Sampling-based probability construction explains individual differences in risk preference

A thesis submitted in fulfilment of the requirements

for the degree of Master of Science (By Research)

by

Ankoju Bhanu Prakash

20128402

DEPARTMENT OF COGNITIVE SCIENCE

INDIAN INSTITUTE OF TECHNOLOGY KANPUR

September 2022

Certificate

It is certified that the work contained in this thesis entitled **"Sampling-based probability construction explains individual differences in risk preference"** by **Ankoju Bhanu Prakash** has been carried out under my supervision and that it has not been submitted elsewhere for a degree.

Dr. Nisheeth Srivastava
Associate Professor
Department of Computer Science and Engineering and Cognitive Science
Indian Institute of Technology Kanpur

Declaration

This is to certify that the thesis titled **"Sampling-based probability construction explains individual differences in risk preference"** has been authored by me. It presents the research conducted by me under the supervision of **Dr. Nisheeth Srivastava**.

To the best of my knowledge, it is an original work, both in terms of research content and narrative, and has not been submitted elsewhere, in part or in full, for a degree. Further, due credit has been attributed to the relevant state-of-the-art and collaborations with appropriate citations and acknowledgments, in line with established norms and practices.

A. Bhanuprakash

Ankoju Bhanu Prakash
Roll No. 20128402
Cognitive Science Department
Indian Institute of Technology Kanpur

Abstract

Name of the student: **Ankoju Bhanu Prakash**

Roll No: **20128402**

Degree for which submitted: **M.S (By Research)**

Department: **Cognitive Science**

Thesis title: **Sampling-based probability construction explains individual differences in risk preference**

Thesis supervisors: **Dr. Nisheeth Srivastava**

Month and year of thesis submission: **September 2022**

We present a how model for subjective probability distortions observed in people's behaviour in simple frequency estimation tasks and risky decisions [30]. The existing models [33, 13] of probability distortions assume encoding-based assumptions, inconsistent with evidence that people can reproduce probabilities veridically when elicited using graphical methods and motor movements [15, 29]. While in our model, we assume that probability distortions occur because people read out probability judgments as biased averages from working memory contents. Our model demonstrates the inverse S-shaped distortion of probability judgments in the simulation. Moreover, it also shows a clear relationship between working memory size and probability distortions, i.e. greater working memory capacity should lead to greater overweighting of small probabilities, which should lead to a particular fourfold pattern of risk preference as a function of working memory capacity. We conducted an experiment with human participants by considering cognitive ability measurements as a proxy for working memory capacity to validate our predictions. The model's predictions are consistent with the empirical results in three of four quadrants (HPG, LPG and HPL) and with earlier empirical studies of the relationship between cognitive ability and risk preference. Our results support a role for sampling during assessment of risky prospects, which in turn explains differences in probability distortions seen across different elicitation methods.

Acknowledgements

I am thankful to Dr. Nisheeth Srivastava for providing valuable guidance which has enhanced my clarity towards life and research values.

Contents

List of Figures

List of Tables

List of Publications

Publications from Thesis

1. Sampling-based probability construction explains individual differences in risk preference *https://escholarship.org/uc/item/9pm4x231*.

Chapter 1

Introduction

Consider the following gamble.

(a) You will win a reward of 100 for sure.

(b) There is a 20% chance that you will win a reward of Rs.500.

As per expected utility theory, an option with highest expected value would be preferred over other options, where the expected value is defined as $\sum_{i=1}^{n} p_i x_i$, where p_i is the probability of the event x_i. If we evaluate the above choice under the risk problem, as per expected utility theory. We can see that both the options are equally good, the expected value of A is (1*100) $100, and the expected value of B is also (0.8 * 0 + 0.2 * 500) $100. So there is no additional value we can yield by choosing one option over the other in the above gamble. But have you preferred one over the other in the above gamble?

There is a general consensus that people's decisions differ from the expected utility theory's rational decisions in a risky situation [23]. Despite the standard modelling approach, the expected utility theory does not explain people's decisions under risk. Prospect theory [30] is one of the theories proposed that explain the disparity between people's decisions and the expected utility theory's rational decisions in a risky situation.

1.1 Prospect Theory

As per prospect theory, a risky situation's outcome is evaluated as a relative measure with some reference point, which is a subjective value. The outcome can be treated as either gain or a loss, i.e., if the difference with respect to a reference point is positive, the outcome is classified as a profit. And similarly, the outcome is classified as a loss if the difference is negative with respect to the reference point.[30].

Prospect theory hypothesizes following, in gain conditions, the behaviour of the value function for changes in wealth is concave. For example, winning 200 over 100 appears more pleasant than winning 1200 over 1000. And in loss conditions, the behaviour of the value function for changes in wealth is convex. That is, a loss of 200 over 100 seems more awful than a loss of 1200 over 1000. Furthermore, the value function is multiplied by the decision weight, which is a subjective probability estimate of the true stated probability of that event [1]. And the weighting function exhibits an inverted S-shaped relationship, meaning that it tends to underestimate the large probability values and overestimate the low probability values, as shown in Figure 1.1[30].

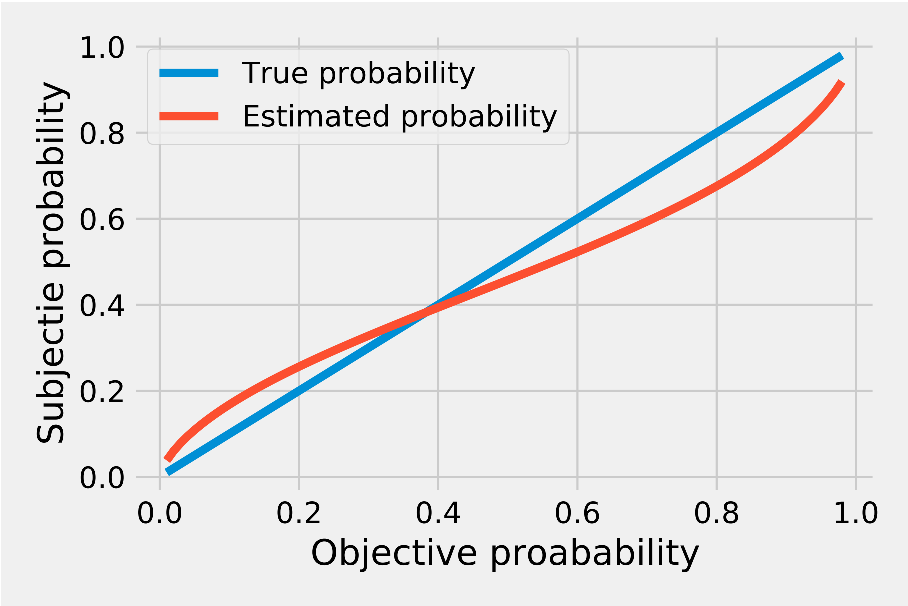

FIGURE 1.1: This figure demonstrates overestimation of low probabilities and underestimation of high probabilities.

Based on these posits, prospect theory characterises people's risk attitude in four quadrants as follows i.e., risk-averse in high probable gain and low probable loss conditions, and risk-seeking in low probable gain and high probable loss conditions[30, 18].

[1] $v(500)*\pi(0.2)+v(0)*\pi(0.8)$ is the value of the prospect in our gamble, where $v(x)$ is the value function and $\pi(p(x))$ is the decision weight for event x.

1.1.1 Prospect theory parameters correspond to stable individual differences in cognition

Given recent advances in our ability to estimate individual-level parameters for prospect theory [22], Glöckner and Pachur studied temporal consistency of prospect theory's parameter values by conducting experiments with risky choice problems across two sessions. They found that correlations for prospect theory's parameters within participants across sessions show a large effect size, suggesting that they correspond to stable individual differences in cognition [14].

But what are the cognitive processes that the prospect theory parameters map to? Pachur et al. investigated how prospect theory parameters can be interpreted with regard to attention allocation using the process tracing paradigm. And they found that the individual differences in prospect theory's parameters were systematically related to individual differences in attention paid to gains/losses and probability information[23].

Specifically, with respect to probability distortions, Zhang and Maloney proposed the following two-parameter linear form that best fits probability distortions observed in several studies.

$$Lo(w(p)) = \gamma \; Lo(p) + \; (1\gamma) \; Lo(p_0) \tag{1.1}$$

, where $Lo(p) = log\,(\frac{p}{1-p})$, $w(p)$ is the subjective probability estimate, p is true probability, p_0 is a crossover parameter determining where the inverse S-shaped distortion function switches from under-weighting to over-weighting , and γ is the linear transform slope parameter, which also determines the magnitude of probability distortion. Zhang and Maloney also proposed that assuming this linear log odds representation of probability in the brain is sufficient to explain the observed probability distortions across various studies. However, this representational claim is consistent with a large number of theoretical possibilities, [10, 9, 20] and thus does offer limited process-level understanding.

Goldstein and Rothschild investigated whether people encode distorted probability values or whether distortions were due to the elicitation process used. They found that people are able to reproduce probability distributions more accurately when elicited with graphical methods than standard methods. Suggesting people have accurate representation, but distortions are induced by the elicitation procedure. Combined with classic studies showing that frequency encoding in humans is significantly veridical [17], such findings suggest that cognitive processes during *retrieval* may be more likely to produce probability distortions.

1.2 Cognitive ability discerns risk aversion profiles

There are several studies [11, 4, 7], that share a common observation that people with higher cognitive ability have a risk-seeking attitude in certainty-equivalence experiments with low probable gains [11, 4] and are risk-averse in high probability gains [11, 7]. In an example from [11] study, participants who scored high on Cognitive Reflection Test demonstrated a great tendency to accept risky choices that lead to gains when an expected utility calculation favoured the risky choice, but crucially, even when it is not [11].

This behaviour of participants is consistent with the viewpoint of probability estimation. For low probability gains, participants with greater cognitive ability appear to be more risk-seeking, consistent with over-weighting the low probability gain option. Similarly, for high probability gains, such participants are more risk-averse, consistent with over-weighting the low probability non-gain option.

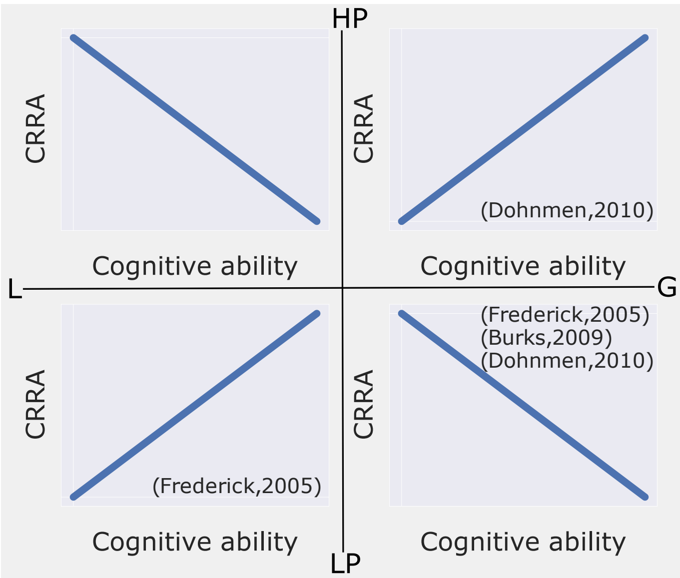

Figure 1.2: Risk sensitivity increases with measures of cognitive ability, as described in a number of behavioural studies. Multiple measures of risk preference and cognitive ability have been used in different studies. This figure plots the expected variation in the coefficient of relative risk aversion (CRRA) with the increase in cognitive ability in the four patterns of behaviour observed in the prospect theory view of risk aversion, along with references to field studies that support the prediction in the particular quadrant.(HP - High Probable, LP- Low Probable, G- Gain, L- Loss)

We outline the expected relationship between cognitive ability and risk aversion in Figure 1.2, where we consider the source of risk preference lies in the over-weighting of low-probable lottery outcomes. The papers referenced in the figure show evidence consistent with the prediction relevant for each quadrant. Thus, convergent evidence across studies suggests a relationship between cognitive ability and probability distortions.

In this report, we present sampling-based model of probability judgment for risky prospects that illuminates the relationship between cognitive ability and probability distortions. We explain the theoretical details of the model in the next chapter.

Chapter 2

Probability by sampling

According to prospect theory, when given a choice between risky prospects, people behave as if they were constructing a subjective probability estimate $w(p)$ based on the provided prospect risk p. If we consider this process hypothesis seriously, we should ask: how do people associate p to w? We propose that they do this by sampling from mental simulations, which have recently proven to be successful in explaining people's perception of physical situations [27], as well as biases in probability judgments [35]. Suppose, for example, when given a choice between risky prospects, people attempt mentally to simulate the lottery. And every simulation run will yield one of the payout options of the lottery. Multiple simulations of the lottery end up with different results. The frequencies of the payout options from the accumulated sample of lottery results from the mental simulations may inform the subjective probability estimates of Risky prospects.

Focusing on probability judgments for evaluating binary prospects, for simplicity, our probability-by-sampling model assumes that,

1. Observers possess a veridical, possibly noisy, internal probability scale.
2. When asked to reflect on a risky binary prospect, observers sample multiple abstract lotteries parameterized by the prospect risk, as read off the internal scale.
3. The outcomes of these simulated lottery draws are stored in working memory.
4. Observers sample from the lottery until either working memory capacity is reached[1], or both prospects have occurred at least once during sampling.

[1] If memory sampling fails to retrieve a sample of the low probability outcome by the time capacity is reached, the model returns a probability of 0.01 for the low probability outcome.

5. Observers read out the average occurrence of the salient option as their subjective probability estimate for it.

Of these assumptions, #1 follows standard psychophysical premises, #2 is the key sampling assumption of our approach, #3 follows standard assumptions about the role of working memory made in nearly all symbolic cognitive architectures [32], #4 is a novel assumption made based on the indicator variable 'counting predictor' from [28]. By definition, the counting predictor is the sample count at which all possible options have occurred at least once while sampling, which substantially improved the sampling duration predictions in a decision from experience paradigm [28] and #5 is standard. Thus, the novelty of our model lies in assumptions #2 and #4.

Formally, where $\mathcal{I}_m$ is an indicator function that takes the value 1 if the low probability outcome is sampled in the m^{th} memory slot, and 0 otherwise. Also, $\mathcal{M}$ represents the set of memory slots in working memory filled up at the time $w(p)$ is read out (up to maximum capacity), which in turn is determined by the number of samples it takes to see two distinct outcomes during sampling.

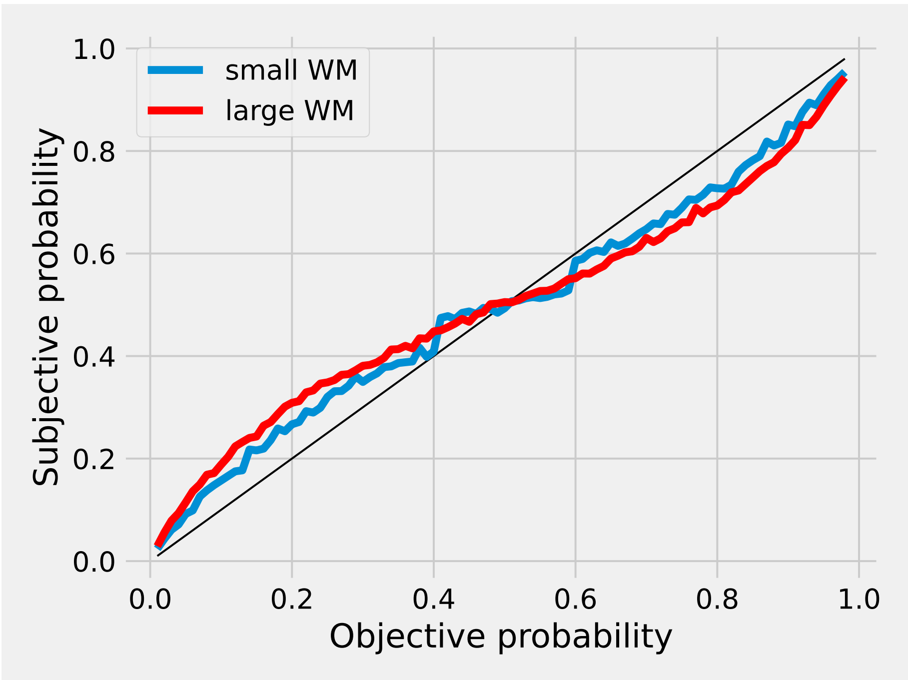

FIGURE 2.1: Subjective probability judgments were extracted from the probability-by-sampling model for cohorts of 1000 observers sampled from low (blue) and high (red) working memory capacity pools.

We conducted an *in silico* experiment, by sampling 1000 probability-by-sampling observers with working memory capacities sampled from normal distributions with means $\mu_{low} = 5, \mu_{high} = 10$ and SD = 1. Observers from low and high Working memory capacity groups responded to binary prospects across all possible probability values (quantized in steps of 0.01), producing subjective probability estimates for all these values. Figure 2.1 plots the average of these estimates for both WM size groups.

There are two important observations. One, we see that probability-by-sampling observers produce an inverse-S shaped distortion of probabilities [30], based only on the retrieval stage assumptions. This is consistent with the empirical evidence that people can reproduce probabilities veridically in some elicitation formats [29, 15]. And two, we note that the high WM group shows greater probability distortion than the low WM group. These observations remain consistent across multiple numeric values of our simulation parameters, but with working memory sizes greater than 12 probability distortions fade away.

Let's consider the prospect "20% chance of winning Rs.500 otherwise, you win nothing" to understand why the probability-by-sampling model generates inverse S-shaped probability distortions. As the probability-by-sampling observers sample simulated outcomes until they see both outcomes at least once, and then average over the outcomes to read out the lower probability, there are two possibilities. Either they do not sample the low probability option, or they will sample the low probable outcome once and terminate sampling. In our example, for working memory size 4, the probability that the low-probable outcome(20% chance of winning Rs.500) will not be sampled is $0.8^4 = 0.409$ [2], so the readout probability by the observer will be 0 less than 50% of the time. So, for more than half the time, the low probable outcome is sampled, and the read-out value will be inflated, for example, if the observer has sampled the following sequence of outcomes [H - 1, L - 0](where H be the high probable outcome, represented as 1 and L be the low probable outcome, represented as 0), the readout probability will be $(1 + 0)/2 = 0.5$, as the observer terminates the sampling process since it has seen both the outcomes. Averaged across the population, this asymmetry yields the probability of over-weighting.

Similarly, for large working memory sizes, let's say 8, the chance of not sampling the low probability outcome at least once reduces still further to $0.8^8 = 0.17$. In 83% of cases (1 - $0.8^8 = 0.83$), the observer will sample the low-probability outcome at least once and read out a subjective probability estimate equal to or greater than the objective probability.

[2] If p is the probability of an event, then the probability of the event occurring x times in a binomial experiment, of n trials is $nCx * p^x * (1 - p)^x$

And a number of instances where readout probability values as zero are much lower for high working memory observers than low working memory observers. Averaged across observers, this leads to a greater probability of over-weighting for high WM observers.

We calculated the slope of linear transformation γ for the observed average read-out probabilities corresponding to the population of observers and true probability values, from Equation 1.1. Figure 2.2 shows the slope of linear transformation γ (also represents the magnitude of probability distortions) decreases with working memory size. This indicates that the high WM group shows greater probability distortion than the low WM group(lower γ higher the distortions).

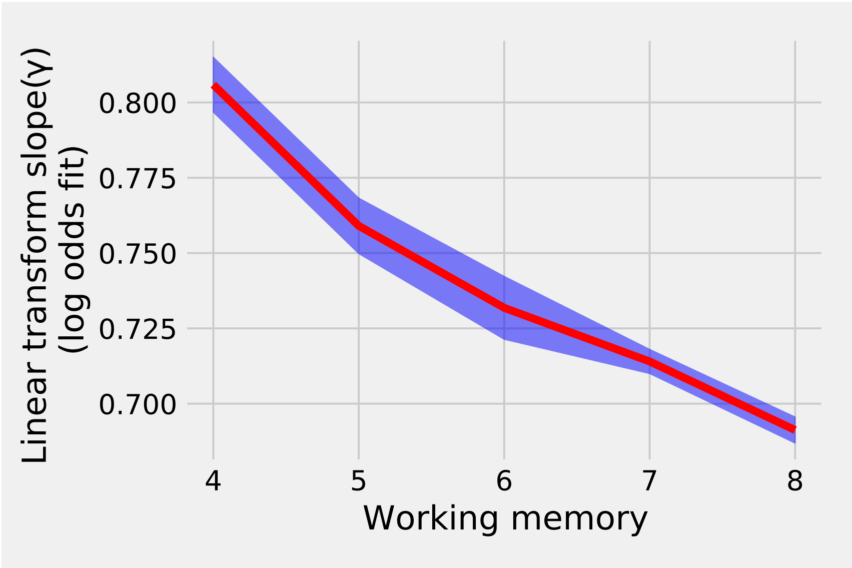

FIGURE 2.2: This figure plots the relationship between the slope of a linear transformation (γ) vs working memory size, obtained by simulating 200 observes with working memory capacities ranging from 4 to 8. The mean and the standard deviation for the slope parameter are calculated by constructing a sample of gamma fits by conducting simulations multiple times, in this case we ran simulations for 30 times.

To summarise, the key novelty of the probability-by-sampling account of probability distortions is the assumption that observers mentally simulate lottery outcomes until they have seen at least one instance of both lottery prospects. The model demonstrates the inverse S-shaped distortion of probability judgments, using only retrieval-stage assumptions. And it indicates a clear relationship between working memory size and probability distortions, i.e. greater working memory capacity should lead to greater overweighting of small probabilities. Based on this relationship between cognitive ability and probability distortions, we outline the expected relationship between cognitive ability and risk aversion in Figure 1.2. We conducted an experiment with human participants to verify our predictions. In the next chapter, we discuss the experimental procedure in detail.

Chapter 3

Methods

While previous studies partially support the existence of the fourfold pattern illustrated in Figure 1.2, differences in protocols, analysis methods and operationalization of both independent and dependent variables make it difficult to assess the net weight of the evidence. To address this concern, we conducted an experiment to measure risk aversion as the CRRA coefficient of isoelastic utility functions in certainty equivalence problems selected to represent each of the four quadrants(HPG, LPG, LPL and HPL) for participants with different cognitive ability levels, as measured by RSPM (Raven's Standard Progressive Matrices). We expected to see the specific relationship pattern between cognitive ability and risk aversion, as predicted in Figure 1.2 as the outcome of this experiment.

3.1 Subjects

We invited participants via email and social media. 156 participants (58 female, 98 male) responded and provided consent for participation. Out of 156 participants who appeared for the IQ test, 122 participants (47 female, 75 male) expressed interest in participating in the online risk-preference study. The mean age of the participants was 23.83 years. Since this was a between-subject design, participants were assigned to one of the four quadrants randomly at the time of experiment participation. All experimental protocols were approved by an Institutional Review Board. Participants signed a consent form describing all experimental procedures before participating in the study. Each participant was compensated for their time.

3.2 Measuring cognitive ability

It is well-known that working memory capacity is strongly correlated with general cognitive ability, as measured by progressive matrices tests [12]. So to measure cognitive ability, we used Raven's Standard Progressive(SPM) Matrices [24] containing 60 questions. We designed a website to administer the test online. Participants were shown puzzles from SPM one by one on the screen with corresponding options. They had to answer the puzzles by clicking one of the options. The raw scores (number of correct responses) obtained for each participant were converted to standard SPM percentiles using the SPM manual. There was no time limit for the test. Out of 122 participants, four participants whose SPM's standard score was 5 were excluded from analysis since their test duration was less than six minutes for 60 questions, suggesting random responses leaving us with 118 participants(44 female, 74 male) for the risk preference experiment. Participants were assigned randomly to the four quadrants of the experiment, and the breakdown is given in the below table. The average time to complete the IQ test by the participants was 35.06min, and our sample's average SPM percentile score was 62.7, suggesting that it was representative.

Table 3.1: Participants count in all quadrants.

Quadrant	Participant Count
HPG	30
LPG	25
LPL	35
HPL	28

Table 3.1 documents the participant count for all four quadrants.

3.3 Measuring risk preference

We measured risk preference for each participant using a choice table, which had 20 rows. Every participant provided their preference for each row of the table.[7]. The choice tables used, follow the ones used in Dohmen et al..

TABLE 3.2: Sample choice table, from high probable gain quadrant

	Lottery Prospect	Accept the safe amount
1	90% chance of winning amount 16,200	10100
2	90% chance of winning amount 16,200	10400
3	90% chance of winning amount 16,200	10700
4	90% chance of winning amount 16,200	11000
5	90% chance of winning amount 16,200	11400
6	90% chance of winning amount 16,200	11700
7	90% chance of winning amount 16,200	12000
8	90% chance of winning amount 16,200	12300
9	90% chance of winning amount 16,200	12700
10	90% chance of winning amount 16,200	13000
1 1	90% chance of winning amount 16,200	13300
12	90% chance of winning amount 16,200	13600
13	90% chance of winning amount 16,200	14000
14	90% chance of winning amount 16,200	14300
15	90% chance of winning amount 16,200	14600
16	90% chance of winning amount 16,200	14900
17	90% chance of winning amount 16,200	15300
18	90% chance of winning amount 16,200	15600
19	90% chance of winning amount 16,200	15900
20	90% chance of winning amount 16,200	16000

Table 3.2, a sample choice table, belongs to the High Probable Gain quadrant,note that the expected value of the gamble matches the safe amount at 15^{th} row $(0.9 * \$16,200 = \$14,580 \approx \$14,600)$. The lottery amounts and payoffs for all the choice tables were derived from Frederick study 1 and then converted to equivalent local currency by considering Purchasing Power Parity in 2005 and inflation (2005-2021). The lottery amount remained the same while the safe option increased systematically for every row; a rational agent would be willing to take risks until the safe amount is less than the expected value of the gamble and then switches to the safe option. We ensured the safe amount matches the expected value of gamble at 15^{th} /16^{th} row in all choice tables. we followed [7], where the safe amount crosses the expected value of the gamble at 15^{th} row.

For gain-based problems, we asked the participants to choose whether to buy a lottery ticket that could fetch them lottery money with some uncertainty or accept the safe amount. Each quadrant had a set of five choice tables with different lottery amounts and payoffs. The choice tables in these quadrants represent either high probable or low probable gain conditions. The order of the choice table presented to each participant was randomized. We presented participants one row at a time and asked to choose whether 'to buy the lottery ticket (risky)', which can fetch them a lottery amount with some uncertainty, or 'Not to buy the lottery ticket (safe)' and accept the safe amount.

Once the participant switched from the risky option to the safe option, the algorithm asked the participant whether they would accept all higher safe amounts or not (see also [7]) if they responded yes, the algorithm considered all other safe options in the table as their preferences, and the participant was progressed to the new choice table. Otherwise, the participant had to decide on the rest of the table manually and then be presented with a new choice table. Following Dohmen et al., we also informed participants that one row from one of the five-choice tables would be randomly selected, and they would be rewarded with the amount proportional to the choice they made in that selected row to encourage participants to choose according to their true preferences for each row.

The same procedure was used for loss-based problems, except that the problems were framed as a choice to buy insurance costing a small fixed amount or retain a small probability of suffering a larger loss. And here, the uncertain loss amount remained the same while the insurance premium decreased systematically for every row; a rational agent would be willing to take risks until the insurance cost is greater than the expected value of the gamble and then switches to the insurance cover. We ensured the insurance cost matches the expected value of gamble at 15^{th} /16^{th} row in all choice tables. we followed [7] study, where the safe amount crosses the expected value of the gamble at 15^{th} row.

Out of 590 instances(118 participants * 5 choice tables), there were 16 incidents where participants switched from a risk option to a safe option multiple times. The sixteen instances can be classified into two scenarios. Scenario-1, they selected safe options consecutively. One example, a participant switched from a risky option to a safe option at 10^{th} row and selected a safe option again in the next rows(11,12) and then moved to the next choice table. We considered the first switch (in this example, 10^{th} row) as their risk preference. And in scenario -2, they switched from a risky option to a safe option and again selected the risky option, then switched to a safe option. One example, a participant switched from a risky option to a safe option at 4^{th} row and then selected the risky option in the 5^{th} row and continued the risky option till row 10 and switched to a safe option at 11^{th}

row. Here we considered the latest switch(here row 11) as their risk preference. And in another example, a participant switched from the risky option to a safe option at 4^{th} row and then selected the risky option in the 5^{th} row and continued the risky option till the end of the table. Here we considered the last row(20^{th} row) as their risk preference.

The coefficient of relative risk aversion(CRRA) was calculated from an individual's utility function [4]. We follow Burks et al. in assuming that the participant's utility for the lottery would be at the midpoint of $safe_i$ and $safe_j$. (where 'i' and 'j' refers to the steps when the participant prefers to take the risk at $safe_i$, but switches to the safe option at $safe_j$.). The individual's utility function is then given by,

$$u(c) = \frac{c^{1-\sigma}}{1-\sigma}, \tag{3.1}$$

where σ is the CRRA coefficient, we are interested in measuring.

Following Burks et al. and assuming expected utility maximization, the equation below holds when a participant switches their lottery preference between cells i and j of the table and is solved analytically for lottery utility and then numerically for σ to obtain the coefficient of relative risk aversion,

$$p\ u(lottery) = 0.5\, u(safe_i) + 0.5\, u(safe_j), \tag{3.2}$$

where p corresponds to the stated probability of winning the lottery. The same procedure was used to estimate CRRA in loss conditions as well.

We calculated CRRA for every row in all twenty choice tables(4 quadrants * 5 choice tables) using a numerical solver, we capped CRRA values to avoid large numbers. Figure 3.1, plots CRRA values for every choice table in all quadrants, as we move down the table, the CRRA values decrease, indicating a more risk-seeking nature.

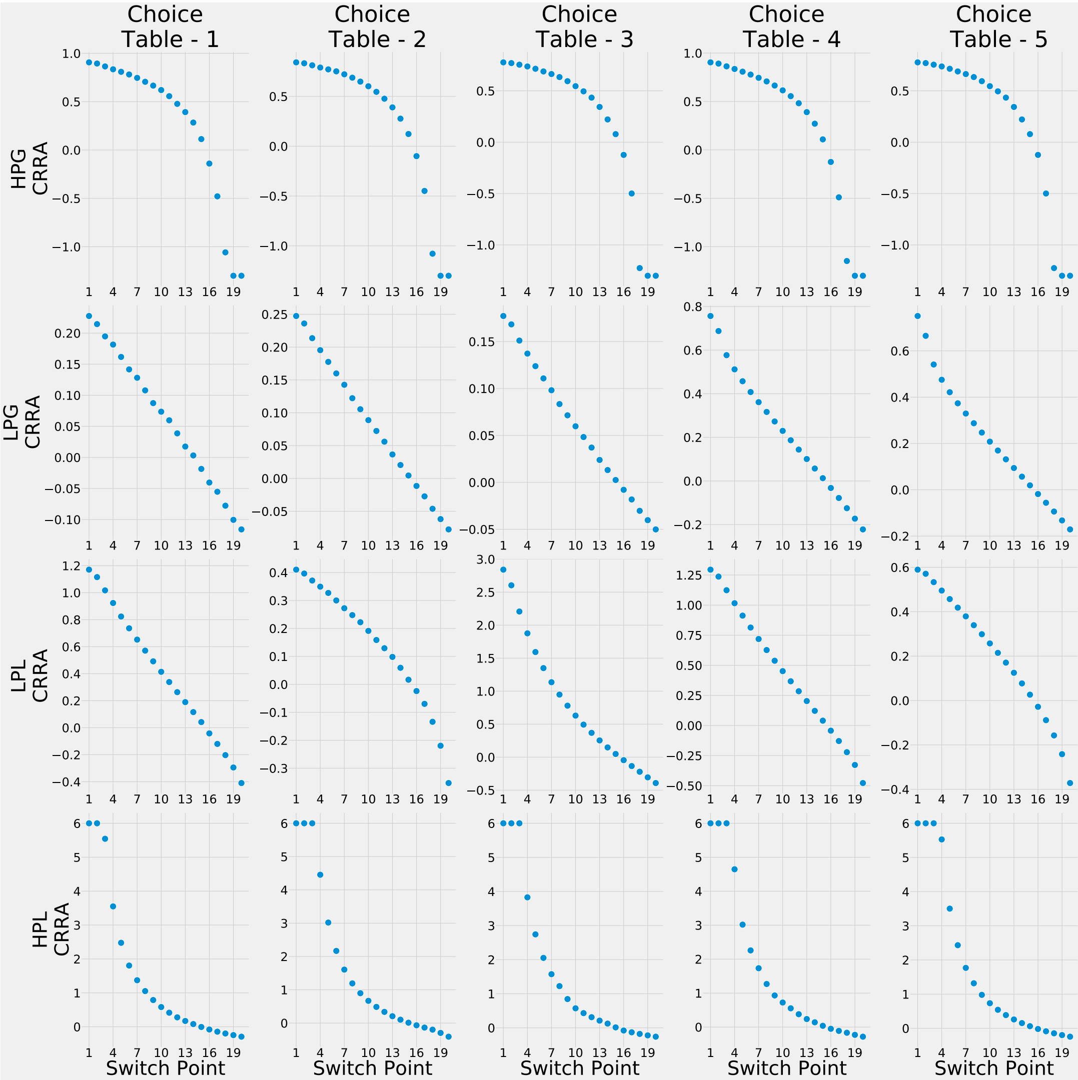

FIGURE 3.1: This figure plots, how CRRA changes as we move down a table, for every choice table in all quadrants.

Chapter 4

Results

The mean CRRA estimates for all five choice tables seen by participants in each quadrant are shown in Figure 4.1. In every quadrant, for each choice table, we found the best fit line relating CRRA to IQ. To obtain a summary measure of the trend across choice tables for each quadrant, we shifted the CRRA points from each choice table to a common intercept (the average intercept across the best fit lines). We then replotted the points using individual slope values from the table-wise best fit lines. Finally, we fitted a linear regression to the combined CRRA estimates (see the rightmost column in Figure 4.1).

Table 4.1: Average slope in all quadrants.

Quadrant	Sign Prediction	Coefficient	p	f^2
HPG	+	0.31	0.03	0.17
LPG	-	-0.03	0.03	0.17
LPL	+	-0.21	0.09	0.1
HPL	-	-2.41	0.000008	1.06

Table 4.1 documents the slope value of the combined regression for all four quadrants, alongside the predicted coefficient sign, as seen in Figure 1.2. We note that the measured coefficients are directionally consistent with our predictions in three of four quadrants. Results for three quadrants (high probability gains, low probability gains and high probability losses) were statistically significant at the traditional 0.05 alpha-error level and displayed medium effect sizes ($f^2 > 0.15$) [5]. For the low probable loss quadrant, we see small effect sizes ($f^2 > 0.02$), with the relationship failing to meet statistical significance.

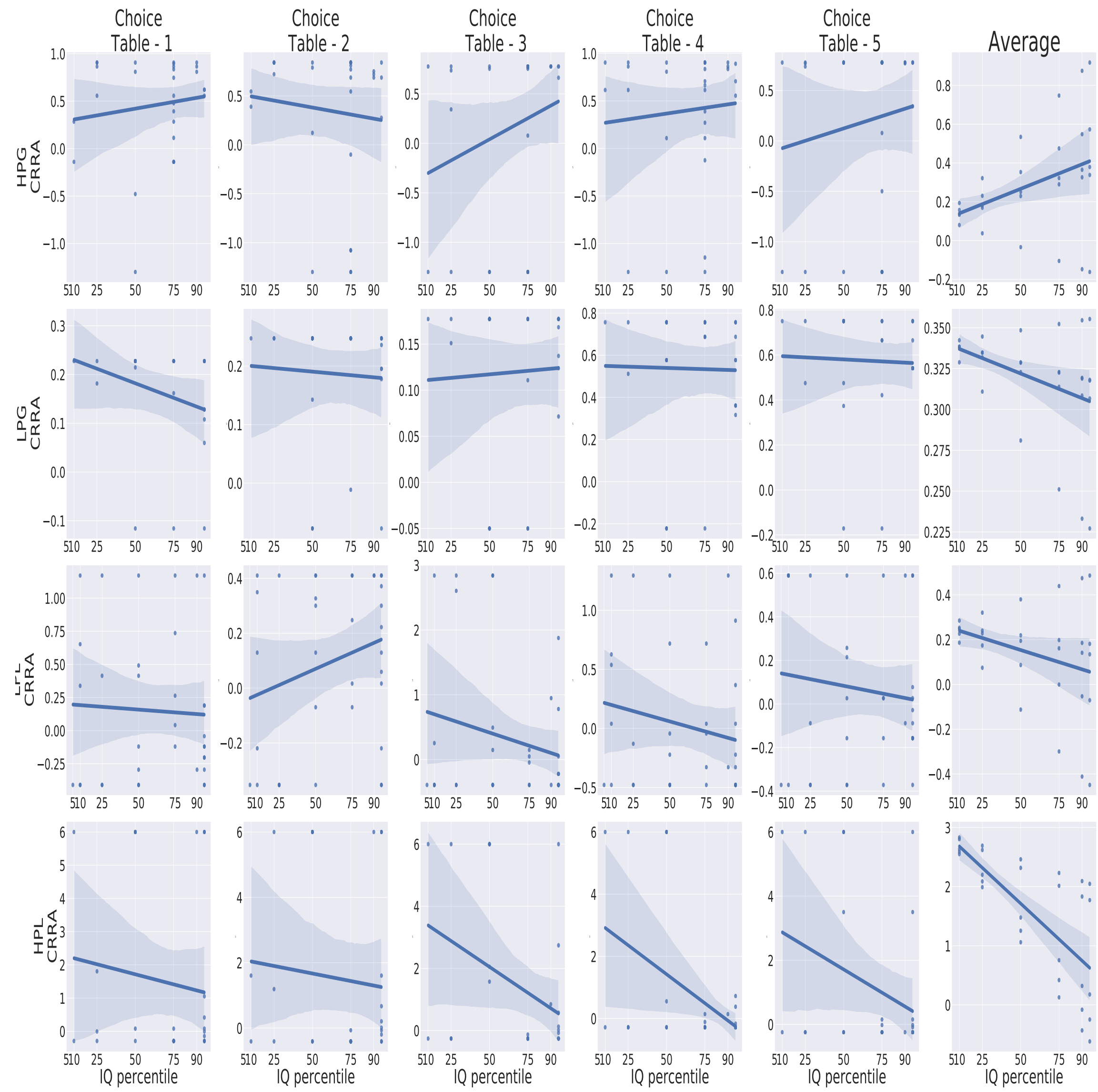

FIGURE 4.1: This figure plots the relation between CRRA vs IQ for every choice table in all the quadrants and the average plot to show the overall trend in a quadrant.

To verify that the observed relationships between cognitive ability and risk preference are not an artefact of our data pooling procedure across choice tables, we fit a hierarchical linear regression model for every quadrant separately. We model the relationship between

CRRA and IQ in each quadrant as follows

$$CRRA_i = slope_i * IQ + intercept_i + \epsilon$$
$$slope_i \sim \mathcal{N}(\mu_{slope}, {\sigma^2}_{slope})$$
$$intercept_i \sim \mathcal{N}(\mu_{intercept}, {\sigma^2}_{intercept})$$
$$\epsilon \sim HalfCauchy(5)$$

where $slope_i$, $intercept_i$ are the slope and intercept parameters for the choice-table 'i' in a quadrant and ϵ is noise. We used Gaussian and half-Gaussian priors, respectively, for our two mean and two standard deviation hyperparameters.

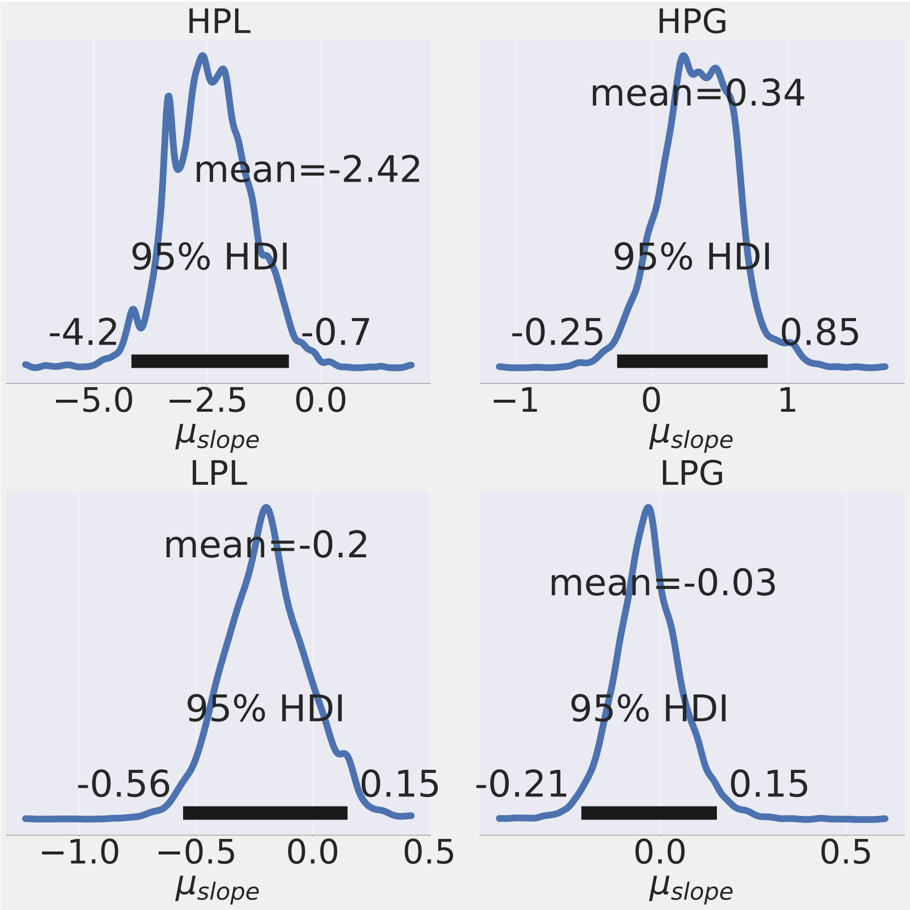

FIGURE 4.2: This figure plots the posterior distribution for the mean distribution from which slopes of all quadrants are sampled with 95% credible intervals.

We fit this model using PyMC3's NUTS sampler using 2 chains of 2000 draw iterations with 1000 tuning steps. The key parameter of interest for us is the mean of the distribution of

μ_{slope} from which slopes for different choice sets are sampled. Figure 4.2 plots the quadrant-wise posterior distributions for μ_{slope} from the fitted model. The key observation is that the MAP estimates of μ_{slope} reliably track the average slope estimates we obtained in our pooled analysis, suggesting that the pattern seen in the previous analysis is not an artefact of the data pooling procedure.

Chapter 5

Discussion

In this work, we presented a sampling-based probability judgment model for risky prospects based only on retrieval-stage assumptions. The model demonstrated the inverse S-shaped distortion of probability judgments. And also showed a clear relationship between the working memory size and probability distortions, i.e. greater working memory capacity should lead to greater over-weighting of small probabilities. And further predicted a pattern of the relationship between cognitive ability and risk aversion in the fourfold pattern of risk attitudes. The model's predictions are consistent with the empirical results in three of four quadrants (HPG, LPG and HPL) and with earlier empirical studies of the relationship between cognitive ability and risk preference [11, 4, 7]. Andersson et al. proposed that the relationship between cognitive ability and risk aversion is spurious, and the direction of correlation depends on the behavioural noise and the biased risk elicitation method. In the current study, all the choice tables in every quadrant are biased in the same direction, yet we still see a positive and negative correlation between cognitive ability and risk aversion. This suggests that the relationship between cognitive ability and risk aversion does not depend solely on the bias in the risk elicitation method and the behavioural noise. However, the model's prediction for the low probable loss quadrant is directionally inconsistent with the participant's data, and the observed correlation between cognitive ability and risk preference also failed to meet statistical significance. Further work is needed to verify it.

While the previous models of probability distortions [10, 9] that use retrieval-specific assumptions are context- and task-specific models. The probability-by-sampling model, oriented towards prospect risk in our presentation, can be extended to other tasks easily. Considering the frequency estimation task, we need to assume that observers need to sample tokens until they sample the one they are estimating the frequency of for once, and then average across the token count to produce a frequency estimate. And also, the other previous models of probability distortions [13, 33], based on encoding-based assumptions, are inconsistent with the empirical evidence [15] that people can reproduce probabilities veridically when these are elicited using graphical methods. Whereas Probability-by-sampling is a task-general retrieval-based model of probability distortions, able to accommodate the possibility of veridical encoding of frequency information [17] and the possibility of near-veridical retrieval of probability information using non-symbolic elicitation procedures [15].

Bibliography

[1] Ola Andersson, Håkan J Holm, Jean-Robert Tyran, and Erik Wengström. Risk aversion relates to cognitive ability: Preferences or noise? *Journal of the European Economic Association*, 14(5):1129–1154, 2016.

[2] Bhanu Prakash Ankoju and Nisheeth Srivastava. Sampling-based probability construction explains individual differences in risk preference. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, volume 44, 2022.

[3] Nicholas C Barberis. Thirty years of prospect theory in economics: A review and assessment. *Journal of Economic Perspectives*, 27(1):173–96, 2013.

[4] Stephen V Burks, Jeffrey P Carpenter, Lorenz Goette, and Aldo Rustichini. Cognitive skills affect economic preferences, strategic behavior, and job attachment. *Proceedings of the National Academy of Sciences*, 106(19):7745–7750, 2009.

[5] Jacob Cohen. The effect size. *Statistical power analysis for the behavioral sciences*, pages 77–83, 1988.

[6] Joshua R De Leeuw. jspsych: A javascript library for creating behavioral experiments in a web browser. *Behavior research methods*, 47(1):1–12, 2015.

[7] Thomas Dohmen, Armin Falk, David Huffman, and Uwe Sunde. Are risk aversion and impatience related to cognitive ability? *American Economic Review*, 100(3): 1238–60, 2010.

[8] Jacques Dutka. On the st. petersburg paradox. *Archive for History of Exact Sciences*, 39(1):13–39, 1988.

[9] Craig R Fox and Yuval Rottenstreich. Partition priming in judgment under uncertainty. *Psychological Science*, 14(3):195–200, 2003.

[10] Craig R Fox and Amos Tversky. A belief-based account of decision under uncertainty. *Management science*, 44(7):879–895, 1998.

[11] Shane Frederick. Cognitive reflection and decision making. *Journal of Economic perspectives*, 19(4):25–42, 2005.

[12] Keisuke Fukuda, Edward Vogel, Ulrich Mayr, and Edward Awh. Quantity, not quality: The relationship between fluid intelligence and working memory capacity. *Psychonomic bulletin & review*, 17(5):673–679, 2010.

[13] Samuel Gershman and Robert Wilson. The neural costs of optimal control. *Advances in neural information processing systems*, 23:712–720, 2010.

[14] Andreas Glöckner and Thorsten Pachur. Cognitive models of risky choice: Parameter stability and predictive accuracy of prospect theory. *Cognition*, 123(1):21–32, 2012.

[15] Daniel G Goldstein and David Rothschild. Lay understanding of probability distributions. *Judgment & Decision Making*, 9(1), 2014.

[16] Charles R. Harris, K. Jarrod Millman, Stéfan J van der Walt, Ralf Gommers, Pauli Virtanen, David Cournapeau, Eric Wieser, Julian Taylor, Sebastian Berg, Nathaniel J. Smith, Robert Kern, Matti Picus, Stephan Hoyer, Marten H. van Kerkwijk, Matthew Brett, Allan Haldane, Jaime Fernández del Río, Mark Wiebe, Pearu Peterson, Pierre Gérard-Marchant, Kevin Sheppard, Tyler Reddy, Warren Weckesser, Hameer Abbasi, Christoph Gohlke, and Travis E. Oliphant. Array programming with NumPy. *Nature*, 585:357–362, 2020. doi: 10.1038/s41586-020-2649-2.

[17] Lynn Hasher and Rose T Zacks. Automatic processing of fundamental information: the case of frequency of occurrence. *American psychologist*, 39(12):1372, 1984.

[18] Ralph Hertwig, Greg Barron, Elke U Weber, and Ido Erev. Decisions from experience and the effect of rare events in risky choice. *Psychological science*, 15(8):534–539, 2004.

[19] Wei Ji Ma, Masud Husain, and Paul M Bays. Changing concepts of working memory. *Nature neuroscience*, 17(3):347–356, 2014.

[20] André CR Martins. Probability biases as bayesian inference. *Judgment and Decision Making*, 1(2):108, 2006.

[21] Wes McKinney et al. Data structures for statistical computing in python. In *Proceedings of the 9th Python in Science Conference*, volume 445, pages 51–56. Austin, TX, 2010.

[22] Håkan Nilsson, Jörg Rieskamp, and Eric-Jan Wagenmakers. Hierarchical bayesian parameter estimation for cumulative prospect theory. *Journal of Mathematical Psychology*, 55(1):84–93, 2011.

[23] Thorsten Pachur, Michael Schulte-Mecklenbeck, Ryan O Murphy, and Ralph Hertwig. Prospect theory reflects selective allocation of attention. *Journal of Experimental Psychology: General*, 147(2):147, 2018.

[24] John C Raven, John Hugh Court, and John Earle Raven. *Standard progressive matrices*. Australian Council for Educational Research Limited, 1989.

[25] John Salvatier, Thomas V. Wiecki, and Christopher Fonnesbeck. Probabilistic programming in python using PyMC3. *PeerJ Computer Science*, 2:e55, apr 2016. doi: 10.7717/peerj-cs.55. URL `https://doi.org/10.7717/peerj-cs.55`.

[26] Skipper Seabold and Josef Perktold. statsmodels: Econometric and statistical modeling with python. In *9th Python in Science Conference*, 2010.

[27] Kevin A Smith and Edward Vul. Sources of uncertainty in intuitive physics. *Topics in cognitive science*, 5(1):185–199, 2013.

[28] Nisheeth Srivastava, Johannes Müller-Trede, Paul R Schrater, and Edward Vul. Modeling sampling duration in decisions from experience. In *CogSci*, 2016.

[29] Julia Trommershäuser, Laurence T Maloney, and Michael S Landy. Statistical decision theory and the selection of rapid, goal-directed movements. *JOSA A*, 20(7):1419–1433, 2003.

[30] Amos Tversky and Daniel Kahneman. Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty*, 5(4):297–323, 1992.

[31] Guido Van Rossum and Fred L. Drake. *Python 3 Reference Manual*. CreateSpace, Scotts Valley, CA, 2009. ISBN 1441412697.

[32] Peijun Ye, Tao Wang, and Fei-Yue Wang. A survey of cognitive architectures in the past 20 years. *IEEE transactions on cybernetics*, 48(12):3280–3290, 2018.

[33] Hang Zhang and Laurence T Maloney. Ubiquitous log odds: a common representation of probability and frequency distortion in perception, action, and cognition. *Frontiers in neuroscience*, 6:1, 2012.

[34] Hang Zhang, Xiangjuan Ren, and Laurence T Maloney. The bounded rationality of probability distortion. *Proceedings of the National Academy of Sciences*, 117(36): 22024–22034, 2020.

[35] Jian-Qiao Zhu, Adam N Sanborn, and Nick Chater. The bayesian sampler: Generic bayesian inference causes incoherence in human probability judgments. *Psychological review*, 127(5):719, 2020.