WHICH RISK ATTITUDE COMPONENTS ARE RELATED TO INTELLIGENCE?

—— DISENTANGLING THE INFLUENCE OF, AND STUDYING THE INTERACTION BETWEEN COGNITIVE ABILITY AND EXPERIENCE ON PROBABILITY WEIGHTING TOWARDS RISK

Tingying Dong 406455td

Supervisor: Peter Wakker

Which Risk Attitude Components Are Related to Intelligence?

—— Disentangling the influence of, and studying the interaction between cognitive ability and experience on probability weighting towards risk

ABSTRACT. Prospect theory explained the violations of traditional theories of choice under risk and it provided more behavioural foundations showing that the evaluation of probabilities could be subjective along with outcomes. Probability weighting consists of two components: likelihood insensitivity and pessimism. The two-system theory suggested a plausible mechanism of how cognitive limitations may influence people’s understanding of probabilities and therefore on the overweighting of rare events. This study conducted a laboratory experiment to investigate the effect of, and the interaction between cognitive ability and experience on probability weighting. Among the 32 participants who were highly educated and young with a small range of age, cognitive ability and direct experience from sampling were both found to be negatively correlated with likelihood insensitivity. Pessimism was found to be influenced by experience although the direction was not clear. As for the interaction, cognitive ability affected the willingness to sample and how the participants learned from the procedure of sampling to moderate their likelihood insensitivity and pessimism levels. Compared to previous studies, this research involved a model of probability weighting to study the influence of both cognitive ability and experience, and investigated the complex interaction between them. Suggestions of further investigation and possible applications on paternalistic policies are discussed.

KEY WORDS: probability weighting, likelihood insensitivity, pessimism, cognitive ability, experience.

Contents

1. Introduction

Theories of choice under risk have been discussed for centuries. Since the failure of the early expected value theory representatively demonstrated by the St. Petersburg paradox, to the expected utility theory (Bernoulli, 1738; Von Neumann & Morgenstern, 1944), the evaluation of outcomes started to be assumed subjective. From the later violation suggested by Allais' paradox (Allais, 1953) representatively, to original prospect theory (Kahneman & Tversky, 1979), more behavioural foundations were provided and the evaluation of given probabilities also started to be considered as subjective. Inspired by original prospect theory, rank-dependent utility (Quiggin, 1982), the new version of prospect theory, sometimes called cumulative prospect theory (Tversky & Kahneman, 1992), inverse-S shaped probability weighting function (Tversky & Fox, 1995), and bounded subadditivity (Tversky & Wakker, 1995) were developed to give more explanations. In the end, likelihood insensitivity and pessimism were combined to describe people's risk attitudes (Wakker, 2010). People's worse understanding on uncommon probabilities and their tendency to overweight rare events reflect cognitive limitations and the possibility to be corrected by proper learning. And therefore, my research question to disentangle the effect of, and investigate the interaction between cognitive ability and experience on probability weighting towards risk was formed.

The two-system theory (Stanovich & West, 2000; Kahneman & Frederick, 2002; Kahneman, 2003) was developed to explain the mechanism of cognitive process. It was used as the main theory to explain the possible mechanisms of the effect of cognitive ability and experience. System 1 is related to intuitive thinking and System 2 is related to deliberative thinking. The thinking process of System 1 is effortless so that it would generate fast but possibly problematic results. Especially when facing unfamiliar situations, for example, extreme probabilities, erred results could result from making use of heuristics and attribute substitution. System 2 is effortful and could be used to detect if the results generated by System 1 are biased, although this procedure could not assure the mistakes to be totally avoided. Cognitive ability could be a reflection of the performance of the two systems. Experience could influence the accessibility of related information when heuristics are used.

Many studies have confirmed the existence of a correlation between cognitive ability and risk attitude (Frederick, 2005; Dohmen et al., 2008; Burks et al., 2009; Cokely and Kelley, 2009; Benjamin et al., 2013; Choi et al., 2014), while some studies did not find a correlation (Eckel et al., 2012; Andersson et al., 2013; Tymula et al., 2013; Petrova et al., 2014). However, most of them only generally described the relation or related cognitive ability with risk aversion. Very few used a model of probability weighting to detect the effect of cognitive ability on likelihood insensitivity and pessimism, which left room for my study.

Some studies investigated the influence of experience on risk attitude (Kunreuther et al., 2001; Gayer, 2010; Hertwig et al.2004; van de Kuilen 2009) and most involved probability weighting and explained possible mechanisms and crucial criteria. But the two distinguished components of likelihood insensitivity and pessimism did not receive much attention in those experiments and discussions. Therefore, further study of them is still worthwhile to develop detailed explanation.

Among the studies of cognitive ability and experience, some delivered some suggestions on the interaction between them (Heckman, 2006; Heckman et al., 2006; Dohmen et al., 2008). It seemed plausible that they could influence each other, because experience provides the pool of information for System 1 to refer to, whilst cognitive ability reflects the quality of combining the two systems. The interaction was thought to be complex in previous studies and was barely investigated. My study may be one of the first ones trying to investigate it in more detail.

Three hypotheses were therefore proposed for choice making under risk: (1) higher cognitive ability is correlated with a better understanding of possibilities and less subjective probability weighting; (2) more experience of dealing with probabilities is correlated with a better understanding of possibilities and less subjective probability weighting; (3) cognitive ability and experience moderate each other's influence.

An experiment including four parts was designed and it used 32 participants. The first part collected demographic information to give control, based on some benchmarks (Flynn et al., 1994; Halek & Eisenhauer, 2001; Holt & Laury, 2002; Frederick, 2005; Choi et al., 2014; Frey et al. 2015) and the necessity for my study. The second part measured participants' cognitive ability by the three-question Cognitive Reflection Test (Frederick 2005). The higher the score (range from 0 to 3 points), the higher the cognitive ability. The third part measured participants' probability weighting using price lists (Tversky & Fox, 1995; Brandstatter et al., 2002), where rare events were involved. Certainty equivalents were recorded from participants' choices and nonlinear regression was performed to analyse the data under a model of probability weighting where the level of likelihood insensitivity and pessimism were reflected by the two parameters. The fourth part provided the procedure of gaining direct experience from sampling before measuring the participants' probability weighting again.

All three hypotheses were confirmed to some extent by the results of the experiment. Among highly educated young people (the main characteristics of my sample group), they would overweight rare events. Cognitive ability was found to be negatively correlated with likelihood insensitivity. Direct experience had negative influence on likelihood insensitivity and varied influence on pessimism. Some interactions were detected and they were indeed relatively complex. Cognitive ability affected participants' willingness to sample, as well as how they learned from the experience to moderate likelihood insensitivity and pessimism. The results were compared to benchmarks, and possible applications to policy were discussed.

The rest of this thesis is organised as follows: Section 2 reviews the theories of choice under risk, two-system theory, and previous studies on the influence of cognitive ability and experience. Section 3 explains the methodology used to design the experiment. Section 4 introduces the participants, the procedure, and reports the results of the experiment in detail. Section 5 provides discussion of the results and a conclusion. In the end there are an appendix and the references.

2. Literature Review

2.1 Development of the theories of choice under risk

To model, analyse, and predict people's preferences toward risky choices, different theories have been developed. A prospect (list of outcomes with associated probabilities) is denoted as (p_1: x_1, …, p_n: x_n), where p_i is the probability of outcome x_i. Under the assumption of rationality, the optimal choice is given by the highest expected value, which is $EV=\sum p_i x_i$. The equation shows that when considering the expected value, the given probabilities (p_i) and outcomes (x_i) are directly used and no interpretations concerning specific individuals are needed. In this manner, analysing people's choices with expected value is simple and convenient.

However, this simplicity leads to its failure on explaining complex yet common phenomena which involving more factors. The St. Petersburg paradox (Shafer, 1988; Gigerenzer & Selten, 2001) is a representative phenomenon among those. I will not introduce the paradox in detail here; but in short, it indicates that the amount of money people are willing to pay to participate in a gamble, which provides an infinite expected value, is much smaller than the expected value. This phenomenon violates the theory. Also it often happens that, when indicating indifference between a sure amount and a gamble, this sure amount is higher or lower than the expected value of the gamble.

Bernoulli (1738) is known as the first one to use the notion of expected utility. Many studies toward its implications were done, and many different axiom systems were developed (Edwards, 1954). The model that von Neumann and Morgenstern developed is considered to be the best (Georgescu-Roegen, 1954). Von Neumann and Morgenstern's expected utility model (1944) alters the given objective outcomes into subjective utility. Under the assumption of expected utility, $EU=\sum p_i u(x_i)$. The utility $u(x_i)$ is translated from the given objective outcome x_i based on individual's own evaluation, and the decision is derived from the maximization of EU.

Due to the subjectively evaluated utility and the formation of EU, the optimal choices may differ between individuals, and it reflects the difference in people's characteristics through the different evaluated utilities for the same amount of outcomes. Risk attitude, as a main characteristic, is categorised into three types: risk aversion, risk neutrality, and risk seeking. These differ in the preference towards a prospect and its expected value. Risk aversion is keeping a preference for the expected value rather than the prospect. Risk neutrality is keeping indifference between the prospect and its expected value. And Risk seeking is keeping a preference for the prospect rather than its expected value. These definitions of risk attitudes hold under all utility models, including the expected utility model here and prospect theory to be discussed later. The three different risk attitudes under expected utility have utility functions in different shapes, which are concave, linear, and convex, respectively. There are various measures of risk aversion. The absolute measure of risk aversion developed by Pratt (1964) and the relative measure of risk aversion are classic ones, and the latter one is most used when assuming expected utility.

The method of expected utility received a behavioural foundation with its subjective part on utility, and most of the time it explains and predicts better than expected value. However, there are still empirical findings violating expected utility (Starmer, 2000). Among them, the Allais' paradox is representative (Allais, 1953). According to empirical findings, between the prospects of (0.1: €5M, 0.89: €1M, 0.01: 0) and (1, €1M) (M: million), most people would choose the latter one; and between the prospects of (0.1: €5M, 0.9: 0) and (0.11: €1M, 0.89: 0), most people would choose the former one. However, under expected utility, people should either choose the former for both sets of prospects, or the latter. This version of Allais' paradox is also called the common consequence effect. Similarly, most people would prefer the prospect of (1: €3000) rather than (0.8: €4000, 0.2: 0); and prefer the prospect of (0.2: €4000, 0.8: 0) rather than (0.25: €3000, 0.75: 0), which again violates the predicted preferences under expected utility. This violation is called common ratio effect. These two effects are the violations of the independence axiom, an axiom that should hold under expected utility.

To explain the violations and develop a better fitting decision model, the consideration of probability sensitivity was included. The probabilities in a prospect are also possibly to be subjectively evaluated by people when decisions are being made. The most influential model is prospect theory proposed by Kahneman and Tversky in 1979. According to prospect theory, people's choices under risk are systematically inconsistent with what expected utility theory would exhibit and predict. To be specific, when facing rare events, which means the outcomes with small probabilities, people tend to overweight the probabilities. It's for both gains and losses. In this process, probabilities are replaced by decision weights. And correspondingly, other probabilities, especially the large ones, receive lower decision weights than what the objective probabilities would imply if directly used.

Kahneman and Tversky (1979) proposed a modified model in which people make decisions by maximising: $V(p: x, q: y)=\pi(p)v(x)+\pi(q)v(y)$. $(p: x, q: y)$ is a prospect, where outcome x could be entailed with probability p and outcome y could be entailed with probability q. $\pi(p)$ and $\pi(q)$ are the decision weights for probabilities p and q, and $v(x)$ and $v(y)$ are the subjectively evaluated outcomes for the corresponding objective outcomes. When discussing the weighting function, Kahneman and Tversky assumed that the reason for ignoring or overweighting rare events and neglecting or exaggerating the difference between highly probable and certain events was that people's ability of comprehending and evaluating extreme probabilities was limited. This description suggested the possible influence of cognitive ability and experience on decision weighting which will be discussed later.

Although the model of prospect theory was better aligned with risky choices and indeed reasonably explains Allais' paradox (Kahneman & Tversky, 1979), it generated other problems. A crucial one was that when probability weighting function became nonlinear, the predictions generated from this model may violate monotonicity. Monotonicity is also called stochastic dominance, under which, lowering an outcome in a prospect would always makes it less preferable compared to the original prospect. Whereas involving nonlinear weighting function could result in opposite consequences (more explanation and discussion can be found in the book by Wakker, 2010).

To overcome this problem, Quiggin (1982) proposed rank-dependent utility, in which the probabilities weighted were the ones that allowed people to entail at least, instead of the exact number of, the given amount of outcomes. In this manner, the probabilities and the rank of the outcomes both mattered. With ranked outcomes from high to low as $x_1 \geq x_2 \geq \ldots \geq x_n$ and the prospect therefore as $\mathbf{p}=(p_1: x_1, p_2: x_2, \ldots, p_n: x_n)$, the decision weight for each outcome is $\pi_i= w(p_1+\ldots+p_i) - w(p_1+\ldots+p_{i-1})$. And accordingly, people's choices under risk should result from maximising RDU(**p**)=$\sum \pi_i u(x_i)$. Under rank-dependent utility, pessimism and optimism are used to describe the shapes of the curves of people's probability weighting function. When the function is convex, it represents pessimism (risk aversion) because the worse the rank of the outcome the more decision weight it receives. And when the function is concave, it represents optimism (risk seeking) because the better the rank of the outcome the more decision weight it receives. Also the degree of risk aversion or risk seeking can be enhanced with consideration of both the shape of utility function and probability weighting function.

Tversky and Kahneman (1992) revealed, from an experiment in which the subjects chose between various certain amounts of cash and a risky prospect in a series of price lists, that in the gains domain of risky choices, people tend to be risk seeking when coping with low probability and risk averse with high probability (in the losses domain the results were contrary). This pattern had been confirmed by several studies (Fishburn & Kochenberger, 1979; Payne et al., 1981; Cohen et al., 1987; Wehrung, 1989), and there were studies attempting to explain it with utility functions under the assumption of expected utility. However, a plausible explanation was not proposed until the probability weighting function was included. Tversky and Fox (1995) generated a typical weighting function with varied sensitivity to different possibility intervals. An inverse-S was shaped as the changes of probability in low-probability interval (close to 0) and high-probability interval (close to 1) have greater impact than the change in the middle interval.

Tversky and Wakker (1995) formalised the inverse-S shape with the concept of bounded subadditivity (SA). Their work was based on the framework of prospect theory (Tversky & Kahneman, 1992) which allows different treatments on gains and losses compared to rank-dependent utility model. They focused on gains domain of the weighting function. As they defined, "w satisfies bounded subadditivity, or subadditivity (SA) for short, if there exist constants $\xi \geq 0$ and $\xi' \geq 0$ such that $w(q) \geq w(p + q) - w(p)$ whenever $p + q \leq 1 - \xi$ and $1- w(1 - q) \geq w(p +q) - w(p)$ whenever $p \geq \xi'$ ". They called ξ and ξ' boundary constants and assumed them being independent from the objective probabilities p and q. The levels of ξ and ξ' reflect individual's characteristic difference. They also measured the degree of SA to see the degree of departure from expected utility and proposed that it was reasonable to make use of this method to measure the departure from rationality if expected utility is assumed to be rational.

Similar constants were used by Wakker (2010), but named differently as best-rank boundary and worst-rank boundary, to describe likelihood insensitivity, with which he combined with pessimism as two components of probabilistic risk attitude. Under the inverse-S shaped weighting function, the curve is divided into three regions: best rank region, insensitivity region, and worst rank region. The curve in the insensitivity region is flatter than in the best and worst rank regions, caused by people's overweighting on the extremes and insufficiency in distinguishing intermediate probabilities. It reflects cognitive limitations on understanding probabilities, and it is possible to be corrected by proper learning, which led to my research question to explore the influence of cognitive ability and experience on probability weighting towards risk and disentangle them as well as study their interaction.

2.2 The two systems theory

The two distinct systems of cognitive process were defined by Stanovich and West (2000) as System 1 and System 2. System 1 processes spontaneously and is in absence of careful attention and analysis, while System 2, in contrary, processes in need of effort, concentration, making use of knowledge, and other mental operations. In comparison, System 1 stands for intuition and System 2 stands for reasoning. Kahneman (2003) explained the distinction and interaction between the two systems, and how this would cause heuristics and biases. Different accessibility of information in people's mind could explain the difference of these two systems. Kahneman defined accessibility as "the ease (or effort) with which particular mental contents come to mind". And it is determined by how the thought is produced and evoked. The effortlessness and effortfulness of accessing certain thoughts and information can vary. When dealing with uncommon situations, or under pressure, or some other circumstances, people may use heuristic thoughts to make the thinking process easier, despite that they can possibly cause biases. If assuming that System 2 monitors System 1, which means that System 2 can correct the problematic judgment that System 1 produces, then the heuristic judgment could be explained by both System 1 failing to generate plausible judgement and System 2 failing to detect and correct it properly.

In the old model (Kahneman & Tversky, 1974), representativeness, availability, and anchoring are the three heuristics of judgment. And Kahneman & Frederick (2002) used attribute substitution to explain the possible mechanism. When facing a given objective probability, people may use a heuristic probability that is more easily to access than the original one as a substitute. Representativeness heuristics occur when people could easily access to some similar scenarios in their minds and the objective probability is rather difficult to understand, so that the possibilities of the things happening in these scenarios would be judged as the probabilities instead of the objective ones. Availability heuristics occur when people could access to some examples in their memory, and judge the probabilities basing on how the probability of different choices is attributed in these examples, instead of basing on the given objective probabilities. And therefore, due to this mechanism, the difficulty in accurately understanding rare events could rise from the lack of experience and/or deviated experience. From this mechanism it can be assumed that processors who have higher cognitive ability may be better on using System 2 to detect the intuitive attribute substitution generated by System 1, and who have been trained in statistics and have more experience in judging probability may generate less heuristic and biased judgments with System 1. Although in these situations the intuitive errors still could not be totally avoided.

2.3 Cognitive ability and choice under risk

Lubinski and Humphreys (1997) suggested that the differences in levels in general intelligence should be incorporated into the analysis of certain social phenomenon to generate better understanding of the causes for individual or groups' maladaptive or highly adaptive behaviours. Their work inspired the research investigated by Frederick (2005). He pointed out that effects of cognitive abilities on decision making, including probability weighting, risk attitude, time preference and a wide range of topics, had received little attention. The reasons are that most researchers may be more interested in average effects instead of individual differences, and for many studies the subject pools are college students so that they are more homogenous and assumed to be less varied in cognitive ability, and what is more, some terms related to cognitive ability are seen as discrimination. However, Frederick believed that as an important component in forming people's decision pattern, cognitive ability should not be ignored.

Frederick introduced a three-question Cognitive Reflection Test (CRT) as a sufficient yet simple measure of cognitive ability. He discussed the mechanism of CRT with the two-system theory. When trying to answer the questions in CRT, people may differ in processing with System 1 or 2; because the questions are designed to be looked simple at first and seemingly could be directly processed with System 1, while actually System 2 is required to generate correct answers, while the calculations of the numbers are still easy.[1] He discussed the relations between both cognitive ability and time preferences and cognitive ability and risk preferences, but here I only look into the latter one, because it is more closely related to my study.

Frederick assigned subjects scoring 0 (answering none of the questions correctly) in CRT to a "low" group and subjects scoring 3 (answering all of the questions correctly) to a "high" group. He argued that the subjects scoring 1 or 2 also gave intermediate responds to other measures about time and risk preferences in the experiment, so the analysis would only focus on the two "extreme" groups. It seems plausible when he further argued that most of

[1] The Cognitive Reflection Test will be explained more in detail in the following Methodology section.

the subjects are college students who do not differ a lot in cognitive ability in the first place and the “extreme” groups may overcome this problem. Yet there were still subjects from different sources so it is not very clear whether this technique is plausible enough. To measure the risk preferences, he did not use any model to generate parametric level, but simply listed several choices between certain outcomes (gains or losses) and outcomes (gains or losses) with given probabilities to provide a description of risk preferences.

The results showed that in the gains domain the “high” group is more willing to gamble no matter whether the expected values are higher or lower than the paired certain outcomes. Frederick suggested that the reason may be that these people do not only rely their choices on expected value. And in the losses domain, this group was less risk seeking. He also simply discussed that the preferences of “low” group in gains and losses domain are seemingly consistent with prospect theory and the preferences of “high” group are not, which pushed the application of the findings a bit closer to probability weighting. In the end he critically discussed the relation between cognitive ability and decision-making. By bringing up the issues of the existence of the “correctness” of preferences and the possible circumstances and definitions for it, he ended the discussion with a relatively neutral conclusion that despite of the characteristics of the preferences, people different in CRT scores did make different choices. His findings of the influence of cognitive ability on risk preferences are valuable even though no exact analysis or regression was performed. It provides basic support for this assumption and has inspired following researches.

Similarly, Dohmen et al. (2008) examined the relations between cognitive ability and risk aversion and between cognitive ability and impatience. Here I also focused on the former relation. There were monetary incentives in the experiments. Risk aversion was tested with real-stakes lotteries which were organized similarly to price list. Cognitive ability was measured with two tests similar to certain modules of Wechsler Adult Intelligence Scale (WAIS), and selected from the nonverbal and verbal sections for each. Characteristics of demographic, such as finance, health, and attitudes were collected and controlled with a questionnaire which was a remarkable advantage of their experiment and analysis. Their subjects were randomly selected from the adults who were living in Germany and the number was over one thousand, which as they said were representative of the population. And from this aspect their results are more plausible to be treated as general ones and would be an improvement compared to Frederick’s (2005).

Dohmen et al. found a systematically negative relation between cognitive ability and risk aversion. This result was also found in a later study conducted by Burks et al. (2009). And despite of education and financial situation, cognitive ability independently influences risk aversion. However, Dohmen et al. only measured risk aversion with choices between certain amounts of safe money and 50-50 lotteries, and therefore they did not relate their studies to the possible impact from different levels of cognitive abilities on likelihood insensitivity and attitudes toward rare events. At the same time, interestingly, in their discussion of the relation between cognitive ability and risk aversion, they brought up the possibility of a complex relation and that risk preferences may also affect the accumulation of cognitive skills throughout life time, which I assumed can be similarly referred as an aspect of knowledge or experience. They referred to the studies of Heckman (2006) and Heckman et al. (2006) which indicated that cognitive and noncognitive skills are complexly related, and argued that the component of their research was closer to the innate cognitive ability due to the selection of measurement which in some level exceeded the possible inverse influence. Here I thought Dohmen et al. showed an attitude to disentangle cognitive ability and experience, although they may not very consciously notice it and did not make further discussion.

Cokely and Kelley (2009) also showed a tendency to explain the influence from both cognitive ability and experience. According to their results, the choices made by subjects with higher cognitive reflectiveness (cognitive ability), higher working memory span (experience), and better understanding of probabilities (cognitive ability and/or experience) showed more consistency with expected value. In their experiment, subjects verbally reported their thinking process for the choices, and from this the results should be explained by the elaborative heuristic search processes instead of the computing of expected value. So that the richness of related experience of understanding and memory, as well as the ability of System 2 detecting errors, was crucial.

Benjamin et al. (2013) also took expected-value maximisation as “normative” choice, yet they measured cognitive skills with a standardised math test, and their subjects were high school students. They found less risk aversion in small stakes and less short-term discounting with subjects with higher measured cognitive ability. They also explained their results with the two-system theory. Also they noticed that although the skills vary among subjects, highly skilled people could still be biased, which supported that the intuitive errors could not be totally avoided.

With a different perspective of ‘rational choice’, Choi et al. (2014) based their analysis on comparing the ‘heterogeneity’ behaviours to expected value maximisation. Their opinion was that varied decision-making ability, which could be influenced by cognitive ability, education, or financial literacy, can lead to varied decision-making quality, and therefore would bring heterogeneity in choices. To support their opinion, they measured the consistency between subjects' choices under risk and maximised expected utility. With a large-scale panel data containing a representative sample, their analysis mainly contained three parts, focusing on the consistency and the possible causes and correlations. As they explained, the methodology they used overcame the identification and measurement difficulties in decision-making ability judging and defining. Their proposition of a suitable measure was the consistency of individual choice under risk and the Generalised Axiom of Revealed Preference (GARP), that was to test whether the data generated from the experiment could be rationalised by a utility function, and if not, how much they complied with GARP judging by the Critical Cost Efficiency Index (CCEI) developed by Afriat (1972).

Considering the most related parts with my study, Choi et al. descriptively described the consistency, and the heterogeneity was at a considerable level. And there was a positive correlation between education level and consistency. Demographic factors such as gender and age also have influence. They also discussed whether other tests of cognitive ability could be the substitute of CCEI. Although they argued that the Cognitive Reflection Test (CRT) by Frederick (2005) was not able to be a substitute when the purpose was to explain wealth accumulation; they still found that when adding the scores of CRT as a factor, the coefficient of it was economically large and the coefficient of CCEI was reduced. And there was a positive, although insignificant, correlation between CRT and CCEI. These results supported that the decision-making ability could be captured by Cognitive Reflection Test at some level, and therefore the decision-making quality could also be explained by it.

A discussion of how to define "normative" or "rational" choice should be noticed. Different studies assumed different standards for it, which, we could see from the studies introduced above, were the maximisation of expected value and the maximisation of expected utility. Since my study only focus on how probability weighting is influenced, I would not look into the value/utility part of the equation.

Apart from the studies above, some recent researches did not find a clear relation between cognitive ability and choice under risk. Eckel et al. (2012) measured cognitive ability using numeracy tests, which was part of the Educational Testing Service's (ETS) Adult Literacy and Life Skills Survey (ALL). They did not find a relation between cognitive ability and risk aversion. Similarly, in the study of Tymula et al. (2013), a numeracy test (Ofstedal et al., 2005) was used to test the correlation between the score and risk and ambiguity attitudes, and the correlation was not found. However, as I see in the above studies where the correlation was found, the absence of correlation may be explained by the measure of cognitive ability. Cognitive ability should not be limited to maths literacy and, instead of which, it should be measured by a more multi-dimensional test. In Eckel et al. (2012)'s study, the numeracy tests mainly measured the subjects' use of maths method in real life, but as Cokely and Kelley (2009) argued, during the thinking process towards risky choices, people may only use heuristic thoughts instead of computing the payoffs. Meanwhile, in another respect, in this study they found the result that higher levels of patience was correlated to less risk aversion, which showed a relation between patience and risk attitude. Patience could be related to a better use of System 2 which would exceed impulsive thinking to some extent. If it can be seen as one aspect of cognitive ability, a weak correlation was still found.

Worth mentioning is a study conducted by Petrova et al. (2014) focusing on the influence from both numeracy and emotion and their interaction used another numeracy test which was different and brought many thoughts. First, some extreme probabilities were included in their experiment, for example, 1% and 99%. Following this, it was shown from the results that their subjects overweighted small probabilities and underweighted large probabilities, which confirmed the inverse-S shaped probability weighting function. Second, they did not find a correlation between numeracy level and probability weighting, but they were partly associated. And meanwhile, with a reappraisal process of emotion, high numeracy did have an influence on subjects' decision to be more in line with a normative model, which suggests that there was an interaction between emotion and numeracy. Third was that the Berlin Numeracy Test (Cokely et al., 2012) they used was a unique and efficient numeracy test which specifically focused on people's understanding of probabilities.[2]

Andersson et al. (2013) studied the relation between cognitive ability and risk preference (mainly risk aversion in their study), and argued that the noisy decision making made it difficult to identify this relation. Because with erred choices, the estimation of the risk attitude may be deviated. They used 50-50 gambles and varied payoffs in their price lists to minimise probability weighting. The subjects in their study did not face rare events. They used a cognitive ability test called IST 2000 R, which showed essentially similar results compared to CRT. They compared different tasks by which the level of risk aversion was measured, and varied correlations with cognitive ability were shown from these tasks. They suggested that the mistakes people would make differed on the directions according to the design of the tasks, and this bias could explain the variation of correlation depending on the tasks. In further analysis, they found that cognitive ability was correlated to noisy behaviour instead of risk attitude. Their opinion was that the human nature of making mistakes should be taken into account when interpret the possible relation. In their discussion, they also mentioned the correlation between education and risk preference and suggested the possible correlation was also led by the relation between education and noisy choice making.

[2] The Berlin Numeracy Test will be introduced in more detailed in the following Methodology section.

Based on the review of these studies of the relation between cognitive ability and risk attitude, and with the assumption of prospect theory, my first hypothesis is that people with higher cognitive ability will have a better understanding of (extreme) probabilities and be less spontaneous when making decisions and using heuristics, and less overweight rare events so that their probability weighting functions are less inverse-S shaped.

2.4 Experience and choice under risk

Kunreuther et al. (2001) studied on people's tendency to neglect on the negative events which have extremely low probability to happen but would lead to high consequences of accidents or disasters once happened. These events are similar to the losses domain of choice under risk. Although their focus was on the losses domain, some results and conclusions are still essential and can be introduced to explain some corresponding phenomenon in the gains domain of choice under risk. They based their study on the concept of evaluability (Hsee, 1996; Hsee et al., 1999), under the hypothesis of which, it would be difficult for people to evaluate the given attribute (can be seen as a prospect) if they could not compare the attribute with what is meaningful to them. And therefore it provided an explanation of why people were empirically found to hardly interpret well on low probabilities. Kunreuther et al. proposed that with richer information and perceiving a feeling of the risk by themselves would help people to be more sensitive to the likelihood of those rare events. And I interpret their usage of the word "sensitive" here as less underestimating or less underweighting the low probabilities of losses.

In the experiments, Kunreuther et al. tested how to more effectively improve people's understanding of the probability of the accident of toxic chemical release. They used comparison scenarios which were different in the familiarity to the subjects, in the level of the low probability, and in the richness of the content to test whether these factors would help and influence subjects' evaluation on the probability of the given rare event. With results from five studies all together, they concluded that the extent of similarity of the probabilities and the relevance of the given event and comparison points do not have influence; but when having to compare several low-probability events together, the comparison points get more helpful. Also, the comparison scenarios should provide richer information in their content of the interpretation of how the qualitative feelings and quantitative scale of the risk are associated. And therefore people could learn more from the comparisons and understand the low probabilities better.

Kunreuther et al. did not compare the situations with and without the procedure of learning. Also they did not combine their findings with probability weighting function or any other choice models. But still, their research well discussed that people do respond to a proper learning procedure and how to effectively provide information for people to get experience of rare events and understand low probability. Their results should also be applicable to the gains domain of choice under risk, that people would be able to learn from given references and develop a better understanding of the probabilities of rare events of gains such as a lottery, as long as the references are rich in their content.

Gayer (2010) discussed the mental process of comprehending and evaluating objective probabilities. Similar to Kunreuther et al. (2001), she held the idea that there was a learning process for people to evaluate probabilities. The differences were that she did not only focus on rare events but also on probabilities in a general scale, and the source of learning she based on was people's own experience and memory. Due to the size of the experience pool and accessibility, evaluating objective probabilities from experience may be biased and therefore lead to the distortion of the probabilities. She suggested that the experience from the past, the realisation of the probabilities, and the similarity between the objective given probabilities of past experience and the risk being evaluated are three main components within this process.

To be specific, when facing a familiar probability that is easy to be accessed in their memories, people recall both the experience with identical probabilities and the experience with similar probabilities to evaluate this given probability. And when facing an odd or unfamiliar probability, they may recall the experience which has similar situation and try to evaluate the probability by learning from it. This could be explained by representativeness heuristics. She argued that the distortion of probabilities, especially the overweighting of rare events, is possibly caused by this procedure. And since different individuals have different accessibility for experience, it would be natural that the intersections of their weighting functions and the diagonal are varied and flexible. She also mentioned that the availability heuristics play an important part. People's memories have many sources including their own life experience as well as their families' and friends' experience and news from the media and so on, and out of this the extreme information would tend to receive more attention. Thus, people may be biased when comprehending the low probabilities from these kinds of experience.

From the perspective of how the mechanism works when facing risky choices with unknown probabilities and subjects could only learn from experience, Hertwig et al. (2004) developed their study basing on the empirical findings which indicated the tendency of underweighting rare events, instead of overweighting which the prospect theory would predict. They defined decision from description and decision from experience as two different methods of how the risky choices are provided. Decision from description is for the respondents to choose between the risky choices that are described visually with the payoffs and their probabilities, and generally they could only choose for one time. Under the situation of decision from experience, the risky choices are presented in such a way that the respondents could make repeated choices and receive feedbacks to learn the distributions. They suggested that the latter one would be a better reflection of the situations in the real world where the description of the risky choices are always absent and people have to make the decision basing on their experience. The underweighting of rare events is more often to be found in the experiments designed with decision from experience. They disentangled the effects from direct experience and repeated decisions, and the results showed that direct experience is the main reason.

The methods people use to collect information and how the information affect their decisions are two crucial factors to consider. Hertwig et al. suggested that the subjects in their experiment made their decisions from experience basing on the information they received from sampling, and therefore the sample size would have a great impact. Thumb rule indicates that the frequency of the rare events happening would be even smaller than expected (considering the objective given probabilities) if the sample size is small in binomial distributed risky choices. And therefore the probability of the rare events would be underestimated/underweighted and the rare events may be even less preferable due to the limited sample size.

Also, more recent sampled choices would have greater impact on people's decisions, so that even with large samples the rare events could be underweighted, due to the lower likelihood of the rare events to occur recently comparing to the common events. To test this assumption, Hertwig et al. divided the data from the experiment into two halves following the sequence of sampling, and found that the samples from the second half (more recent ones) have higher ability in predicting the subjects' choices. And therefore they concluded that limited sample size and the recency effect could both lead to underweighting of the rare events under decision from experience. From their conclusion, I would assume that the sampling process reduces the weights that the respondents put on the rare events. If combining decision from description and decision from experience, the overweighting of rare events in decision from description would be reduced by sampling, while there would be less tendency of underweighting because the probabilities are provided visually to the respondents.

Van de Kuilen (2009) assumed subjective probability weighting and how it could be reduced by a learning process. The foundation of his study is the Discovered Preference Hypothesis

(DPH) developed by Plott (1996). This hypothesis assumes that when subjects in the experiment are able to deliberate and learn from sufficiently provided opportunities and incentives, their preferences would reach consistency. From this, van de Kuilen interpreted that DPH predicts that the subjects' preferences would be consistent with 'the normatively compelling axioms underpinning' expected utility theory, which means that learning from experience could help the subjects to make such choices that are more consistent with the maximisation of expected utility. This sufficient learning process is considered to be formed by both repeatedly making choice under similar scenarios, and direct feedback of the choices.

Aiming to test the influence of direct feedback, van de Kuilen conducted an experiment containing two different treatments: terminal outcome feedback (without direct feedback) and ongoing outcome feedback (with direct feedback). The results showed that the influence of direct feedback is significant. Under the terminal outcome feedback treatment, subjects' probability weighting function continues to deviate from linearity. But under the ongoing outcome feedback treatment, part of the probability weighting function does not significantly differ from linearity after the whole learning process. It could be interpreted that the learning process with direct feedback could help the subjects to make choices that comply better with the maximisation of expected utility, while the absence of direct feedback would not have this effect. Despite of the opinion on learning, his experiment and results provide a perspective that the factors of repeated choices and direct feedback as experience could be included in an experiment with related topics and how it could be conducted.

From the researches above and the assumption of prospect theory, my second hypothesis is that people who are more experienced in dealing with probabilities and risk events would have a better understanding of (extreme) probability; and if they use heuristics, there would be a larger pool of examples providing more similar situations with the given objective probabilities. And therefore, they would less overweight rare events so that their probability weighting functions are less inverse-S shaped. Also, there may be an interaction between cognitive ability and experience, taking the suggestions of Dohmen et al. (2008), Heckman, (2006), and Heckman et al. (2006). It is possible that people with higher cognitive ability could learn more from their experience, and treat the experience more critically. So that even when they use heuristics to solve the problem, they may try to use System 2 more often and better to detect. So my third hypothesis is that higher cognitive ability would strengthen the influence of experience.

To conclude, the research question of my study is to disentangle the influence of cognitive ability and experience on probability weighting towards risk, and detect the interaction between cognitive ability and experience. My hypotheses are:

1) Cognitive ability is correlated with probability weighting. With higher cognitive ability, people will less overweight rare events so that their probability weighting functions are less inverse-S shaped;
2) Experience is correlated with probability weighting. With more experience, people will less overweight rare events so that their probability weighting functions are less inverse-S shaped;
3) Cognitive ability and experience are interacted. Cognitive ability will enhance the influence of experience.

3. Methodology

With the research question and three hypotheses, a laboratory experiment was conducted to collect subjects' demographic information, measure subjects' risk attitude, measure subjects' cognitive ability, and provide the subjects chances to gain experience with probability.[3] With the data collected from the experiment, data analysis was conducted to explore the relations.

3.1 Demographic questions

Many researches have found demographic influences alongside their studies on risk attitude (Flynn et al., 1994; Halek & Eisenhauer, 2001; Holt & Laury, 2002) or the relations between cognitive ability and experience and choice under risk (e.g. Frederick, 2005; Choi et al., 2014; Frey et al. 2015). For example, in most cases males were found to be more risk seeking than females (Croson & Gneezy, 2009), and elderly people were found to perform worse than younger people in cognitive ability tests and therefore would influence the following analysis and comparison. To give more control to the experiment and the data analysis, demographic information including age, gender, nationality, education, and field of study/job were collected from the subjects during the experiment. Along with giving control, education level could be included as a signal of intelligence, and the field of study/job could capture the subjects' previous experience in dealing with probabilities.

[3] The procedure of the experiment will be introduced in the following Experiment section.

3.2 Measure of probability weighting

Under the assumption of prospect theory, I measured subjects' probability weighting function with price lists. It provides researchers with the factors they need to build the model and study the subjects' risk attitude. Since part of the main purposes of my study was to detect whether the subjects would behave with likelihood insensitivity and overweighting rare events, and how their probability weighting could be influenced by cognitive ability and experience, payoffs with extremely low and high probabilities needed to be included.

The design of the price lists was inspired by Brandstatter et al. (2002). Their price lists in one experiment were based on the one outlined by Tversky & Fox (1995), and another experiment about gambling included a series of extreme probabilities (rare events). I combined the two ideas and designed the price lists in my experiment as shown in Figure 1 and 2 below as two examples. There were seven lists in total, with varied probabilities (p = 0.01, 0.05, 0.1, 0.5, 0.9, 0.95, 0.99). In each list, Option A was varied sure amounts of money from €0 to €100 that the subjects could entail, and Option B was a gamble with one of the probabilities with which the subjects could entail €100. In the first row, the secure amount of money given by Option A was always €0, while the outcome from Option B would be either €0 or some amounts that is higher than €0 so that in the first row Option B was always better off. And in the last row, the secure amount was always €100, which was obviously better off compared with the gamble provided by Option B. And when the secure amount of money growing higher and higher from €0 to €100, at some point before reaching the last row, the subjects' preference would switch from Option B to Option A depending on their own evaluation. So the subjects were asked to choose between which two rows their preferences were switched from Option B to Option A. The switching points were recorded to calculate the Certainty Equivalent from each Price List. Each Certainty equivalent was the average of the two secure amounts of the two rows.

One thing to be noticed is that, among these price lists, the gaps between every secure amounts in each list were not necessarily the same, and the gaps were smaller around the expected value calculated with the payoff and probability of Option B. For example, if comparing the two price lists in Fig. 1 and 2, in the first list the secure amounts in the first five rows have smaller gap, and the expected value calculated with the prospect is €1. Meanwhile, in the second list, the five secure amounts around €5, which equals to the expected value, have a smaller gap between each other. This was designed to fit with the rare events which have relatively very small or big expected values, and subjects could be more sensitive with their choices in these intervals.

	OPTION A	OPTION B
1	€0 for sure	€100 with the probability of 0.01
2	€0.5 for sure	€100 with the probability of 0.01
3	€1 for sure	€100 with the probability of 0.01
4	€1.5 for sure	€100 with the probability of 0.01
5	€2 for sure	€100 with the probability of 0.01
6	€5 for sure	€100 with the probability of 0.01
7	€15 for sure	€100 with the probability of 0.01
8	€25 for sure	€100 with the probability of 0.01
9	€50 for sure	€100 with the probability of 0.01
10	€75 for sure	€100 with the probability of 0.01
11	€100 for sure	€100 with the probability of 0.01

Figure 1. The first price list.

	OPTION A	OPTION B
1	€0 for sure	€100 with the probability of 0.05
2	€2.5 for sure	€100 with the probability of 0.05
3	€5 for sure	€100 with the probability of 0.05
4	€7.5 for sure	€100 with the probability of 0.05
5	€10 for sure	€100 with the probability of 0.05
6	€15 for sure	€100 with the probability of 0.05
7	€20 for sure	€100 with the probability of 0.05
8	€40 for sure	€100 with the probability of 0.05
9	€60 for sure	€100 with the probability of 0.05
10	€80 for sure	€100 with the probability of 0.05
11	€100 for sure	€100 with the probability of 0.05

Figure 2. The second price lists.

3.3 Measure of cognitive ability

There are many varied methods of measurements that could capture subjects' cognitive ability, for example, IQ tests, numeracy tests, and so on. Each measurement could have different focuses, different complexities, different lengths, and therefore they could suit into experiments with different aims. In my study, the theoretical assumptions required the method of measurement to be able to reflect the two systems theory instead of only numeracy. Besides it, what also needed to be taken into consideration was that the whole process of the experiment contained many sections so it was already time-consuming; and additionally there was no actual payoff awarded to the subjects and they were not gathered in a lab but stayed at their own places which may cause some distractions during the experiment. As a result, an ideal measure of cognitive ability for this experiment should be time-saving and not too complex.

The Cognitive Reflection Test (CRT) was introduced by Lubinski and Humphreys (1997) and discussed in great detail by Frederick (2005). This test is formed by three questions as shown in Figure 3 below. These are maths questions that are easy from the perspective of computing, but they consume more effort and deliberate thinking to be understood and answered correctly. At first sight the questions may lead the participants to have a quick reaction and answer wrongly if they fail to take some effort to avoid being tricked. For example, in Question 1, an immediate answer may be 10 cents, but additional thoughts would be needed to detect the problematic impulse reaction and solve it with the correct answer which should be 5 cents. This process reflects the theory of two systems (Epstein, 1994; Sloman, 1996; Stanovich & West, 2000; Kahneman & Frederick, 2002), where System 1 is intuitive and effortless thinking without much attention, and System 2 is deliberate and effortful thinking that consumes more attention. In this manner, higher cognitive ability could be judged by using System 2 more appropriately and avoiding the possible mistakes made by System 1. Reviewing this measurement with the previous assumption, if the subject answers wrongly, then it could either be caused by the dependence on System 1 while lack of using System 2, or a fail of System 2 detecting the mistakes. Either way it captures a lower level of cognitive ability. The score is from 0 to 3, representing that a participant answers 0 to 3 questions correctly. And it is assumed that the higher the score one earns in this test, the higher cognitive ability it reflects.

Cognitive Reflection Test (CRT)

Question 1: A bat and a ball cost €1.10. The bat costs €1.00 more than the ball. How much does the ball cost? ___ cents

Question 2: If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets? ___ minutes

Question3: In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake? ___ days

Figure 3. The questions in CRT

Frederick (2005) compared CRT with the Wonderlic Personnel Test (WPT) which is used by employers to test the intellectual abilities of their candidates, "need for cognition" scale (NFC) which asks self-assessed questions related to cognitive ability, and the scores on Scholastic Achievement Test (SAT)/American College Test (ACT) which are college-entrance examinations. In the results of comparisons, in general, these five measures are positively correlated with each other, and all measure some factors in common. But they at the same time have specifically distinguishing focuses. When it comes to the validities of predictions on time and risk preferences, CRT is always the best or second-best one to predict. And obviously this three-question test is much more concise and time-saving compared to other measures. With the consideration of being part of a rich-contexted and time-consuming experiment, measuring the subjects' cognitive ability with CRT would be a reasonable choice.

Another measurement which could also make sense in this situation would be the Berlin Numeracy Test (Cokely et al., 2012; Petrova et al., 2014). It was based on the previous work of Schwartz et al. (1997) and Lipkus et al. (2001), and formed by four maths questions as shown in Figure 4 below. The format and testing process are also brief and time-saving compared with CRT.

Berlin Numeracy Test

Question 1: Out of 1,000 people in a small town 500 are members of a choir. Out of these 500 members in the choir 100 are men. Out of the 500 inhabitants that are not in the choir 300 are men. What is the probability that a randomly drawn man is a member of the choir? Please indicate the probability in percent. ___

Question 2a: Imagine we are throwing a five-sided die 50 times. On average, out of these 50 throws how many times would this five-sided die show an odd number (1, 3 or 5)? ___ out of 50 throws.

Question 2b: Imagine we are throwing a loaded die (6 sides). The probability that the die shows a 6 is twice as high as the probability of each of the other numbers. On average, out of these 70 throws how many times would the die show the number 6? ___ out of 70 throws.

Question 3: In a forest 20% of mushrooms are red, 50% brown and 30% white. A red mushroom is poisonous with a probability of 20%. A mushroom that is not red is poisonous with a probability of 5%. What is the probability that a poisonous mushroom in the forest is red? ___

Figure 4. The questions in Berlin Numeracy Test

Cokely et al. (2012) tested the predictive validity of this measure, and found that its predictive power was superior to other numeracy tests, and CRT was tested to possess similar predictive power. However, since the targeted subject groups are educated and highly educated people, and the main purpose is to measure these people's understanding of the information about risk, the questions are all related to probabilities that are not very easy. So that the skills of solving probability are required to correctly answer them. The thinking process towards these questions is not related to the struggle between the two systems, while only effortful thinking is needed here. So under the assumption of two system theory, CRT is more related to it compared to Berlin Numeracy test. The latter one could rather capture numeracy better. Some researches (i.e Cokely et al., 2009; Liberali et al., 2011) demonstrated that CRT could capture the ability of information searching and encoding, but the predictive validity of numeracy on search and encoding was not clear. Also from a practical point of view, the score deliberating is relatively complex for the researcher, and therefore the designing of the program for computer-based experiments. So in my experiment CRT was a more appropriate measure of cognitive ability. And I expected that the subjects with higher cognitive ability would less overweight rare events and their probability weighting would deviate lessfrom expected utility.

3.4 Gaining experience

Except for subjects' possible real-life experience of probabilities captured by the demographic question of the fields of their study or job, I also provided the subjects with the chance to directly experience the given probabilities before they make their decisions. In Hertwig et al.'s (2004) study on decisions from experience and risky choice under rare events, they introduced Barron & Erev's (2003) and Weber et al.'s (2004) methods of letting subjects experience the unknown probabilities before they make decisions on the choices. The treatment group of their experiment was designed for the subjects to try to learn, or get a sense, of the probabilities by clicking buttons and sampling the outcomes before they made final decisions. And the subjects in the other group learned the probabilities just descriptively. Also rare events were included in their experiment to explore subjects' attitude under the two situations.

I used a similar setting where the subjects also needed to sample and experience the probabilities before they made the decisions, and after each time they clicked the button, an outcome following the distribution would be shown. There were some adaptations to make the process better fit into the purpose of this experiment. To make sure that the subjects had sufficient sampling size, the least number of times of sampling was set to be 5, and they could sample as much as they felt like. The amount of money and the probability to entail the money were shown on the button, so that they were known to the subjects. It was due to the setting of this part of the experiment. The analysis on this part was designed to be within-subjects and the subjects had already known the probabilities from the previous section of Price Lists. The purpose was to explore whether the subjects could learn and develop some more sense of the probabilities after sampling, compared with the decisions they had made before sampling.

The controlling minimum times of sampling and giving direct feedbacks after every time of sample echo the opinion of Van de Kuilen (2009), that with these together with the "learning process", people would be able to make choices more aligned with the maximisation of expected utility. I expected the subjects to less overweight the rare events after sampling, compared to their decisions before.

4. Experiment

4.1 Participants

There were 32 people participating in the experiment. They were recruited from my contacts either at Erasmus University Rotterdam, or through internet. The range of the age of the participants is from 19 years to 31 years, with an average of 22.75 years. There were 22 (68.75%) female participants and 10 (31.25%) male participants. All participants were non-Dutch in nationality. All participants had received high education, among which, 11 (34.38%) reached Bachelor level and 21 (65.62%) reached Master level. The fields of their studies/jobs included public policy, law, computer science, economics, accounting, business administration, finance, media, physics, information system, mathematics, literature, engineering, and psychology, etc. To detect the relations, these fields were divided into two categories based on the requirement of probability knowledge. There were 24 (75%) participants who dealt with probability more often in their fields, and 8 (25) less often.

	Observation	Mean	SD	Min	Max
Age	32	22.75	2.24	19	31
				Observation	Proportion
Gender					
Female				22	68.75
Male				10	31.25
Nationality					
Non-Dutch				32	100.00
Dutch				0	0.00
Education					
Bachelor				11	34.38
Master				21	65.62
Field of study/job					
Probability-related				24	75.00
Probability-unrelated				8	25.00

Table 1. Demographic information (N=32)

4.2 Procedure

A questionnaire (Appendix A) was designed and programmed to be available online for the participants to fill in, and the data was directly collected from the answers. The data was recorded anonymously. The participants were aware that it was an economics experiment and that the purpose of the experiment was to collect data for my master thesis. The participants either filled in the questionnaire using my computer, or received the link of the website and used their own computers; either way, I was available for the participants to ask questions if they needed help.

The full questionnaire contained five parts in total, and before each part an introduction was provided. Part One asked for demographic information including age, gender, nationality, education, and field of study/job, in which the participants needed to choose or fill in the answers that described them the best.

Part Two was the Cognitive Reflection Test, where the three questions were shown separately, and the participants were required to type in their answers after reading each question.

Part Three contained seven price lists. The participants were required to choose their switching point from Option B (a gamble with a given probability) to Option A (sure amounts of money which became better and better). The given probabilities of the risky choices in the seven price lists are 0.01, 0.05, 0.1, 0.5, 0.9, 0.95, and 0.99, and the participants saw one price list at a time. Before the first list was shown, it was explained to the participants that how to state the preferences by making the choice of switching point. Also, the participants got a chance to choose in an “example” price list with the given probability 0.2 as practice.

Part Four combined sampling and price lists. The price lists were the same with the ones in the previous part, only before each price list a button for sampling the risky choice was shown on the screen. The risky choices were written on the buttons and the participants were required to click each button for at least five times to sample. The result of the gamble was shown after each click, being either €0 or €100 following the distribution of the given probability. A practice example was also given in the introduction, with a button sampling the given probability of 0.2 before the example price list was shown.

Part Five contained two questions, asking whether the participants had participated in any economics experiment before and whether they had seen (parts of) the CRT questions, to give more control.

It was a hypothetical experiment and no payment or award was implemented. Studies had discussed the usage of hypothetical choices. Kahneman and Tversky (1979) suggested that hypothetical experiments could be applied to risky choices with high stakes. While Holt and Laury (2002) found it in their research that the participants were less willing to take risks in the real-payoff scenarios compared to hypothetical payoffs; and the different attitudes toward risk was even more distinguishable when facing high payoffs. Some studies (e.g. Schoemaker, 1990; Hogarth and Einhorn, 1990) suggested that when it comes to the gains domain, there would be no or only little difference on the risk attitude between real and hypothetical experiments. Considering the budget and the fact that the risky choices in my experiment were only related to gains domain and the payoffs were only very high, a hypothetical experiment could be reasonable.

The relation between risk attitude and cognitive ability was detected and compared between participants with different cognitive-ability levels. The relation between risk attitude and experience was detected mainly by within-participants investigation, where the change of decisions was measured with decisions made before and after sampling. According to Kahneman (2003), dual-task methods are recommended as one possible design to study System 2, for example to test whether the intuitions could be modified by the intervention of System 2, to be more specific, the effects of training and intelligence. And therefore it could be an ideal method for my study with sampling and cognitive ability. However, it had to be taken into account that there could be a possibility of either anchoring on intuition or overcorrecting the judgment from it.

4.3 Results

4.3.1 Analysis

Before any further data analysis and comparisons can be performed, the parameters in the probability weighting functions needed to be generated. In this study, I mainly followed the procedure of Petrova et al.'s (2014) data preparation method, because their experiment shared similar characteristics with mine.

They first used an equation to normalise certainty equivalents in the range from 0 to 1, and equalled it to probability weighting w(p):

$$w(p)=\frac{CE(p)-min}{max}=\frac{CE(p)-CE(0)}{CE(1)}.$$

The connection with w(p) is that, with the prospect of (p: x, 1-p: 0), CE=u^{-1}w(p)u(x), and assuming a simple utility function u(x)=x, we get w(p)= $\frac{CE(p)}{x}$; the CE for gaining x with 0% to be 0 (CE(0)=0), and the CE for gaining x with 100% to be x (CE(1)=x), then the above equation of w(p) and normalised CE holds. Using this equation in the situation of my experiment, w(p)= $\frac{CE(p)}{100}$[4].

As for the model of the probability weighting functions, there are several with different advantages. The one I used here is a two-parameter function relatively widely discussed and used (Goldstein & Einhorn, 1987; Lattimore et al., 1992; Ostaszewski et al., 1998; Gonzalez & Wu, 1999; Petrova et al., 2014):

$$w(p) = \frac{bp^a}{bp^a+(1-p)^a}.$$

Parameter *a* reflects likelihood insensitivity: when *a* increases, likelihood insensitivity decreases; and the shape of the curve changes from a strong inverse-S (0<*a*<0.69) to less strong (0.69<*a*<1) to S-shaped (*a*>1), while *a*=0.69 shapes an inverse-S fitting commonly found data, and *a*=1 means no S shape. Parameter *b* reflects pessimism/optimism by affecting the elevation of the curve: when *b* increases, pessimism decreases/optimism increases; and *b*=0.77 fits common data. *a*=0.69 and *b*=0.77 will be called "common finding" in the following discussions. Two sets of parameters (before and after the process of sampling, and will be called "original" and "post-sampling" in the following discussions) for each participant were estimated in Stata, using the procedure of Nonlinear Least Squares (NLS).

After the regressions were performed, the data collected from the original choices of eight participants was found to be not covered by the model. And the data collected from the post-sampling choices of four participants was also at this situation. These participants were automatically dropped by Stata during the following analysis. However, one thing to be noticed was that, among these participants, the choices of two were not able to be covered for both original and post-sampling; while six were originally could not be covered but later

[4] The equation selected to perform the analysis in my study was w(p) = $\frac{CE(p)}{CE(0.99)+CE(0.01)}$ in the end, following the same form of Petrova et al.'s (2014). They assumed that the CE for gaining €100 with 100% could be substituted by the sum of the CEs of gaining €100 with 99% and 1%. Also, as I observed, the sums of CE(0.99) and CE(0.01) were very close to CE(1); so the final results would barely deviate from the analysis based on w(p)= $\frac{CE(p)}{100}$.

on could be covered for their post-sampling choices; and two were originally could be covered but later on could not be covered for the post-sampling choices.

4.3.2 General results

With the obtained a_i and b_i (original parameters), and a_{si} and b_{si} (post-sampling parameters), we could roughly interpret from Table 2 and Figure 5 that the mean, as well as the standard deviation, of a_s was a large increase from *a*. On the other hand, there was only a small change between the means of b_s and b. These could suggest that the level of likelihood insensitivity may be more easily influenced by experience (sampling), while the effect on pessimism ay be little.

Parameter	Obs.	Mean	SD	Min	Max
a	24	0.58	0.31	0.03	1.16
b	24	1.01	0.32	0.49	1.90
a_s	28	1.12	0.62	0.03	2.24
b_s	28	0.98	0.48	0.16	2.20

Table 2. Summary of parameters a, b, a_s, b_s.

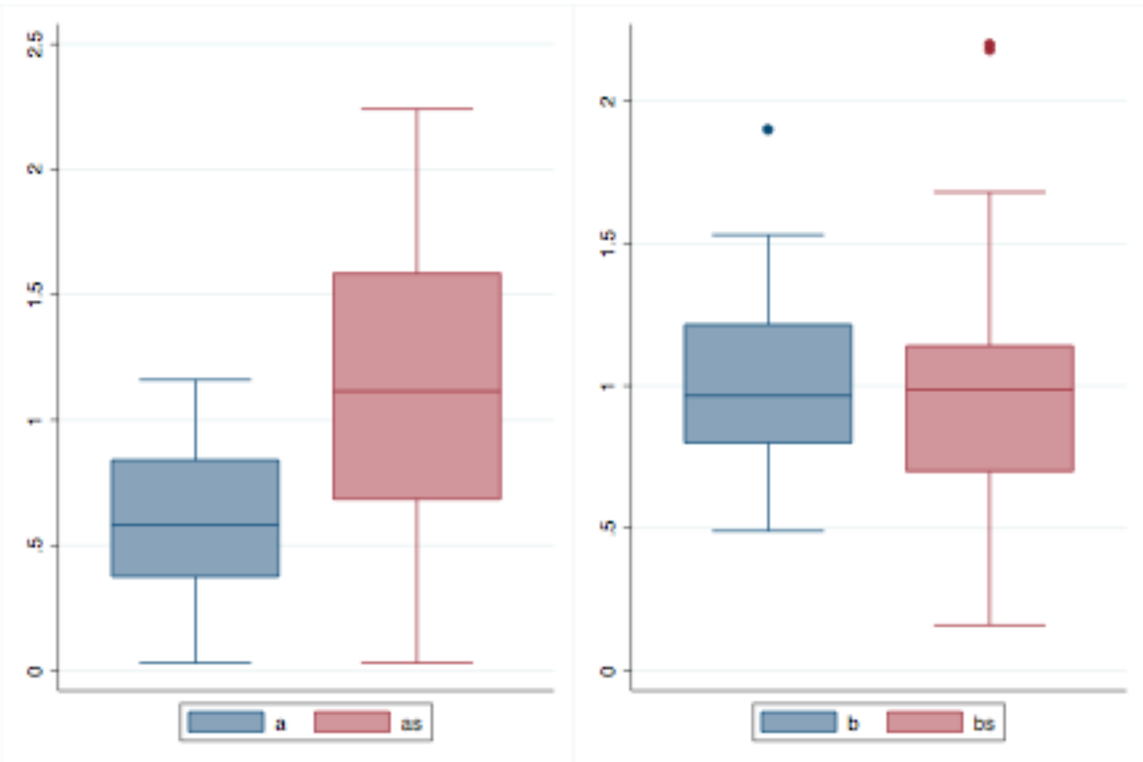


Figure 5. Mean estimates of a and a_s, b and b_s.

Taking the means of a and b (original condition) and a_s and b_s(post-sampling condition), two probability weighting functions were fitted with the model to be compared with the common finding where a=0.69 and b=0.77 (Figure 6). It is obvious from the values and the graph that the original condition was the most likelihood insensitive and the least pessimistic among the three. The post-sampling condition was a lot less likelihood insensitive and was close to linear with even a bit underweighting of the rare events, while the level of pessimism was not much different from the original one. The deviation between the two conditions with the common finding on pessimism may be caused by the use of hypothetical choices instead of real ones, by which the participants may be more willing to take risks compared to the latter. The deviation on likelihood insensitivity between the original condition and the common finding was not significant (p=0.10), while the one between the post-sampling condition and the common finding was much larger and significant (p=0.001), and the shapes of the curve even differed so much with the comparison of inverse-S and S, which could represent a big influence of experience.

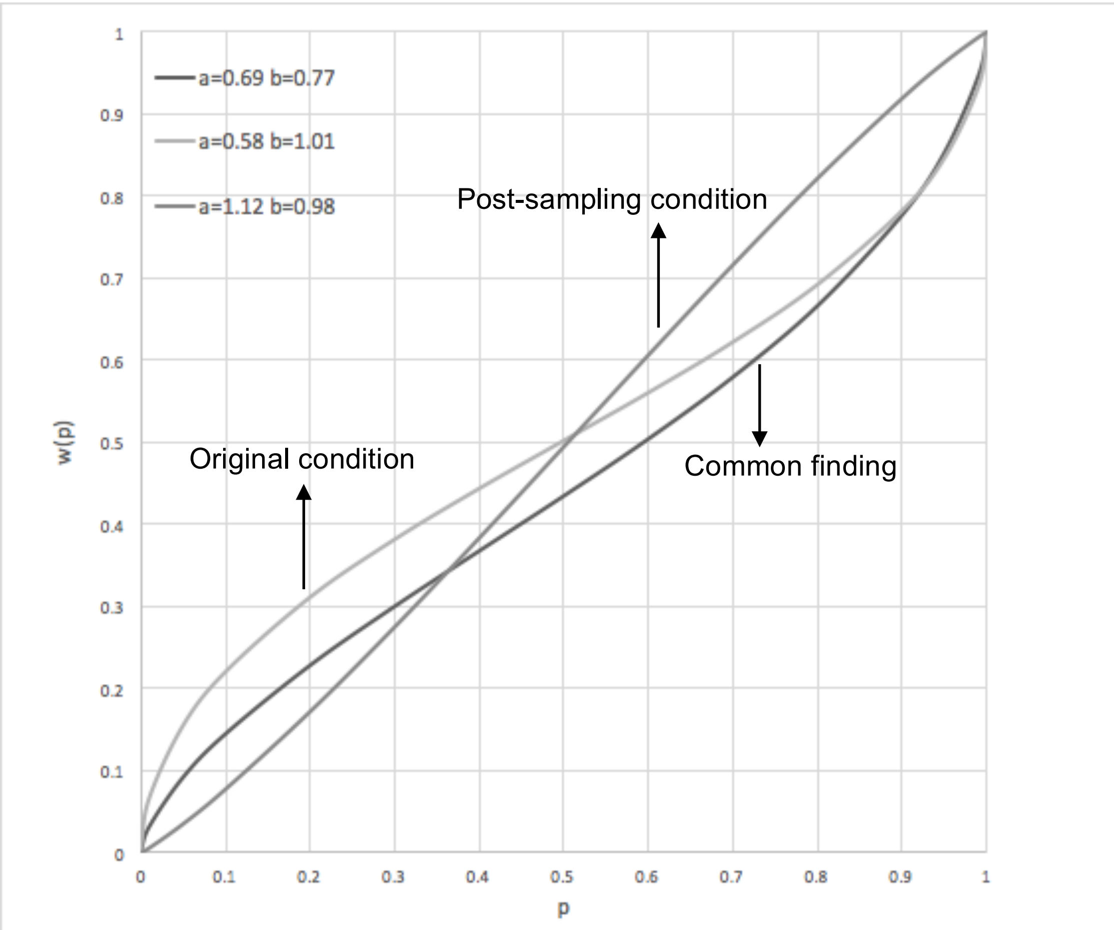


Figure 6. Three fitted probability weighting functions.

4.3.3 Cognitive ability

To study the relation between cognitive ability and risk attitude, CRT was measured for each participant. First take a look at participants' performance in the test (Table 3). The average score among the participants was 1.78, with almost half scoring 2. The second most was 8 participants scoring 3, followed by equally least amount of participants scoring 0 or 1. This distribution may be a result of the generally high education level of the whole group.

	Observation	Mean	SD	Min	Max
CRT	32	1.78	1.01	0	3
Scores				Observation	Proportion
0				5	15.62
1				5	15.62
2				14	43.75
3				8	25.00

Table 3. Summary of CRT scores.

Second, to give more control on the performance of answering the three questions, we needed to see whether the participants' previous knowledge of CRT would influence their scores. A Mann-Whitney U test was performed to detect it, and the null hypothesis that the participants had or had not seen (part) of the CRT questions would score the same in the test was not able to be rejected with over 95% confidence (p=0.57) (see Appendix B). And therefore, the participants' knowledge of CRT did not influence their scores.

Linear regressions were then performed to detect the relation between cognitive ability and likelihood insensitivity (parameter *a*) and pessimism (parameter *b*). Besides the score of CRT, variables including age, gender, education, field of study/job, and experience of economics experiment were used to give control. The results showed that, for likelihood insensitivity no variable had significant a effect on *a* (p>0.05), except for scoring 2 (p=0.01) or 3 (p=0.02) points in the CRT test which had positive correlations. Compared to scoring 0, participants who scored 2 would generate a 0.65 higher *a*, and those who scored 3 would generate a 0.48 higher *a*, ceteris paribus. It could be interpreted that participants with high cognitive ability were less likelihood insensitivity. Also the different scores of CRT were tested to be jointly significant (p=0.01) of the effect on *a*. The same procedure was applied again on *b*. Unfortunately, no variable was found to significantly (p>0.05) influence pessimism, whereas the scores of CRT were not jointly significant (p=0.20). This could mean that among these participants, their cognitive abilities did not influence their level of pessimism. (Appendix C)

The later conducted Mann-Whitney U tests echoed these results. For the convenience of using this test, participants with different levels of cognitive ability were categorised into two groups. Participants scoring 0 or 1 were assigned to the group of low cognitive ability, and those who scored 2 or 3 were assigned to the group of high cognitive ability. The tests were performed to detect whether participants showed different likelihood insensitivities and pessimism if they were from different cognitive ability groups. What we could interpret from the results of the tests was that the likelihood insensitivities between low and high cognitive ability groups were different (p=0.003), while on pessimism it did not show a difference (p=0.62). (Appendix D)

To conclude this part, the effect of cognitive ability on likelihood insensitivity was significant but on pessimism it was not. Participants with higher cognitive ability tended to be less likelihood insensitive, that was to say more likelihood sensitive. The shapes of their probability weighting function were likely to be less inverse-S, compared with the participants with lower cognitive ability. These suggested that high cognitive ability may help participants to better understand probabilities and less overweight rare events.

4.3.4 Experience

Before analysing the effect of sampling, the information of education level and field of study/job was used as participants' general experience of probability, in order to explore its relations with risk preferences. The linear regressions in the previous part suggested no significant correlation, and education and field were tested to not be jointly significant (p=0.59 on *a*, p=0.42 on *b*). The Mann-Whitney U tests also suggested the same. Among the participants, possessing either a Bachelor or Master degree, or studying/working in a field requiring either more or less knowledge of probability did not mean that they were different in their likelihood insensitivity (p=0.91 of education, p=0.83 of field of study/work) and pessimism (p=0.77 and p=0.86) levels. (Appendix E)

In Table 4 the times participants took to sample before each Price List were summed. The tendency of how much effort they were willing to make was rather obvious. At the beginning, facing the rare events with probabilities of 0.01 and 0.05, participants on average sampled 33.97 and 23.03 times. And the maximum times even reached 275. From the third event to sample, the average times started to decrease and remain between 10 and 20 times. In the sampling of the last two events which were also rare events, the average times slightly increased again. It seemed that the participants tend to gradually lose their interest of sampling and gaining more experience; but the generally better knowledge of the probabilities of more common events could also affect, because in the end they seemed to regain the interest to sample the probabilities that they were not familiar with.

Sampling of	Observation	Mean	SD	Min	Max
€100 with the probability of 0.01	32	33.97	52.47	5	275
€100 with the probability of 0.05	32	23.03	17.96	5	70
€100 with the probability of 0.1	32	16.59	14.83	5	66
€100 with the probability of 0.5	32	11.53	6.93	5	37
€100 with the probability of 0.9	32	10.66	8.19	5	43
€100 with the probability of 0.95	32	12.63	11.11	5	53
€100 with the probability of 0.99	32	12.50	11.07	5	56

Table 4. Summary of sampling times.

To look into the effect of gaining experience from sampling, I first chose to use the Wilcoxon test which could make the comparison between within-participants data. The left of Figure 7 reflects the effect on parameter *a*. The null hypothesis was rejected (p=0.00) which suggested that participants' likelihood insensitivity level was lower (*a* became larger) after the sampling procedure. From the figures we could see that most participants (19 out of 22) had their parameter *a* rising after sampling; while only one participant had fallen parameter *a*, and two participants did not have changes. Whereas the right part shows that there was no decrease in pessimism (higher *b*; p=0.90). Half of the participants' parameter *b* decreased, which meant their pessimism increased; at the same time almost half of the participants' parameter *b* increased (pessimism decreased), and also two did not have changes.

```
Wilcoxon signed-rank test

        sign |      obs   sum ranks    expected
-------------+---------------------------------
    positive |        1         3.5         125
    negative |       19       246.5         125
        zero |        2           3           3
-------------+---------------------------------
         all |       22         253         253

unadjusted variance      948.75
adjustment for ties       -0.12
adjustment for zeros      -1.25
                     ----------
adjusted variance        947.38

Ho: a = a_s
             z =  -3.947
    Prob > |z| =   0.0001
```

```
Wilcoxon signed-rank test

        sign |      obs   sum ranks    expected
-------------+---------------------------------
    positive |       11         129         125
    negative |        9         121         125
        zero |        2           3           3
-------------+---------------------------------
         all |       22         253         253

unadjusted variance      948.75
adjustment for ties       -0.12
adjustment for zeros      -1.25
                     ----------
adjusted variance        947.38

Ho: b = b_s
             z =   0.130
    Prob > |z| =   0.8966
```

Figure 7. Results of Wilcoxon tests.

Then Student t-tests were performed to investigate whether there were absolute-changes. The observed absolute-changes between after and before sampling data were compared to 0. From the results (Figure 8) we could tell that both the changes of *a* and *b* were different from 0 (p=0.00 and p=0.00). We could interpret that although there were different directions of changing, the influences of direct experience on both the levels of likelihood insensitivity and pessimism were significant.

```
One-sample t test
------------------------------------------------------------------------------
Variable |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
a_abschg |      22    .4904545    .0853652    .4003983    .3129279    .6679812
------------------------------------------------------------------------------
    mean = mean(a_abschg)                                         t =   5.7454
Ho: mean = 0                                     degrees of freedom =       21

    Ha: mean < 0                 Ha: mean != 0                 Ha: mean > 0
 Pr(T < t) = 1.0000         Pr(|T| > |t|) = 0.0000          Pr(T > t) = 0.0000
```

```
One-sample t test
------------------------------------------------------------------------------
Variable |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
b_abschg |      22    .4645455    .0845941    .3967814    .2886224    .6404685
------------------------------------------------------------------------------
    mean = mean(b_abschg)                                         t =   5.4915
Ho: mean = 0                                     degrees of freedom =       21

    Ha: mean < 0                 Ha: mean != 0                 Ha: mean > 0
 Pr(T < t) = 1.0000         Pr(|T| > |t|) = 0.0000          Pr(T > t) = 0.0000
```

Figure 8. Results of Student t-tests.

Concluding this part, among the participants, general experience of probability including education and field of profession did not have an impact on either likelihood insensitivity or pessimism, whereas direct experience from sampling was related to both. Most participants tended to be more likelihood sensitive after sampling, while the participants were almost equally divided to become either more or less pessimistic.

4.3.5 Interaction of cognitive ability and experience

First, a linear regression was performed with the OLS procedure to explore how the total sample size (the sum of the sample sizes of every sampling procedure) of each participant was influenced. Table 5 shows the results of the regression. Keeping all other variables to stay the same, compared to scoring 0 point, scoring 1 point in the CRT increased the total sample size by 108.06 times for sampling (p=0.004), and scoring 2 points increased the size by 106.25 (p=0.04); scoring 3 points increased the size by 72.46, however, this was not significant (p=0.14). Nevertheless, different scores of CRT were tested to be jointly significant (p=0.02) with the influence on the total sample size. Besides the effect of cognitive ability, some demographic variables also had predictive power. Being female, compared to being male, decreased the size by 76.49 samples (p=0.005), ceteris paribus. Participants with Master's degree sampled 87.55 more times than those with Bachelor's degree (p=0.03), ceteris paribus.

```
Linear regression                               Number of obs =       32
                                                F(  8,     23) =     4.60
                                                Prob > F       =   0.0019
                                                R-squared      =   0.5762
                                                Root MSE       =   65.243
```

smp	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
age	-.8600743	4.483677	-0.19	0.850	-10.13527	8.415118
gender	-76.48878	24.67517	-3.10	0.005	-127.5333	-25.4443
edu	87.54961	37.45539	2.34	0.028	10.06722	165.032
field_prob	33.01724	30.80625	1.07	0.295	-30.71036	96.74483
dcrt2	108.0573	33.83026	3.19	0.004	38.07408	178.0405
dcrt3	106.2507	48.8099	2.18	0.040	5.27973	207.2217
dcrt4	72.45533	47.2334	1.53	0.139	-25.25441	170.1651
par_ex	-42.4231	38.11601	-1.11	0.277	-121.2721	36.42587
_cons	47.91926	106.9472	0.45	0.658	-173.3178	269.1564

```
 ( 1)  dcrt2 = 0
 ( 2)  dcrt3 = 0
 ( 3)  dcrt4 = 0

       F(  3,    23) =    3.86
            Prob > F =    0.0226
```

Table 5. Linear regression on total sample size and joint significance test.

When comparing the situations of original and post-sampling, there was something related to cognitive ability which was worth being noticed. Different from the effect of cognitive ability on likelihood insensitivity and pessimism levels before sampling, post-sampling likelihood insensitivity was not different (p=0.30) between the high and low cognitive ability groups, whereas the pessimism level became different (p=0.004) between the two groups. (Appendix F). This may suggest that after the learning process, both groups made use of the knowledge they just gained instead of their original understanding to evaluate the rare events; meanwhile, across all probabilities, the two groups learned from the experience differently and generated varied evaluation.

Further investigation on the rate of change of parameters *a* and *b* were conducted. The rate of change was calculated with the difference between the original and post-sampling values divided by the original value of *a* (*b*), and the symbols were kept to show the direction of the change. The larger the rate, the more the participant becomes less likelihood insensitive (pessimistic) compared to the original level. The changes of the level of likelihood insensitivity did not differ between the high and low cognitive ability groups (p=0.20). Also, neither the score of CRT (p=0.60, jointly) nor any other listed demographic variables (p>0.05) predicted the changes. (Appendix G). As for the changes of the level of pessimism, belonging to high or low cognitive ability group did not put difference to the rate of change (p=0.097) (Figure 9). In the OLS regression model (Table 6), scoring 1, 2, or 3 points in CRT decreased the rate compared to scoring 0, ceteris paribus, but the effect was not significant (p=0.095, p=0.12, and p=0.40). The three of them were not jointly significant (p=0.35) either. Worth mentioning, gender was correlated with the rate of change of pessimism (p=0.01). Controlling other variables, being female rather than male increased the rate by 1.15.

```
Two-sample Wilcoxon rank-sum (Mann-Whitney) test

       crt_l |      obs    rank sum    expected
-------------+---------------------------------
           0 |        6        91.5          69
           1 |       16       161.5         184
-------------+---------------------------------
    combined |       22         253         253

unadjusted variance       184.00
adjustment for ties        -0.21
                      ----------
adjusted variance         183.79

Ho: b_chg(crt_l==0) = b_chg(crt_l==1)
             z =   1.660
    Prob > |z| =   0.0970
```

Figure 9. Results of Mann-Whitney U test.

Linear regression					Number of obs	= 22
					F(8, 13)	= 1.77
					Prob > F	= 0.1722
					R-squared	= 0.5711
					Root MSE	= .51196

b_chg	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
age	.2033144	.1320901	1.54	0.148	-.082049	.4886778
gender	1.152732	.3924065	2.94	0.012	.3049895	2.000475
edu	-.9539948	.4885958	-1.95	0.073	-2.009542	.1015522
field_prob	-.0556827	.2710437	-0.21	0.840	-.6412371	.5298716
dcrt2	-.5583767	.3102942	-1.80	0.095	-1.228727	.1119732
dcrt3	-.5076866	.3034823	-1.67	0.118	-1.16332	.1479469
dcrt4	-.3252217	.3770904	-0.86	0.404	-1.139876	.4894325
smp	.0032522	.0016951	1.92	0.077	-.0004098	.0069141
_cons	-4.723161	2.876649	-1.64	0.125	-10.93778	1.49146

```
 ( 1)  dcrt2 = 0
 ( 2)  dcrt3 = 0
 ( 3)  dcrt4 = 0

       F(  3,    13) =    1.19
            Prob > F =    0.3509
```

Table 6. Linear regression on rate of change of b and joint significance test.

To sum up this part, there were interactions between both cognitive ability and experience themselves and their influences on participants' probability weighting. First, cognitive ability had influence on participants' sampling sizes, which allowed them to access more or less direct experience. Second, the significance of the effects of cognitive ability on likelihood insensitivity and pessimism was different between the situations before and after sampling. Third, lower cognitive ability was weakly correlated with larger change of pessimism towards the direction of being less pessimistic after gaining the experience, while there was no such difference between the comparison of the two groups when it came to likelihood insensitivity.

5. Discussion and Conclusion

From the results we could firstly see that most participants' choices under risk were able to be covered by the two-parameter probability weighting function, with inverse-S shaped curves. Participants did show poorer understanding of probabilities. The average level of pessimism was close to neutral, which put difference with the condition of commonly found data (b=0.77). The experimental setting of hypothetical choices, by which people would be more willing to take risks, could explain this deviation (Holt & Laury, 2002). Demography had some influence in some situations but not much compared to other studies (e.g. Halek & Eisenhauer, 2001; Holt & Laury, 2002; Frederick, 2005; Choi et al., 2014; Frey et al. 2015), and the reason may be the homogeneity of the characteristics of the participants and the relatively small sample group. The participants were highly educated, relatively young with a small range of age, where the variety was far from representing a wider population.

The three hypotheses were supported to some extent by the results. Higher cognitive ability was correlated with lower likelihood insensitivity (better discrimination), which represented relatively better understanding of (extreme) probabilities. Pessimism (elevation) was not significantly correlated with cognitive-ability level. This effect echoed the results of the study by Petrova et al. (2014), where higher numeracy was correlated with discrimination via higher variance of emotion, while they also did not find a clear relation with elevation. Providing that participants' previous knowledge and experience of probability did not influence the value of the parameters, the accessibility of the information in their minds should not differ much (Kahneman, 2003). With the assumption of two-system theory, it could be explained by the better usage of System 2 to perform deliberative thinking among people with high cognitive ability. Since the high score of CRT could reflect the tendency of using System 2 to correct intuitive reaction generated by System 1 and the possibility of being more patient (Frederick, 2005), it would be plausible to say that the high cognitive ability group would apply more effortful thinking compared to the low cognitive ability group even when they both use heuristics to interpret (Cokely et al. 2012). And therefore, they would deliver less deviated subjective evaluation and less overweight rare probabilities. At the same time, the existence of likelihood insensitivity proved that the errors from spontaneity may not be fully corrected.

On average level, direct experience from sampling mainly had influence on likelihood insensitivity while not much on pessimism. Most participants had an increased *a*; while their *b*s changed to different directions. To put it more straight-forward, direct experience could negatively influence likelihood insensitivity; while the influence on pessimism was also significant but could be either positive or negative, from which we could see that individuals learned from direct experience differently to moderate their pessimism. Being more likelihood sensitive after gaining direct experience could suggest that the sampling procedure helped the participants to understand probabilities better. They developed the knowledge or sense of these probabilities that were not normal in their daily life. More information was stored in their minds for them to approach even if they used heuristics. In this respect the theories from previous studies (Kunreuther et al., 2001; Gayer, 2010; Hertwig et al., 2004; Van de Kuilen, 2009) were strongly supported.

However, there was a tendency to overly correct on the average level of overweighting rare events and on average the participants slightly underweighted the extreme probabilities. This confirmed the findings of Hertwig et al. (2004) even further because in my experiment the decision from experience was combined with decision from description. They suggested that the underweighting after sampling could be caused by limited sample size and recency effect. Among my participants the problem with limited sample size did exist. There was a tendency of decreasing the times to sample along with the whole sampling procedure. Although they relatively sampled more times on rare events, the sizes for them to get a better feeling of the numbers were still quite small compared to the given distribution.

Larger standard deviation in both likelihood insensitivity and pessimism was also found after the sampling procedure, which may suggest the individual difference of being affected by experience. On the individual level, general experience did not show a significant effect. It could have resulted from the small variation of the education level of the participants. Another plausible reason could be that the field of profession is not the only source for people to develop experience related to probabilities. As Gayer (2010) suggested, there are multiple sources for memories, including their own life experience as well as the experience from people around them. So that in comparison, their profession may only play a small part.

The interactions between cognitive ability and experience was also confirmed. Cognitive ability was correlated with sampling size. Although this correlation was not necessarily positive, scoring 0 point in CRT had a negative effect on the total sample size compared to other scores. Being less patient may be the reason. Among the participants scoring 1 to 3 points, there seemed to be a slightly negative effect of the score on the sample size. The reason was not clear. I would assume that those with higher scores may be more confident with their ability of induction so that they felt less necessity to sample a lot. Further investigation could be conducted on this issue.

After gaining direct experience, likelihood insensitivity between different cognitive ability groups became indifferent, and the change rate of parameter *a* was not influenced by the level of cognitive ability. What could happen during the learning process was that individuals were influenced by the experience to different extent despite the level of cognitive ability, although most developed better understanding of rare probabilities and became more likelihood sensitive. In other words, among these participants, cognitive ability did not influence how much they learned from sampling. This was a highly educated group so that in general they possessed high learning ability, and the variation on cognitive ability among them did not play an important role in learning anymore.

Pessimism between the two groups became different from each other. So as the rate of change, that low cognitive ability group was slightly more influenced in changing to less pessimistic. These meant that participants with low cognitive ability systematically changed more in the direction of less pessimism after sampling compared to those with higher cognitive ability who tended to change less towards this direction or to change to the direction of more pessimism. It seemed to be opposite the finding that higher cognitive ability is related to less risk aversion (Dohmen et al., 2008; Burks et al., 2009; Benjamin et al., 2013), which could be caused by the effect of direct experience in general or the detailed content of the experience. More explanation may be delivered after further study.

To sum up, among highly educated young people, subjective probability weighting towards choices under risk was confirmed. Cognitive ability was negatively related to likelihood insensitivity; direct experience was negatively correlated with likelihood insensitivity and correlated with pessimism in a more complex way; the interaction between the effect of cognitive ability and experience on probability weighting was confirmed, and it performed differently on likelihood insensitivity and pessimism. The two-system theory seemed plausible to explain the main findings. To my knowledge this is the first study trying to disentangle the effect of cognitive ability and experience, and to find a clearer interaction between them. Some down sides would be that the characteristics of the participants in my sample group were relatively homogeneity, and the sample size was rather small. And therefore, further study is called for to investigate whether these conclusions would shed light on a wider population. Due to some technical problems there were no real incentives involved, and the environment for the experiment was not strictly controlled, which might lead to some deviation compared to well controlled lab experiment.

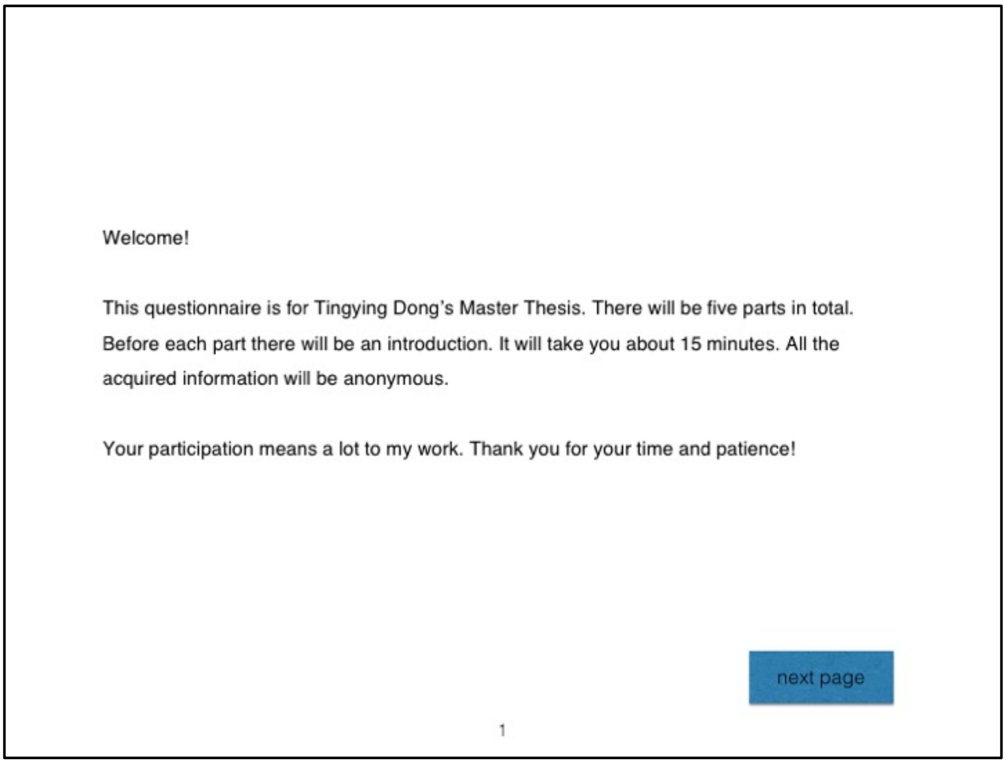
The findings could also be applied to the real world, especially to the design of policies and regulations. Andersson et al. (2013) suggested that the existence of the relation between cognitive ability and risk attitude could have crucial influence on how policies would be implicated. Choi et al. (2014) discussed that decision-making ability could be manipulated by policies to achieve goals of economic outcomes, because different decision-making ability would lead to varied understanding of economic outcomes. In their definition of decision-making ability, it could be influenced by cognitive ability and education for example. The type of policies mentioned in the studies were libertarian paternalism and asymmetric paternalism.

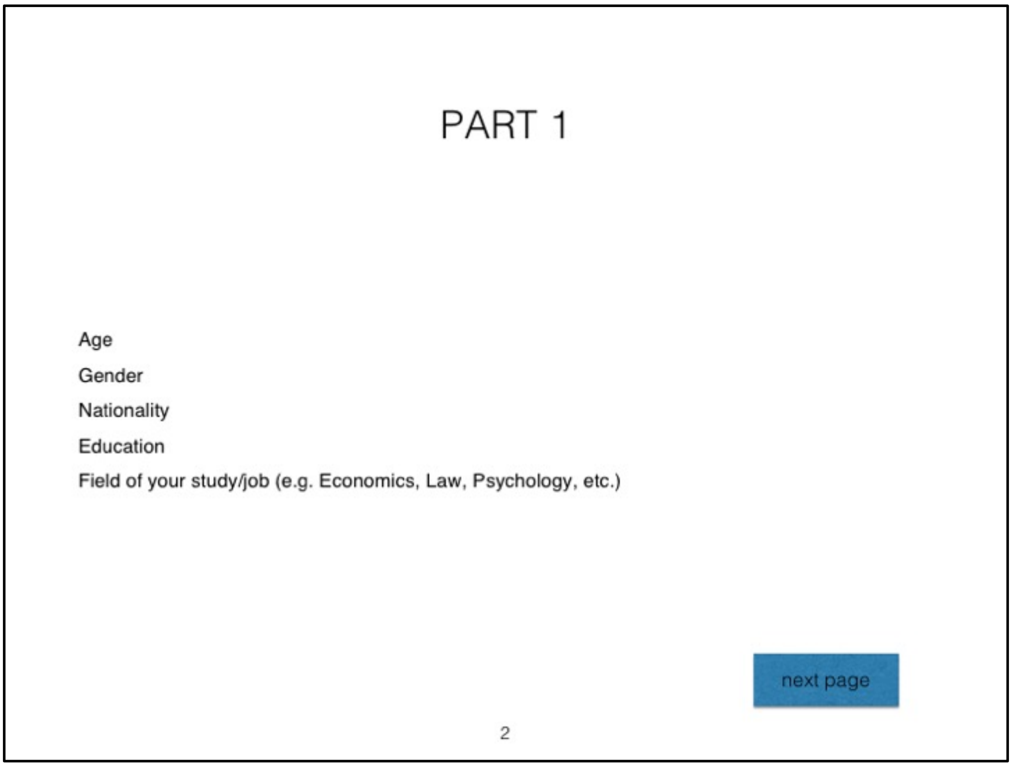
In Thaler and Sunstein (2003)'s opinion, paternalism is the idea of designing a policy to positively influence the targeted groups. Libertarian paternalism is the paternalistic action that is acceptable for libertarians, because it avoids. It is applicable to both private and public organisations in a wide range. They believed that it was a false assumption that people’s choices always follow their best interest, and therefore making use of libertarian paternalism would help people to be better off. According to Camerer et al. (2003), asymmetric paternalistic policies would largely benefit people who fail to make choices in their best interest, while do not generate harm to people who are rational. They shared the opinion that a main reason to use paternalism would that (some) people are not perfectly rational, which was confirmed by my findings that cognitive and experience limitations led to worse understanding of (extreme) probabilities and subjective probability weighting, which in the end would cause deviation of choices from their best interest. Although again there is the limitation of sample group in my study, it could provide support to the policies toward the groups with similar characteristics.

Appendix

A. The online questionnaire

Here is the questionnaire that was designed and programmed to be available online for the participants to fill in.





PART 2

Introduction: In this part, you will answer three questions. Please only fill in numbers into the blank spaces.

next page

3

PART 2

Question 1:

A bat and a ball cost €1.10. The bat costs €1.00 more than the ball.

How much does the ball cost?

___ cents

next page

4

PART 2

Question 2:

If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets?

___ minutes

next page

5

PART 2

Question3:

In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake?

___ days

next page

6

PART 3

Introduction: In this part, you will see seven tables. Each table contains eleven rows with two options each. Option A entails that you can obtain a particular amount of money for sure. Option B entails that you can obtain €100 with some probability and €0 otherwise. From Row 1 to Row 11, Option A gets better and better.

Look at the table on the right as an example. In each row, you have to think of which option you'd like to choose from the two there. In Row 1, you won't obtain any money under Option A, yet there's a chance to obtain €100 under Option B. So everyone will prefer Option B here. In Row 11, you would obtain €100 for sure under Option A, yet there's only a chance to obtain this amount under Option B. So everyone will prefer Option A here. From Row 2 to Row 10, people's preferences may vary. Let's say in Row 2 you still prefer Option B, and in Row 3 you prefer Option A. Since Option A gets better and better, you would keep choosing Option A from Row 3 to 11. Then it means that your preference switches from B to A between Rows 2 and 3 in this table.

	OPTION A	OPTION B
1	€0 for sure	€100 with the probability of 0.2
2	€10 for sure	€100 with the probability of 0.2
3	€20 for sure	€100 with the probability of 0.2
4	€30 for sure	€100 with the probability of 0.2
5	€40 for sure	€100 with the probability of 0.2
6	€50 for sure	€100 with the probability of 0.2
7	€60 for sure	€100 with the probability of 0.2
8	€70 for sure	€100 with the probability of 0.2
9	€80 for sure	€100 with the probability of 0.2
10	€90 for sure	€100 with the probability of 0.2
11	€100 for sure	€100 with the probability of 0.2

Notice: There's no right or wrong answers. Please just make your decisions depending on your own preference and choose the options that suit you the best. Your preference is what interests me in this investigation.

Before the real questions, you will see an example to practice on the next page.

next page

7

PART 3

Example

My preference switches from Option B to Option A between Rows _____.

	OPTION A	OPTION B
1	€0 for sure	€100 with the probability of 0.2
2	€10 for sure	€100 with the probability of 0.2
3	€20 for sure	€100 with the probability of 0.2
4	€30 for sure	€100 with the probability of 0.2
5	€40 for sure	€100 with the probability of 0.2
6	€50 for sure	€100 with the probability of 0.2
7	€60 for sure	€100 with the probability of 0.2
8	€70 for sure	€100 with the probability of 0.2
9	€80 for sure	€100 with the probability of 0.2
10	€90 for sure	€100 with the probability of 0.2
11	€100 for sure	€100 with the probability of 0.2

This is the end of the example. Now come the real questions.

next page

8

PART 3

Table 1

My preference switches from Option B to Option A between Rows _____.

	OPTION A	OPTION B
1	€0 for sure	€100 with the probability of 0.01
2	€0.5 for sure	€100 with the probability of 0.01
3	€1 for sure	€100 with the probability of 0.01
4	€1.5 for sure	€100 with the probability of 0.01
5	€2 for sure	€100 with the probability of 0.01
6	€5 for sure	€100 with the probability of 0.01
7	€15 for sure	€100 with the probability of 0.01
8	€25 for sure	€100 with the probability of 0.01
9	€50 for sure	€100 with the probability of 0.01
10	€75 for sure	€100 with the probability of 0.01
11	€100 for sure	€100 with the probability of 0.01

next page

9

PART 3

Table 2

My preference switches from Option B to Option A between Rows _____.

	OPTION A	OPTION B
1	€0 for sure	€100 with the probability of 0.05
2	€2.5 for sure	€100 with the probability of 0.05
3	€5 for sure	€100 with the probability of 0.05
4	€7.5 for sure	€100 with the probability of 0.05
5	€10 for sure	€100 with the probability of 0.05
6	€15 for sure	€100 with the probability of 0.05
7	€20 for sure	€100 with the probability of 0.05
8	€40 for sure	€100 with the probability of 0.05
9	€60 for sure	€100 with the probability of 0.05
10	€80 for sure	€100 with the probability of 0.05
11	€100 for sure	€100 with the probability of 0.05

next page

10

PART 3

Table 3

My preference switches from Option B to Option A between Rows _____.

	OPTION A	OPTION B
1	€0 for sure	€100 with the probability of 0.1
2	€5 for sure	€100 with the probability of 0.1
3	€10 for sure	€100 with the probability of 0.1
4	€15 for sure	€100 with the probability of 0.1
5	€20 for sure	€100 with the probability of 0.1
6	€25 for sure	€100 with the probability of 0.1
7	€30 for sure	€100 with the probability of 0.1
8	€40 for sure	€100 with the probability of 0.1
9	€60 for sure	€100 with the probability of 0.1
10	€80 for sure	€100 with the probability of 0.1
11	€100 for sure	€100 with the probability of 0.1

next page

11

PART 3

Table 4

My preference switches from Option B to Option A between Rows _____.

	OPTION A	OPTION B
1	€0 for sure	€100 with the probability of 0.5
2	€10 for sure	€100 with the probability of 0.5
3	€20 for sure	€100 with the probability of 0.5
4	€30 for sure	€100 with the probability of 0.5
5	€40 for sure	€100 with the probability of 0.5
6	€50 for sure	€100 with the probability of 0.5
7	€60 for sure	€100 with the probability of 0.5
8	€70 for sure	€100 with the probability of 0.5
9	€80 for sure	€100 with the probability of 0.5
10	€90 for sure	€100 with the probability of 0.5
11	€100 for sure	€100 with the probability of 0.5

next page

12

PART 3

Table 5

My preference switches from Option B to Option A between Rows _____.

	OPTION A	OPTION B
1	€0 for sure	€100 with the probability of 0.9
2	€20 for sure	€100 with the probability of 0.9
3	€40 for sure	€100 with the probability of 0.9
4	€60 for sure	€100 with the probability of 0.9
5	€70 for sure	€100 with the probability of 0.9
6	€75 for sure	€100 with the probability of 0.9
7	€80 for sure	€100 with the probability of 0.9
8	€85 for sure	€100 with the probability of 0.9
9	€90 for sure	€100 with the probability of 0.9
10	€95 for sure	€100 with the probability of 0.9
11	€100 for sure	€100 with the probability of 0.9

next page

13

PART 3

Table 6

My preference switches from Option B to Option A between Rows _____.

	OPTION A	OPTION B
1	€0 for sure	€100 with the probability of 0.95
2	€20 for sure	€100 with the probability of 0.95
3	€40 for sure	€100 with the probability of 0.95
4	€60 for sure	€100 with the probability of 0.95
5	€80 for sure	€100 with the probability of 0.95
6	€85 for sure	€100 with the probability of 0.95
7	€90 for sure	€100 with the probability of 0.95
8	€92.5 for sure	€100 with the probability of 0.95
9	€95 for sure	€100 with the probability of 0.95
10	€97.5 for sure	€100 with the probability of 0.95
11	€100 for sure	€100 with the probability of 0.95

next page

14

PART 3

Table 7

My preference switches from Option B to Option A between Rows _____.

	OPTION A	OPTION B
1	€0 for sure	€100 with the probability of 0.99
2	€25 for sure	€100 with the probability of 0.99
3	€50 for sure	€100 with the probability of 0.99
4	€75 for sure	€100 with the probability of 0.99
5	€85 for sure	€100 with the probability of 0.99
6	€95 for sure	€100 with the probability of 0.99
7	€98 for sure	€100 with the probability of 0.99
8	€98.5 for sure	€100 with the probability of 0.99
9	€99 for sure	€100 with the probability of 0.99
10	€99.5 for sure	€100 with the probability of 0.99
11	€100 for sure	€100 with the probability of 0.99

next page

15

PART 4

Introduction: In this part, you will first click the button and sample the given probabilities for at least 5 times (after that as many times more as you like) by yourself. After sampling, you will have to choose the switching point in the tables as in Part 3.

Here is an example:

Sample: €100 with the probability of 0.2

Result x:

You will see the table on next page

next page

16

PART 4

Example

My preference switches from Option B to Option A between Rows _____.

	OPTION A	OPTION B
1	€0 for sure	€100 with the probability of 0.2
2	€10 for sure	€100 with the probability of 0.2
3	€20 for sure	€100 with the probability of 0.2
4	€30 for sure	€100 with the probability of 0.2
5	€40 for sure	€100 with the probability of 0.2
6	€50 for sure	€100 with the probability of 0.2
7	€60 for sure	€100 with the probability of 0.2
8	€70 for sure	€100 with the probability of 0.2
9	€80 for sure	€100 with the probability of 0.2
10	€90 for sure	€100 with the probability of 0.2
11	€100 for sure	€100 with the probability of 0.2

This is the end of the example. Now come the real questions.

next page

17

PART 4

Sample: €100 with the probability of **0.01**

Result x:

next page

18

PART 4

Table 1

My preference switches from Option B to Option A between Rows _____.

	OPTION A	OPTION B
1	€0 for sure	€100 with the probability of 0.01
2	€0.5 for sure	€100 with the probability of 0.01
3	€1 for sure	€100 with the probability of 0.01
4	€1.5 for sure	€100 with the probability of 0.01
5	€2 for sure	€100 with the probability of 0.01
6	€5 for sure	€100 with the probability of 0.01
7	€15 for sure	€100 with the probability of 0.01
8	€25 for sure	€100 with the probability of 0.01
9	€50 for sure	€100 with the probability of 0.01
10	€75 for sure	€100 with the probability of 0.01
11	€100 for sure	€100 with the probability of 0.01

next page

19

PART 4

Sample: €100 with the probability of **0.05**

Result x:

next page

20

PART 4

Table 2

My preference switches from Option B to Option A between Rows _____.

	OPTION A	OPTION B
1	€0 for sure	€100 with the probability of 0.05
2	€2.5 for sure	€100 with the probability of 0.05
3	€5 for sure	€100 with the probability of 0.05
4	€7.5 for sure	€100 with the probability of 0.05
5	€10 for sure	€100 with the probability of 0.05
6	€15 for sure	€100 with the probability of 0.05
7	€20 for sure	€100 with the probability of 0.05
8	€40 for sure	€100 with the probability of 0.05
9	€60 for sure	€100 with the probability of 0.05
10	€80 for sure	€100 with the probability of 0.05
11	€100 for sure	€100 with the probability of 0.05

next page

21

PART 4

Sample: €100 with the probability of **0.1**

Result x:

next page

22

PART 4

Table 3

My preference switches from Option B to Option A between Rows _____.

	OPTION A	OPTION B
1	€0 for sure	€100 with the probability of 0.1
2	€5 for sure	€100 with the probability of 0.1
3	€10 for sure	€100 with the probability of 0.1
4	€15 for sure	€100 with the probability of 0.1
5	€20 for sure	€100 with the probability of 0.1
6	€25 for sure	€100 with the probability of 0.1
7	€30 for sure	€100 with the probability of 0.1
8	€40 for sure	€100 with the probability of 0.1
9	€60 for sure	€100 with the probability of 0.1
10	€80 for sure	€100 with the probability of 0.1
11	€100 for sure	€100 with the probability of 0.1

next page

23

PART 4

Sample: €100 with the probability of **0.5**

Result x:

next page

24

PART 4

Table 4

My preference switches from Option B to Option A between Rows _____.

	OPTION A	OPTION B
1	€0 for sure	€100 with the probability of 0.5
2	€10 for sure	€100 with the probability of 0.5
3	€20 for sure	€100 with the probability of 0.5
4	€30 for sure	€100 with the probability of 0.5
5	€40 for sure	€100 with the probability of 0.5
6	€50 for sure	€100 with the probability of 0.5
7	€60 for sure	€100 with the probability of 0.5
8	€70 for sure	€100 with the probability of 0.5
9	€80 for sure	€100 with the probability of 0.5
10	€90 for sure	€100 with the probability of 0.5
11	€100 for sure	€100 with the probability of 0.5

next page

25

PART 4

Sample: €100 with the probability of **0.9**

Result x:

next page

26

PART 4

Table 5

My preference switches from Option B to Option A between Rows _____.

	OPTION A	OPTION B
1	€0 for sure	€100 with the probability of 0.9
2	€20 for sure	€100 with the probability of 0.9
3	€40 for sure	€100 with the probability of 0.9
4	€60 for sure	€100 with the probability of 0.9
5	€70 for sure	€100 with the probability of 0.9
6	€75 for sure	€100 with the probability of 0.9
7	€80 for sure	€100 with the probability of 0.9
8	€85 for sure	€100 with the probability of 0.9
9	€90 for sure	€100 with the probability of 0.9
10	€95 for sure	€100 with the probability of 0.9
11	€100 for sure	€100 with the probability of 0.9

next page

27

PART 4

Sample: €100 with the probability of **0.95**

Result x:

next page

28

PART 4

Table 6

My preference switches from Option B to Option A between Rows _____.

	OPTION A	OPTION B
1	€0 for sure	€100 with the probability of 0.95
2	€20 for sure	€100 with the probability of 0.95
3	€40 for sure	€100 with the probability of 0.95
4	€60 for sure	€100 with the probability of 0.95
5	€80 for sure	€100 with the probability of 0.95
6	€85 for sure	€100 with the probability of 0.95
7	€90 for sure	€100 with the probability of 0.95
8	€92.5 for sure	€100 with the probability of 0.95
9	€95 for sure	€100 with the probability of 0.95
10	€97.5 for sure	€100 with the probability of 0.95
11	€100 for sure	€100 with the probability of 0.95

next page

29

PART 4

Sample: €100 with the probability of **0.99**

Result x:

next page

30

PART 4

Table 7

My preference switches from Option B to Option A between Rows _____.

	OPTION A	OPTION B
1	€0 for sure	€100 with the probability of 0.99
2	€25 for sure	€100 with the probability of 0.99
3	€50 for sure	€100 with the probability of 0.99
4	€75 for sure	€100 with the probability of 0.99
5	€85 for sure	€100 with the probability of 0.99
6	€95 for sure	€100 with the probability of 0.99
7	€98 for sure	€100 with the probability of 0.99
8	€98.5 for sure	€100 with the probability of 0.99
9	€99 for sure	€100 with the probability of 0.99
10	€99.5 for sure	€100 with the probability of 0.99
11	€100 for sure	€100 with the probability of 0.99

next page

31

PART 5

Did you ever participate in Economics Experiments before?

Did you ever see (parts of) the three maths questions (bat and ball, machines making widgets, and patch of lily pads) before?

32

This is the end of the questionnaire.

Thank you every much for your help!

33

B. Mann-Whitney U test of the influence of participants' knowledge of CRT

```
Two-sample Wilcoxon rank-sum (Mann-Whitney) test

       crt_ex |      obs    rank sum    expected
--------------+---------------------------------
            0 |       19       299.5       313.5
            1 |       13       228.5       214.5
--------------+---------------------------------
     combined |       32         528         528

unadjusted variance      679.25
adjustment for ties      -72.08
                     ----------
adjusted variance        607.17

Ho: crt(crt_ex==0) = crt(crt_ex==1)
             z =  -0.568
    Prob > |z| =   0.5699
```

C. Linear regressions on *a* and *b*

```
Linear regression                                      Number of obs =      24
                                                       F(  8,    15) =    3.54
                                                       Prob > F      =  0.0168
                                                       R-squared     =  0.4936
                                                       Root MSE      =  .27326

------------------------------------------------------------------------------
             |               Robust
           a |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |  -.0446758   .0579443    -0.77   0.453    -.1681812    .0788296
      gender |   .1716381   .1656318     1.04   0.316    -.1813978     .524674
         edu |   .1641198   .1806769     0.91   0.378     -.220984    .5492236
  field_prob |   .1169555   .1401448     0.83   0.417    -.1817559     .415667
       dcrt2 |   .1131095   .1940353     0.58   0.569    -.3004669    .5266859
       dcrt3 |   .6509186   .2121284     3.07   0.008     .1987776     1.10306
       dcrt4 |    .484161   .1794851     2.70   0.017     .1015977    .8667244
      par_ex |  -.0855243   .1720317    -0.50   0.626    -.4522012    .2811525
       _cons |   .8421907   1.284198     0.66   0.522    -1.895012    3.579393
------------------------------------------------------------------------------

 ( 1)  dcrt2 = 0
 ( 2)  dcrt3 = 0
 ( 3)  dcrt4 = 0

       F(  3,    15) =    5.94
            Prob > F =    0.0070
```

```
Linear regression                                      Number of obs =      24
                                                       F(  8,    15) =    1.37
                                                       Prob > F      =  0.2868
                                                       R-squared     =  0.2482
                                                       Root MSE      =  .34271

------------------------------------------------------------------------------
             |               Robust
           b |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |  -.0477846   .0592446    -0.81   0.433    -.1740615    .0784922
      gender |  -.0033731    .184186    -0.02   0.986    -.3959563      .38921
         edu |   .3293012   .2447635     1.35   0.198    -.1923998    .8510023
  field_prob |   .0859669   .1496935     0.57   0.574    -.2330972    .4050311
       dcrt2 |   .4499886   .2282465     1.97   0.067    -.0365073    .9364845
       dcrt3 |   .3060793   .2219001     1.38   0.188    -.1668895     .779048
       dcrt4 |   .2883618   .1730143     1.67   0.116    -.0804094     .657133
      par_ex |  -.3725729   .2319462    -1.61   0.129    -.8669544    .1218087
       _cons |   1.697531   1.316373     1.29   0.217    -1.108252    4.503313
------------------------------------------------------------------------------

 ( 1)  dcrt2 = 0
 ( 2)  dcrt3 = 0
 ( 3)  dcrt4 = 0

       F(  3,    15) =    1.73
            Prob > F =    0.2041
```

D. Mann-Whitney U tests of the effect of cognitive ability level on *a* and *b*

```
Two-sample Wilcoxon rank-sum (Mann-Whitney) test

       crt_l |      obs    rank sum    expected
-------------+---------------------------------
           0 |        6        30.5          75
           1 |       18       269.5         225
-------------+---------------------------------
    combined |       24         300         300

unadjusted variance      225.00
adjustment for ties       -0.20
                     ----------
adjusted variance        224.80

Ho: a(crt_l==0) = a(crt_l==1)
             z =  -2.968
    Prob > |z| =   0.0030
```

```
Two-sample Wilcoxon rank-sum (Mann-Whitney) test

       crt_l |      obs    rank sum    expected
-------------+---------------------------------
           0 |        6        82.5          75
           1 |       18       217.5         225
-------------+---------------------------------
    combined |       24         300         300

unadjusted variance      225.00
adjustment for ties       -0.29
                     ----------
adjusted variance        224.71

Ho: b(crt_l==0) = b(crt_l==1)
             z =   0.500
    Prob > |z| =   0.6168
```

E. Tests of the influence of education level and field of study/work as general experience

```
 ( 1)  edu = 0
 ( 2)  field_prob = 0

       F(  2,    15) =    0.55
            Prob > F =    0.5876
```

(Test of joint significance on *a*)

```
 ( 1)  edu = 0
 ( 2)  field_prob = 0

       F(  2,    15) =    0.92
            Prob > F =    0.4210
```

(Test of joint significance on *b*)

```
Two-sample Wilcoxon rank-sum (Mann-Whitney) test

         edu |      obs    rank sum    expected
-------------+---------------------------------
           0 |        9       110.5       112.5
           1 |       15       189.5       187.5
-------------+---------------------------------
    combined |       24         300         300

unadjusted variance      281.25
adjustment for ties       -0.24
                     ----------
adjusted variance        281.01

Ho: a(edu==0) = a(edu==1)
             z =  -0.119
    Prob > |z| =   0.9050
```

```
Two-sample Wilcoxon rank-sum (Mann-Whitney) test

  field_prob |      obs    rank sum    expected
-------------+---------------------------------
           0 |        5        65.5        62.5
           1 |       19       234.5       237.5
-------------+---------------------------------
    combined |       24         300         300

unadjusted variance      197.92
adjustment for ties       -0.17
                     ----------
adjusted variance        197.74

Ho: a(field_~b==0) = a(field_~b==1)
             z =   0.213
    Prob > |z| =   0.8311
```

```
Two-sample Wilcoxon rank-sum (Mann-Whitney) test

         edu |      obs    rank sum    expected
-------------+---------------------------------
           0 |        9       117.5       112.5
           1 |       15       182.5       187.5
-------------+---------------------------------
    combined |       24         300         300

unadjusted variance      281.25
adjustment for ties       -0.37
                     ----------
adjusted variance        280.88

Ho: b(edu==0) = b(edu==1)
             z =   0.298
    Prob > |z| =   0.7654
```

```
Two-sample Wilcoxon rank-sum (Mann-Whitney) test

  field_prob |      obs    rank sum    expected
-------------+---------------------------------
           0 |        5          65        62.5
           1 |       19         235       237.5
-------------+---------------------------------
    combined |       24         300         300

unadjusted variance      197.92
adjustment for ties       -0.26
                     ----------
adjusted variance        197.66

Ho: b(field_~b==0) = b(field_~b==1)
             z =   0.178
    Prob > |z| =   0.8589
```

F. Mann-Whitney U tests of the effect of cognitive ability level on a_s and b_s

```
Two-sample Wilcoxon rank-sum (Mann-Whitney) test

       crt_l |      obs    rank sum    expected
-------------+---------------------------------
           0 |        8        95.5         116
           1 |       20       310.5         290
-------------+---------------------------------
    combined |       28         406         406

unadjusted variance      386.67
adjustment for ties       -0.21
                     ----------
adjusted variance        386.46

Ho: a_s(crt_l==0) = a_s(crt_l==1)
             z =  -1.043
    Prob > |z| =   0.2970
```

```
Two-sample Wilcoxon rank-sum (Mann-Whitney) test

       crt_l |      obs    rank sum    expected
-------------+---------------------------------
           0 |        8         172         116
           1 |       20         234         290
-------------+---------------------------------
    combined |       28         406         406

unadjusted variance      386.67
adjustment for ties       -0.42
                     ----------
adjusted variance        386.24

Ho: b_s(crt_l==0) = b_s(crt_l==1)
             z =   2.849
    Prob > |z| =   0.0044
```

G. Mann-Whitney U test, linear regression, and joint significance test on rate of change of *a*

```
Two-sample Wilcoxon rank-sum (Mann-Whitney) test

       crt_l |      obs    rank sum    expected
-------------+---------------------------------
           0 |        6        86.5          69
           1 |       16       166.5         184
-------------+---------------------------------
    combined |       22         253         253

unadjusted variance      184.00
adjustment for ties       -0.10
                     ----------
adjusted variance        183.90

Ho: a_chg(crt_l==0) = a_chg(crt_l==1)
             z =   1.290
    Prob > |z| =   0.1969
```

```
Linear regression                               Number of obs =       22
                                                F(  8,    13) =     1.93
                                                Prob > F      =   0.1402
                                                R-squared     =   0.4950
                                                Root MSE      =   .72717
```

a_chg	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
age	-.1138923	.1828701	-0.62	0.544	-.5089592	.2811746
gender	.49495	.4718768	1.05	0.313	-.5244779	1.514378
edu	-.2413331	.6850819	-0.35	0.730	-1.721362	1.238696
field_prob	.5912114	.3428221	1.72	0.108	-.1494107	1.331834
dcrt2	.7461031	.8056193	0.93	0.371	-.9943316	2.486538
dcrt3	-.1986374	.5346253	-0.37	0.716	-1.353625	.9563503
dcrt4	.0238352	.5562245	0.04	0.966	-1.177815	1.225485
smp	.001008	.0020636	0.49	0.633	-.0034502	.0054661
_cons	2.636002	4.141619	0.64	0.536	-6.311422	11.58343

```
 ( 1)  dcrt2 = 0
 ( 2)  dcrt3 = 0
 ( 3)  dcrt4 = 0

       F(  3,    13) =    0.66
            Prob > F =    0.5907
```

References

Afriat, S. N. (1972). Efficiency estimation of production functions.*International Economic Review*, 568-598.

Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école américaine. *Econometrica: Journal of the Econometric Society*, 503-546.

Andersson, O., Tyran, J. R., Wengström, E., & Holm, H. J. (2013). Risk aversion relates to cognitive ability: Fact or fiction?.

Barron, G., & Erev, I. (2003). Small feedback- based decisions and their limited correspondence to description- based decisions. *Journal of Behavioral Decision Making*, *16*(3), 215-233.

Benjamin, D. J., Brown, S. A., & Shapiro, J. M. (2013). Who is ‘behavioral’? Cognitive ability and anomalous preferences. *Journal of the European Economic Association*, 11(6), 1231-1255.

Bernoulli, D. (1954). Exposition of a new theory on the measurement of risk. *Econometrica: Journal of the Econometric Society*, 23-36.

Brandstätter, E., Kühberger, A., & Schneider, F. (2002). A cognitive- emotional account of the shape of the probability weighting function. *Journal of Behavioral Decision Making*, *15*(2), 79-100.

Burks, S. V., Carpenter, J. P., Goette, L., & Rustichini, A. (2009). Cognitive skills affect economic preferences, strategic behavior, and job attachment. *Proceedings of the National Academy of Sciences*, 106(19), 7745-7750.

Camerer, C., Issacharoff, S., Loewenstein, G., O'donoghue, T., & Rabin, M. (2003). Regulation for Conservatives: Behavioral Economics and the Case for" Asymmetric Paternalism". *University of Pennsylvania law review*, *151*(3), 1211-1254.

Cohen, M., Jaffray, J. Y., & Said, T. (1987). Experimental comparison of individual behavior under risk and under uncertainty for gains and for losses. *Organizational behavior and human decision processes*, *39*(1), 1-22.

Cokely, E. T., Galesic, M., Schulz, E., Ghazal, S., & Garcia-Retamero, R. (2012). Measuring risk literacy: The Berlin numeracy test. *Judgment and Decision Making*, *7*(1), 25.

Cokely, E. T., & Kelley, C. M. (2009). Cognitive abilities and superior decision making under risk: A protocol analysis and process model evaluation. *Judgment and Decision Making*, *4*(1), 20.

Cokely, E. T., Parpart, P., & Schooler, L. J. (2009). On the link between cognitive control and heuristic processes. In *Proceedings of the 31st annual conference of the cognitive science society* (pp. 2926-2931).

Croson, R., & Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic literature*, *47*(2), 448-474.

Dohmen, T. J., Falk, A., Huffman, D., & Sunde, U. (2008). Are risk aversion and impatience related to cognitive ability? IZA Discussion Papers, No. 2735 .

Eckel, C. C., Grossman, P. J., Johnson, C. A., de Oliveira, A. C., Rojas, C., & Wilson, R. K. (2012). School environment and risk preferences: Experimental evidence. *Journal of Risk and Uncertainty*, *45*(3), 265-292.

Edwards, W. (1954). The theory of decision making. *Psychological bulletin*, 51(4), 380.

Epstein, S. (1994). Integration of the cognitive and the psychodynamic unconscious. *American psychologist*, *49*(8), 709.

Fishburn, P. C., & Kochenberger, G. A. (1979).Two-piece von Neumann-Morgenstern utility functions*. *Decision Sciences*,*10*(4), 503-518.

Flynn, J., Slovic, P., & Mertz, C. K. (1994). Gender, race, and perception of environmental health risks. *Risk analysis*, *14*(6), 1101-1108.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic perspectives*, 25-42.

Frey, R., Mata, R., & Hertwig, R. (2015). The role of cognitive abilities in decisions from experience: Age differences emerge as a function of choice set size. *Cognition*, *142*, 60-80.

Georgescu-Roegen, N. (1954). Choice, expectations and measurability. *The Quarterly Journal of Economics*, 503-534.

Gigerenzer, G., & Selten, R. (2001). Rethinking rationality. *Bounded rationality: The adaptive toolbox*, *1*, 12.

Goldstein, W. M., & Einhorn, H. J. (1987). Expression theory and the preference reversal phenomena. *Psychological review*, *94*(2), 236.

Gonzalez, R., & Wu, G. (1999). On the shape of the probability weighting function. *Cognitive psychology*, *38*(1), 129-166.

Halek, M., & Eisenhauer, J. G. (2001). Demography of risk aversion. *Journal of Risk and Insurance*, 1-24.

Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science*, 312(5782), 1900-1902.

Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior (No. w12006). *National Bureau of Economic Research.*

Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological science*, *15*(8), 534-539.

Hogarth, R. M., & Einhorn, H. J. (1990). Venture theory: A model of decision weights. *Management science*, *36*(7), 780-803.

Holt, C. A., & Laury, S. K. (2002). Risk aversion and incentive effects. *American economic review*, *92*(5), 1644-1655.

Hsee, C. K. (1996). The evaluability hypothesis: An explanation for preference reversals between joint and separate evaluations of alternatives. *Organizational behavior and human decision processes*, 67(3).

Hsee, C. K., Loewenstein, G. F., Blount, S., & Bazerman, M. H. (1999). Preference reversals between joint and separate evaluations of options: a review and theoretical analysis. *Psychological bulletin*, 125(5), 576.

Kahneman, D. (2003). A perspective on judgment and choice- mapping bounded rationality. *American psychologist*, 58(9), 697.

Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. *Heuristics and biases: The psychology of intuitive judgment*, *49*.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, 263-291.

Kunreuther, H., Novemsky, N., & Kahneman, D. (2001). Making low probabilities useful. *Journal of risk and uncertainty*, 23(2), 103-120.

Lattimore, P. K., Baker, J. R., & Witte, A. D. (1992). The influence of probability on risky choice: A parametric examination. *Journal of Economic Behavior & Organization*, *17*(3), 377-400.

Liberali, J. M., Reyna, V. F., Furlan, S., Stein, L. M., & Pardo, S. T. (2012). Individual differences in numeracy and cognitive reflection, with implications for biases and fallacies in probability judgment. *Journal of Behavioral Decision Making*, *25*(4), 361-381.

Lipkus, I. M., Samsa, G., & Rimer, B. K. (2001). General performance on a numeracy scale among highly educated samples. *Medical Decision Making*,*21*(1), 37-44.

Lubinski, D., & Humphreys, L. G. (1997). Incorporating general intelligence into epidemiology and the social sciences. *Intelligence*, 24(1), 159-201.

Ofstedal, M. B., Fisher, G. G., & Herzog, A. R. (2005). Documentation of cognitive functioning measures in the Health and Retirement Study. *Ann Arbor, MI: University of Michigan*.

Ostaszewski, P., Green, L., & Myerson, J. (1998). Effects of inflation on the subjective value of delayed and probabilistic rewards. *Psychonomic Bulletin & Review*, *5*(2), 324-333.

Payne, J. W., Laughhunn, D. J., & Crum, R. (1981). Note—Further Tests of Aspiration Level Effects in Risky Choice Behavior. *Management Science*,*27*(8), 953-958.

Petrova, D. G., Pligt, J., & Garcia- Retamero, R. (2014). Feeling the numbers: on the interplay between risk, affect, and numeracy. *Journal of Behavioral Decision Making*, *27*(3), 191-199.

Plott, C.R. (1996), Rational individual behavior in markets and social choice processes: the Discovered Preference Hypothesis, in Arrow, K., Colombatto, E., Perleman, M. and Schmidt, C., (eds.),*Rational Foundations of Economic Behavior*, Macmillan and St. Martins London, pp. 225–250.

Pratt, J. W. (1964). Risk aversion in the small and in the large. *Econometrica: Journal of the Econometric Society*, 122-136.

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior & Organization*, 3(4), 323-343.

Schoemaker, P. J. (1990). Are risk-attitudes related across domains and response modes?. *Management science*, *36*(12), 1451-1463.

Schwartz, L. M., Woloshin, S., Black, W. C., & Welch, H. G. (1997). The role of numeracy in understanding the benefit of screening mammography. *Annals of internal medicine*, *127*(11), 966-972.

Shafer, G. (1988). St. Petersburg Paradox. *Encyclopedia of statistical sciences*.

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological bulletin*, *119*(1), 3.

Stanovich, K. E., & West, R. F. (2000). Advancing the rationality debate. *Behavioral and brain sciences*, 23(05), 701-717.

Starmer, C. (2000). Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk. *Journal of economic literature*, 332-382.

Thaler, R. H., & Sunstein, C. R. (2003). Libertarian paternalism. *The American Economic Review*, *93*(2), 175-179.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *science*, *185*(4157), 1124-1131.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty*, 5(4), 297-323.

Tversky, A., & Fox, C. R. (1995). Weighing risk and uncertainty. *Psychological review*, 102(2), 269.

Tversky, A., & Wakker, P. (1995). Risk attitudes and decision weights. *Econometrica: Journal of the Econometric Society*, 1255-1280.

Tymula, A., Belmaker, L. A. R., Ruderman, L., Glimcher, P. W., & Levy, I. (2013). Like cognitive function, decision making across the life span shows profound age-related changes. *Proceedings of the National Academy of Sciences*, *110*(42), 17143-17148.

Van de Kuilen, G. (2009). Subjective probability weighting and the discovered preference hypothesis. *Theory and Decision*, *67*(1), 1-22.

Von Neumann, J., & Morgenstern, O. (1944). Theory of games and economic behavior.

Wakker, P. P. (2010). Prospect theory: For risk and ambiguity. Cambridge University Press.

Weber, E. U., Shafir, S., & Blais, A. R. (2004). Predicting risk sensitivity in humans and lower animals: risk as variance or coefficient of variation. *Psychological review*, *111*(2), 430.

Wehrung, D. A. (1989). Risk taking over gains and losses: A study of oil executives. *Annals of Operations Research*, *19*(1), 115-139.