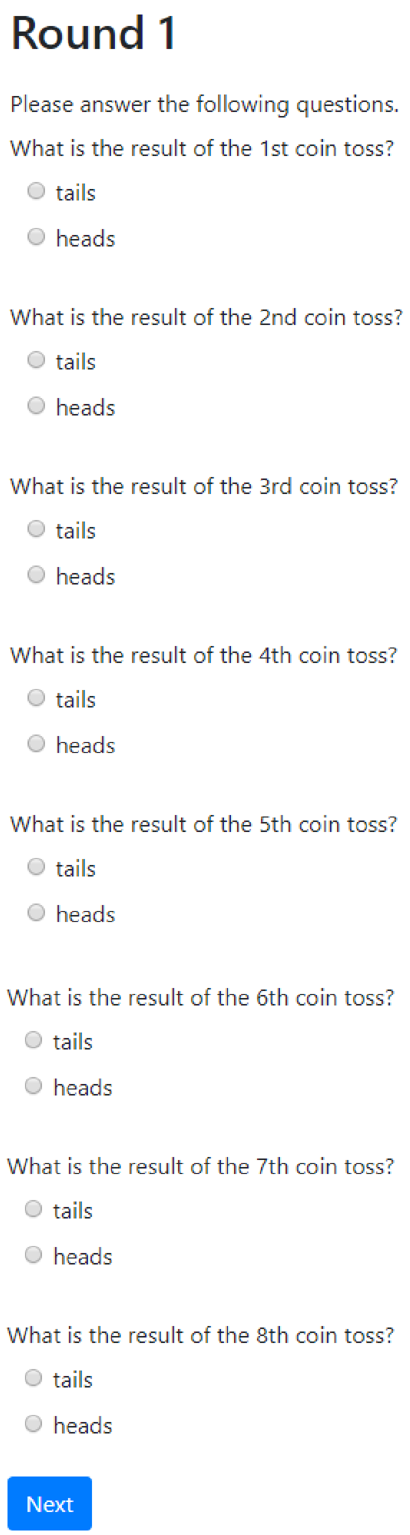


Figure 5: Offer page

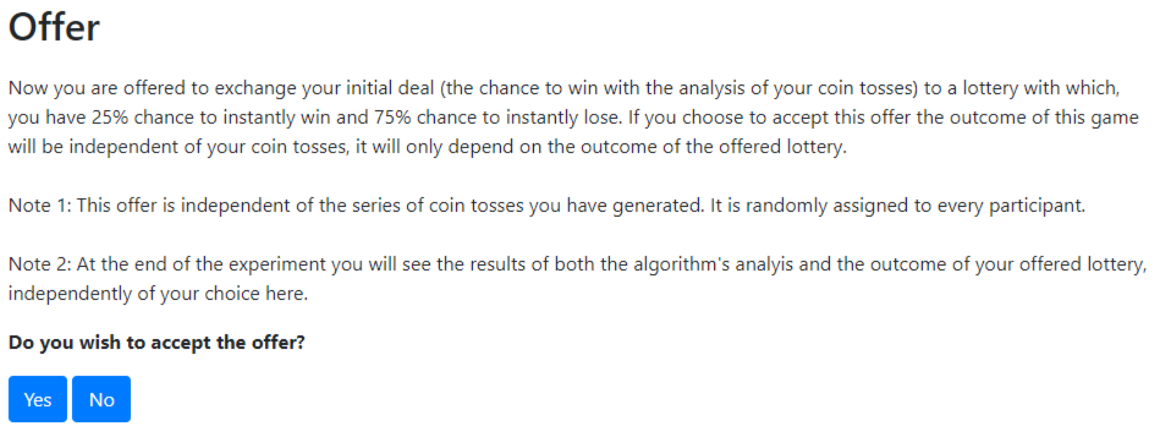


Figure 6: Accepted screen

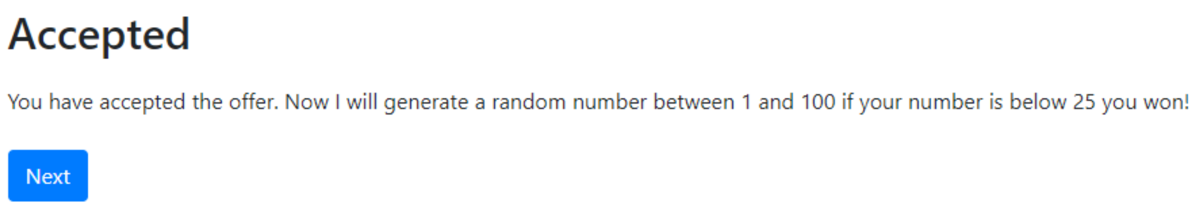


Figure 7: Declined page

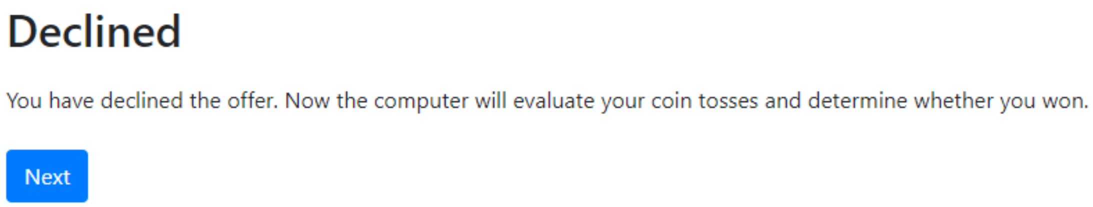


Figure 8: End survey

Survey

Please answer the following questions.

What is your age?

What is your gender?

- ○ Male
- ○ Female
- ○ Other

How many years of education have you had?

What is the highest level of school you have completed or the highest degree you have received?

How often do you bet?

- ○ Often (on average once a month or more often)
- ○ Rarely (on average once a year)
- ○ Never (less often than once a year)

How often do you play card, dice or other games that's core mechanics involve chance?

- ○ Often (on average once a month or more often)
- ○ Rarely (on average once a year)
- ○ Never (less often than once a year)

"Please explain why you accepted or rejected the offer below.":

"Please estimate what is the lowest winning probability, for which you would accept the offer":

Next

Figure 9: Winning page

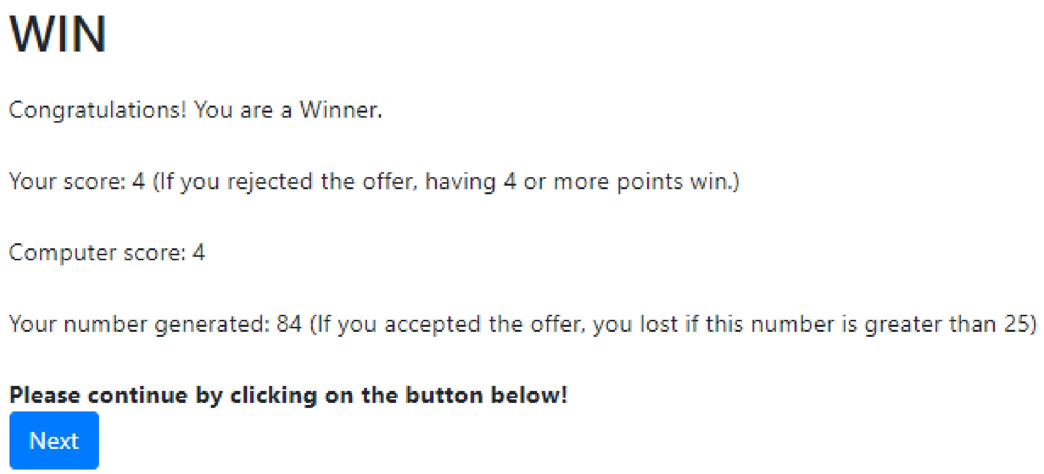


Figure 10: Losing page

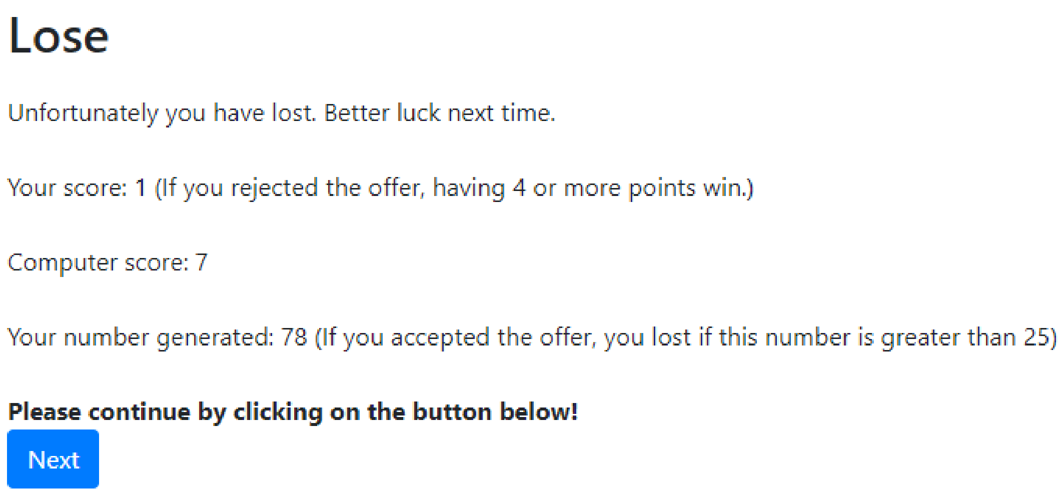


Figure 11: Cognitive test

Survey

Please answer the following questions.

A bat and a ball cost 22 dollars in total. The bat costs 20 dollars more than the ball. How many dollars does the ball cost?

"Which has the higher probability? A: wining the lottery with one try. B: having the same set of numbers drawn for a lottery as the last occasion.":

- A
- B
- They are equal

"What is the expected value of the sum of two fair (6-sided) dice rolls?" :

"Which has the higher probability? A: The sum of two fair (6-sided) dice rolls is 5 B: The sum of two fair (6-sided) dice rolls is 8":

- A
- B
- They are equal

"After flipping a fair coin 10 times you observe that all the 10 flips ended up heads. What is the chance of the the next coin flip ending up as heads?" :

Next