Pre-Registration

Analyzing the Impact of Choice Complexity on Personal Risk Choices

Maohua Nie

Version 1, edited October 4, 2023

1 Study Information

1. Title (required) Analyzing the Impact of Choice Complexity on Personal Risk Choices

2. Authors (required) Maohua Nie, Dr. Sebastian Olschewski, Prof. Dr. Jörg Rieskamp

3. Description (optional)

Individuals make numerous decisions on a daily basis, but often lack the ability to make optimal choices due to the overwhelming amount of information they must process. As cognitive misers, people tend to avoid complex options and prefer simpler and more intuitive alternatives, even if they are not necessarily superior. Previous research has examined the complexity structure of options and has suggested that it can significantly influence individuals' risk choices. Among these studies, compound lotteries are often used as a manipulation of complexity. However, it should be noted that most studies define complexity solely in terms of the number of alternatives or the number of features, and the predominant experimental design involves comparing individuals' preferences for choices with different numbers of potential outcomes. Nonetheless, complexity can also refer to the cognitive effort required to process information, even if the number of potential outcomes remains constant. In other words, options with higher complexity may demand more time and cognitive resources to explore their possible outcomes. Importantly, most of the research on complexity has focused on one type of complexity. In addition, cognitive models have hardly been used to explain people's risky choices. To fill this research gap, our study will employ a series of tasks with different operatinalizations of complexity to examine whether individuals' risk choices are affected by differences in complexity resulting from variations in cognitive burden. We will utilize a series of binary decision problems, such as compound lotteries, to manipulate the complexity involved in calculating each outcome and its respective probability. Such design ensures that we can test whether different forms of complexity can have similar effects on people's risky decisions. Furthermore, We will try to develop models in the framework of drift diffusion model, which has not been used widely in prior studies, to analyze the experimental data. The diffusion model is a model of a cognitive process involving a simple two-choice decision, it assumes that the decisions are made in a noisy process, in which people accumulate information over time from a starting point toward one of two options and begin to respond when one of these boundaries is reached. The model is well suited for fast binary choices. Thus, within the framework of the diffusion model, we believe that our analysis of risky choices will be more reliable and robust, and will allow us to learn about people's cognitive paradigms when faced with options of different complexity.

4. Hypotheses

Response bias: People prefer simple options than complex options.
First hypothesis: People have a natural aversion to the complex option.
Second hypothesis: People will discount the expected utility of complex option.
Third hypothesis: People have subjective interpretation of the probabilities.

2 Design Plan

5. Study type (required)

Example: Experiment - A researcher randomly assigns treatments to study subjects, this includes field or lab experiments. This is also known as an intervention experiment and includes randomized controlled trials.

6. Blinding (required)

For studies that involve human subjects, they will not know the treatment group to which they have been assigned. Personnel who interact directly with the study subjects (either human or non-human subjects) will not be aware of the assigned treatments. (Commonly known as "double blind")

7. Study design (required)

within-subject design,

8. Randomization (optional)

Example: We will use block randomization, where each participant will be randomly assigned to one of the four equally sized, predetermined blocks. The random number list used to create these four blocks will be created using the web applications available at http://random.org. More info: Typical randomization techniques include: simple, block, stratified, and adaptive covariate randomization. If randomization is required for the study, the method should be specified here, not simply the source of random numbers.

3 Sampling Plan

In this section we'll ask you to describe how you plan to collect samples, as well as the number of samples you plan to collect and your rationale for this decision. Please keep in mind that the data described in this section should be the actual data used for analysis, so if you are using a subset of a larger dataset, please describe the subset that will actually be used in your study.

9. Existing data (required)

Example: ...

10. Explanation of existing data (optional)

Example: An appropriate instance of using existing data would be collecting a sample size much larger than is required for the study, using a small portion of it to conduct exploratory analysis, and then registering one particular analysis that showed promising results. After registration, conduct the specified analysis on that part of the dataset that had not been investigated by the researcher up to that point.

11. Data collection procedures (required)

Example: Participants will be recruited through advertisements at local pastry shops. Participants will be paid $10 for agreeing to participate (raised to $30 if our sample size is not reached within 15 days of beginning recruitment). Participants must be at least 18 years old and be able to eat the ingredients of the pastries.

12. Sample size (required)

Example: Our target sample size is 280 participants. We will attempt to recruit up to 320, assuming that not all will complete the total task.

13. Sample size rationale (optional)

Example: We used the software program G*Power to conduct a power analysis. Our goal was to obtain .95 power to detect a medium effect size of .25 at the standard .05 alpha error probability.

14. Stopping rule (optional)

Example: We will post participant sign-up slots by week on the preceding Friday night, with 20 spots posted per week. We will post 20 new slots each week if, on that Friday night, we are below 320 participants.

4 Variables

15. Manipulated variables

Gamble EV (one level) – 30-200
Complexity (two levels) – simple (2 outcomes, direct probability) and complex (2 outcomes, indirect probability)
Skewness (three levels) – left, no, right
Variance (SD - three levels) – low (5), medium (10), high (15)
EV Difference (five levels) - 20, -10, 0, 10, 20

16. Measured variables

Choice Task
- Choices
- Attention times (between the options)
- Manipulation Check: MCT about instructions)

Cognitive Ability Task
- Number of correctly solved matrices and questions

At the end of the experiment:
- Demographics (Age, Gender, education level)
- Comments

Overall:
Duration of each element (reaction times)

17. Indices (optional)

Example: We will take the mean of the two questions above to create a single measure of 'brownie enjoyment.'

5 Analysis Plan

18. Statistical models (required)

General note for multilevel models: As a standard, models will be implemented with random intercept for participant ID. More complex analyses (random slopes for predictors) will be added if there is substantial heterogeneity and if the models converge.

For Manipulation of complexity:
Option 1: one way ANOVA of participants' reaction time for stimuli in the easy vs. easy, complex vs. easy, easy vs. complex and complex vs. complex condition.

Option 2: Multilevel linear regression (random intercept) predicting reaction time, with the following predictors:

$$y = \beta_1 \cdot Complex\_n + \beta_2 \cdot EV\_diff + \beta_3 \cdot Var\_diff + \beta_4 \cdot Skew \qquad (1)$$

Complex_n:numbers of complex options in this trial
EV_diff: expected value difference
Var_diff: Standard Deviation difference
Skew: skewness conditions, ns, lr or rl

For H1: Multilevel logistic regression (random intercept) predicting left choice (1=left, 0=right) with the following predictors:

$$p = \beta_1 \cdot Complex + \beta_2 \cdot EV\_diff + \beta_3 \cdot Var\_diff + \beta_4 \cdot Skew \qquad (2)$$

Complex: whether the left option is easy or complex
EV_diff: expected value difference
Var_diff: Standard Deviation difference
Skew: skewness conditions, ns, lr or rl

19. Transformations (optional)

20. Inference criteria (optional)

As far as possible, we rely on Bayesian Statistics (95%-Credible Intervals).

21. Data exclusion

Exclusion:
P who need two attemps to pass the design questions (total attemps > 6)
All trials that has rt in the range of 0-5 and 95-100 in each condition
P who scored 0 in both CA test.
P who has less than 20% accuracy in control trials: EVD 20 or -20
completing the experiment too fast (1/4 of mean time) or too slow (three times mean time)
Other not allowed behaviours (e.g., restarting the experiment)

22. Missing data (optional)

23. Exploratory analysis (optional)

Cognitive ability test: Multilevel logistic regression (random intercept) predicting left choice (1=left, 0=right) with the following predictors:

$$\begin{aligned} p = \beta_1 \cdot Complex + \beta_2 \cdot EV\_diff + \beta_3 \cdot Var\_diff \\ + \beta_4 \cdot Skew + \beta_5 \cdot CA + \beta_6 \cdot CA \cdot Complex \end{aligned} \tag{3}$$

Complex: whether the left option is easy or complex
EV_diff: expected value difference
Var_diff: Standard Deviation difference
Skew: skewness conditions, ns, lr or rl
CA: average accuracy of the two cognitive ability tests

6 Other

24. Other