DOCTOR OF PHILOSOPHY

The Impact of Visual Elements on Rational Decision Making During Risk Elicitation Tasks

Doukianou, Stella

Award date:
2017

Awarding institution:
Coventry University

THE IMPACT OF VISUAL ELEMENTS ON RATIONAL DECISION MAKING DURING RISK ELICITATION TASKS

By

Stella Doukianou

August 2017

A thesis submitted in partial fulfilment of the University's requirements for the degree of Doctor of Philosophy of Research

REGISTRY RESEARCH UNIT

ETHICS REVIEW FEEDBACK FORM (P48719)

(Review feedback should be completed within 10 working days)

Name of applicant: Stella Doukianou ....

Faculty/School/Department: [Faculty of Engineering, Environment and Computing] Computing & The Digital Environment ....

Research project title: Improving a risk assessment task by using interactive visual aids.

Comments by the reviewer

1. Evaluation of the ethics of the proposal:

Data is anonymised and consent forms are now appropriate for the format.

2. Evaluation of the participant information sheet and consent form:

.

3. Recommendation:

(Please indicate as appropriate and advise on any conditions. If there any conditions, the applicant will be required to resubmit his/her application and this will be sent to the same reviewer).

X	Approved - no conditions attached
	Approved with minor conditions (no need to re-submit)
	Conditional upon the following – please use additional sheets if necessary (please re-submit application)
	Rejected for the following reason(s) – please use other side if necessary
	Not required

Name of reviewer: Anonymous ....

Date: 22/01/2017...

Acknowledgements

There are a number of people without whom this thesis would not have been written, and to whom I'm highly grateful.

First of all, I would like to express my deepest gratitude to my director of studies Dr Ian Dunwell, who encouraged me, supported me and guided me throughout this process. Similarly, I would like to express my gratitude to Dr Pamela Kato for her guidance and inspiration and Dr Petros Lameras who provided me insight, support and praise.

I would also like to acknowledge the input of Joshua Cooper, CEO of Hildebrand Technology Ltd. Without his vision, this PhD would not have been possible.

I would like to thank my father Thanasis and my mother Marianna without their support and faith in me, I would not have undertaken this path; my sister Elena whose humour, support and understanding was invaluable to me all this time. I would like also to thank my partner in life Yiannis, whose support and patience brings fulfilment to my life and enables me to pursue my dreams. They all kept me going and this thesis would not have been possible without them.

I would also like to thank all of the students and members of Serious Games Institute at Coventry University who helped me with my experiments and were always keen to take part in my studies.

Finally, I would like to thank my friends and all those who influenced me, inspired me and provided useful feedback. A big heartfelt thank you.

Abstract

Understanding an individual's risk propensity may provide insight into decision-making processes involving risk, and predict behaviour under various policy interventions. There are different ways to measure risk propensity. One of the principal methods is through risk lottery elicitation tasks where validity issues may be caused by participants' irrational choices. Participants' irrational choices are due to misunderstanding of the lotteries in these risk elicitation tasks which are presented as numerical probabilities. Such choices produce a high inconsistency rate as well as invalid predictions of individuals' risk propensity.

This thesis aims to help people understand the lotteries of such risk lottery elicitation tasks and assist in their rational decision making with the use of interactive approaches. To achieve this, the thesis shows the results of three quantitative empirical studies based on the overarching research question: "Can incorporating visual elements and interactive approaches with graphs in a risk elicitation task influence rational decision making?". The first study investigated whether by introducing gamified elements into a risk lottery elicitation task, participants would be influenced to choose rationally. The second study explored the underlying factors, which influence participants' rational choices in options of a risk lottery elicitation task. The third study examined whether using interactive visual media would influence rational choices in a risk lottery elicitation task.

Results from the first experiment with 32 participants showed that, despite an acceptable level of the game's usability, (as revealed from a validated usability satisfaction questionnaire, $M = 6.86$ of maximum 10 and $SD= 0.62$), there was no observable correlation between participant' choices in the game and the standardised task, $r (32) = .120, p < .511$. Results from the second experiment with 60 participants, showed that participants who scored higher in a validated numeracy scale were more likely to answer rationally in the lottery options of a risk elicitation task, $X^2 (1, N= 60) = 4.176, p = .041$. Similarly, participants who scored higher in a validated

cognitive reflection task had greater chance of choosing rationally in the Holt and Laury task, $p < 0.05$. Furthermore, when participants used graphs to reason their choices in the task, they were more likely to choose rationally in the task, $p < .020$. Results from the third experiment with 225 participants, yielded significant evidence that both people who scored lower in the validated numeracy assessment scale and higher, chose more consistently when they interacted with pie charts that presented the lotteries than when they were presented with lotteries as numerical probabilities or passive pie charts, $p < .00$.

The first experiment demonstrates high variance arising from inconsistency in the choices in both tasks. The results highlight the issues with irrational choices in risk lottery elicitation tasks and show that, despite evidenced usability the inconsistency of participants' choices persisted in a gamified context similar to the standard Holt and Laury task. The second experiment results demonstrate that the use of visual methods, which may extend to game, is more likely to result in rational and therefore consistent choices for people with numerical difficulties or impulsive thinking. The third experiment results demonstrate that the use of interactive visual media has potential to increase consistency of choices in risk lottery elicitation tasks especially for people with numerical difficulties.

These findings have implications for the future implementation of risk elicitation tasks that involve lotteries to convey risky situations particularly for audiences with score lower in validated numeracy and cognitive reflection tasks. The evidence provided by this thesis supports the assertion that by providing an interactive graphical presentational way of the lotteries in the task, the consistency of results can be increased. Interactive approaches with visual media were shown to assist on rational decision making for people who scored lower in the validated metrics of numeracy and impulsive thinking confirming empirical evidence which suggested the use of external representations to understand probability problems. Thus, these experiments' results extended the theories supporting the use of external representations to solve probability problems by using specifically pie charts as external representations that assist the rational choice in the Holt and Laury task. This interactive approach, which could be extended through games, may reduce individuals' irrational choices in risk lottery elicitation task and thus allow the accurate estimations of their risk propensity. The latter would help to provide more accurate predictions of individuals' behaviour under various policies, in the context of financial investments.

CONTENTS

LIST OF TABLES

List of figures

1. INTRODUCTION

Risk elicitation tasks are used in experimental procedures by researchers and policy makers to determine an individual's risk propensity in a variety of contexts. In the scope of this thesis, these tasks are examined as applied in a financial uncertainty context. A variety of such tasks have been developed over the years by researchers in the field to elicit risk propensity using the resulting data to estimate someone's decision-making and economic behaviour. Since many economic models rely on such data to make accurate predictions, measuring risk preferences effectively is important not only for individuals' future decision-making but the heterogeneity across the population (Dave et al. 2010). For example, decisions related to investment choices, on how to allocate assets or how much to self-insure or options of insurances, are choices that can be predicted using data from risk elicitation tasks (Guiso and Paella 2004; Hill and Viceisza 2012). This section presents the issues with risk elicitations tasks so far with the most profound issue of participants' irrational choices in lottery tasks, to portray the identified need for another methodology that gamified and interactive assessments could offer.

The literature review presented in Chapter 2, highlights the issues that principal methods of risk assessment in a financial context have shown to have. These issues primarily involve the lack of a differentiating point between risk averse and risk neutral individuals, complexity of the experimental procedure that leads to participants' irrational choices in the task, difficulty to implement the task outside of a laboratory setting and lack of significant predictive value of the task (Eckel and Grossman 2002; Dave et al. 2010). The review then explores the empirical evidence on games that could improve, as it has been shown, the above mentioned issues of the risk elicitation tasks by using visuals and simplifying the complexity of the experimental procedure, implementing questions from questionnaires in a non-invasive way to incorporate differentiating points of risk averse and risk neutrals and distributing behavioural tasks through mobiles which allow the implementation of the task outside the laboratory (Dahling, Whitaker and Levy 2009; van Lankveld et al. 2011; Brown et al. 2015). Even though, in the context of this thesis, the gamification aspect is limited while visualisation and graphical representation of data is the main focus of this research project, the literature review in assessment games highlights the importance of visualisation that could extend beyond games. Following this introduction to risk elicitation tasks, their issues and how gamification was shown to improve these issues, then the literature review focuses on the issue of complexity of the most popular risk elicitation task developed by Holt and Laury (2002), as the core focus of this thesis. The Holt and Laury task (HL) (2002), which has been the most popular task according to the amount of times that has been used from different empirical studies (Anderson and Mewllor 2008; Eckel and Grossman 2002; Dave et al. 2010; Nielsen and Zeller 2013), is considered to have a strong predictive value of someone's risk preferences in similar choices. The way it is formulated can indicate people's irrational choices, which reflect an inconsistent pattern and misunderstanding of the task and exclude them from the data. Even though the Holt and Laury task has the highest predictive value of all the elicitation tasks, it has been criticised of being complex because it involves ten decisions between lotteries with probabilities ranging from .1 and .9 (Table 1). These lotteries presented as numerical probabilities were shown not to be easily comprehended by people with limited numerical skills (Dave et al. 2010; Galarza 2009; Eckel, Engle-Warnick, Johnson, C. 2005). Irrational choices to a risk lottery elicitation task lead to inconsistent choices as the way the task is designed, at the beginning of the task (top rows), Option A has a safer pay off than Option B. But the payoffs of the two lotteries gradually change leading to a reversal in payoffs as the user moves down to the bottom rows (the payoff becomes riskier in Lottery A and safer in Lottery B). However, if the participant changes lottery options from row to row, it leads to an inconsistent pattern. Inconsistent choices can provide misleading information about someone risk attitude and therefore must be excluded from the analysis. Therefore, the literature review on Chapter 2 continues by discussing the potential factors, which influence people's irrational choices on this task to investigate the issue of its complexity. By exploring the factors, it highlights the issue of people's difficulty to understand the numerical presentation of lotteries in risk lottery tasks, which involve probabilities. This thesis defines numeracy level as someone's score in validated numeracy metric that assesses calculations related to probabilities. It highlights the problems with the presentation of probabilities while exploring the use of visual thinking and external representations to improve people's understanding of them. Finally, the literature review, presents the variety of visuals that have been used in risk elicitation tasks to reduce participants' irrational choices while reviewing the graphs which can be used to express probability problems.

The literature review presented in Chapter 2 identifies empirical evidence showing benefits of using gamified and interactive approaches such as prompting individuals to use external graphical representations to come up with rational decision making while choosing from the lottery options in these risk elicitation tasks. The empirical evidence presented briefly in this section and elaborated further in Chapter 2, leading us to the inevitable question "Can incorporating visual elements and interactive approaches with graphs in a risk elicitation task influence rational decision making?".

This is the high-level research question this thesis answers through a series of quantitative experimental studies. Therefore, according to gaps in current evidence-base, the first experiment investigated whether gamification of a risk elicitation task would influence participants to choose rationally. The second experiment investigated the underlying factors, which influence participants' rational choices in options of a risk lottery elicitation task. The third study examined whether using interactive visual media would influence rational choices in a risk lottery elicitation task.

1.1 Gaps in Current Evidence-Base

This research has been conducted on the basis of three main gaps in current evidence-base which exist primarily in the need for a new methodological approach in the risk elicitation that would reduce participants' irrational choices. These main gaps are investigated through three quantitative studies.

1. Empirical findings have identified that the variety of risk elicitation tasks were shown to have certain issues which need to be addressed and new methodologies need to be acquired. Taking into account the benefits and disadvantages of each previous risk elicitation task, an improved measure should: 1) be simple enough so people with numerical difficulties could understand it (Eckel and Grossman 2002; Dave et al. 2010), 2) involve a differentiate point where risk neutrals could be differentiated from risk averse (Holt and Laury 2002), 3) be flexible and not being used only in laboratory settings and finally 4) have great predictive value of individuals choices. According to the literature review presented in Chapter 2, Holt and Laury (2002) risk elicitation task involves a differentiate point, could be used outside of a laboratory and could have great predictive value when is understood. Therefore, only Holt and Laury's complexity need to be resolved to improve the task's predictive value.

In the light of the recent tendency to use games to assess behaviour, little work has been done to explore how games can be best used to enhance risk elicitation exercises. Considerable motivation exists to examine this (Dahling, Whitaker and Levy 2009; van Lankveld et al. 2011; Brown et al. 2015). Games were shown to induce certain behaviours; risk elicitation tasks could be enhanced through better understanding of the specific behaviour.

2. Considering that empirical studies with Holt and Laury task showed high inconsistency rate in participants' choices (Cleveland, Harris and McGrill 1982; Charness and Viceisza 2012; Galarza 2009), further investigation is required into the underlying process of participants when reasoning these options. Several studies showed that numeracy (Dave et al. 2010), cognitive style (Frederick 2005), and educational level (Charness and Viceisza 2012) have an effect on people's inconsistent answers in Holt and Laury task. However, the empirical evidence, so far, are not clear on whether numeracy level, educational attainment, cognitive thinking style influences the most someone's choices in such tasks. If these factors are investigated in depth then potentially an assessment tool could take into account these limitations to help people choose rationally and therefore consistently. For example, if automatic thinkers tend to make more irrational choices compared to reflective thinkers, the assessment tool could involve a procedure that 'forces' automatic thinkers to reflect more on the choices. To address the issues caused by these factors, the interactive approach of reasoning using external presentations has shown to help people to understand and gain insight into the concept of probability problems which are similar to the lotteries presented in risk elicitation tasks like Holt and Laury (Heibert and Carpenter 1992; Greeno and Hall 1997). This interactive approach has not been investigated in conjunction with lottery risk elicitation tasks, which can introduce new ways for participants of filling in the risk elicitation tasks and choosing irrationally. Therefore, the investigation of the factors that influence participants' irrational choice as well as the interactive engagement with external presentational ways could lead us to a different experimental approach.

3. Finally, several studies have experimented with graphs and illustrations to improve the Holt and Laury task and increase participants' choice consistency rate especially for those that have limited numerical skills (Camerer 1989; Habib et al. 2016; Bauremeister and Musshoff 2016). None of these ways though have managed to eliminate participants' irrational choices. According to findings from literature, which support that visuals help people with numerical skills to reason correctly probabilities (Peters 2008) and also that interactive engagement with probability options, can lead to more successful results (Stylianou and Silver 2004), an interactive approach to reasoning while choosing the options in such lotteries has not emerged yet. Therefore, an interactive use of visuals that replace the conventional way of lotteries presentation, which can extend to games, is more likely to result in consistent choices.

The gaps in knowledge, as mentioned earlier, are described in depth in the literature review part of this thesis and explored through the experiments. The literature review and the gaps in knowledge formulate the hypothesis as outlined in the next chapter, which leads to three research sub–questions

1.2 Scope

The primary focus of this thesis is to examine whether gamified and interactive approaches can influence rational decision-making in a risk elicitation task. Interactive approaches, in the context of this thesis is defined as the interaction between user and visual aids; the flexibility for the user to manipulate a graph to present a numerical probability. A series of experiments are conducted based on empirical evidence. The first experiment which is the pilot study, incorporates a risk elicitation task into a gamified context with visuals instead of text for the lotteries. The construct validity of the assessment tool is examined. In the scope of this thesis only construct validity is examined. Further building on findings that show high inconsistency rate on people's choices in the standardised lottery risk elicitation task of Holt and Laury and the gamified approach, the next experiment seeks to understand how these tasks can be deconstructed such that the individual components of these tasks, linked to outcomes, can be more clearly understood. Therefore, the factors that influence participants to choose rationally risk lottery tasks are investigated. Among these factors, the use of external representations to reason the options of a risk lottery task is also examined to determine whether they can assist in rational decision making. Drawing on findings from this study and empirical evidence, another experiment is conducted to determine whether engaging users to interact with graphs while choosing the options would improve their understanding of the task as well as their rational choices. This thesis proposes a new interactive method for the risk lottery task. It poses a central research question then defines and tests three main hypotheses within it by manipulating a variable and observing the outcome. These findings are reported and related to the main Research Question in Chapter 5.

1.3 Contribution of this Thesis

Through a series of three quantitative experimentations directed at identifying if interactive approaches can influence rational choices in a risk elicitation task several insights have been gained and contribute to the empirical evidence presented in the literature review. These contributions are divided into three subsections to reflect each experiment.

- The first experiment highlights the issues with inconsistent choices in lottery risk elicitation tasks, and shows that the inconsistency persists, despite the good usability level, even after the visualisation of the lotteries as roulettes in a gamified context. There were no significant results from the correlation between the game and a validated risk elicitation task. However, since it was a pilot study and the sample size was small (32 participants), there are no significant contributions that could be generalised. Nevertheless, the persistence of participants irrational choices after the visualisation of the numerical lotteries points towards to further exploration of the issues participants have with the display format of these lotteries. Since games are such rich interactive media that do not allow to test a specific element (display format of the lotteries), further investigation could identify the appropriate type of visuals that should be used for risk elicitation tasks before implanting into a game.

- Second experiment demonstrated that the use of graphs from the participants is more likely to result in consistent choices. This finding confirms empirical evidence suggesting that using external representations (Zahner and Corter 2002) and more specifically graphs (Hegarty and Kozhnenikov 1999) leads to successful problem solving in probabilities. These findings also extend those theories by showing a link of using successfully external representation of graphs for rationally decision making in lottery choices which is a different context to probability problems in school settings where external representations have been used before.

- The third experiment showed that the use of interactive graphical media increased the consistency of choices in a risk elicitation task and thus tasks' predictive ability by influencing participants to choose rationally. This is linked to empirical evidence suggesting that people are poor in evaluating the risks from pie charts (Elting et al. 1999), however when these pie charts are drawn by the individuals it is more likely to evaluate the probability correctly, as it was suggested by Hegarty and Kozhnenikov (1999). Third experiment's results, highlight that when individuals draw the pie charts themselves, they do not have the same difficulty in estimating proportions as when they are presented with pie charts that need to evaluate. The results also confirm the theory suggested by Hegarty and Kozhenikov (1999) that graphical representations help people in solving successfully probability problems while taking it a step further by specifically highlighting the importance of pie charts used as external representations. Future research could explore the use of other graphical representation in lottery problems. The finding that interactive approaches with pie charts help individuals that were found to score lower in the Numeracy Scale developed by Lipkus, Samsa and Rimer (2001) or in the Cognitive Reflection task (Frederick 2005) showing an automatic thinking style, to choose rationally and reduce their inconsistent choices in lottery tasks which reflects their understanding of the methodological procedure and their actual risk preferences. Therefore, the task's ability to predict individuals risk propensity is increased. Primarily, the use of external representations of pie charts could solve the problem of inconsistent choices that was identified in a variety of experimental studies with audience who had limited numerical skills (Charness and Viceisza 2012; Galarga 2009; Jacobson and Petrie 2009). These findings open up new paths of the tasks implementation in the field that accurately can predict farmers' risk-taking propensity which could be used to formulate policies for technology adoption or crop selection (Guiso and Paella 2004) Similarly, the task could accurately predict users' risk propensity accurately in any type of investment choices such as allocation of assets (Hill and Viceisza 2012). The interactive pie chart approach could be transferable to any risk elicitation task that demands mathematical sophistication to convey risk. The finding that people understand probability problems when interacting with graphs is not surprising. Corter and Zahmer (2007) have found that students using visual external representations themselves help them to understand and solve the probability problem. Hegarty and Kozhnenikov (1999) have highlighted that graphs help problems solvers more than any other type of visual such as images. Therefore, in the case of risk elicitation tasks, if participants understand the risk by interacting with graphs then the tasks ability to predict their choices is automatically increased. To that end, predictions of a number of individuals would produce accurate outcomes to help the development of certain financial policies. Also, this approach makes presentation of similar probability problems understandable and easily applicable to any type of audience, which could be used by several areas such as health to demonstrate the risks to the patients in a way that could understand the risks, and benefits without being misled.

1.4 Thesis Structure

The remainder of this thesis is structured into the following chapters.

- Chapter 2 presents an overview of the previous research and related work for risk elicitation tasks. It also defines the terms of risk taking, numeracy level and education in the scope of this thesis. Based on challenges created from past elicitation tasks, the literature review provides an overview of gamified assessment tools drawing on the strengths and limitations on previous risk tasks. The next section in the literature review discusses the factors that influence people's inconsistent answers in risk elicitation tasks, providing evidence from previous empirical studies. This section also describes the methods used for presentation of probabilities as well as the ways people use external presentations to think visually and approach a problem. The third section of the literature review presents the effectiveness of using graphs to describe probabilities and how interactive engagement with visuals can help people understand such problems.

- Chapter 3 outlines the general research methodology approach contemplating on research approach, tools and analysis of the results. It presents the methods used for the three quantitative studies. The first study investigated whether by introducing gamified elements into a risk lottery elicitation task, participants would be influenced to choose rationally. The second study explored the underlying factors, which influence participants' rational choices in options of a risk lottery elicitation task. The third study examined whether using interactive visual media would influence rational choices in a risk lottery elicitation task.

- Chapter 4 presents the experimental results from the three experiments. Results from the first experiment with 32 participants showed that despite an acceptable level of the game's usability, (as revealed from a validated usability satisfaction questionnaire, a $M = 7.11$ of maximum 10 in Likert scale and $SD.= 0.5117$), there was no observable correlation between participant' choices in the game and the standardised task, $r\ (32) =.120$, $p < .511$. Results from the second experiment with 60 participants, showed that participants who scored higher in a validated numeracy assessment scale were more likely to answer rationally in the lottery options of a risk elicitation task, $X^2(1) = 4.026$, $p = .045$. Similarly, participants who scored higher in a validated cognitive reflection task had greater chance of choosing rationally in the task, $p < 0.05$. Furthermore, using graphs to reason the options could predict that they would answer rationally in the task, $p < .000$. Results from the third experiment with 225 participants, yielded significant evidence that people who scored lower in the validated numeracy assessment scale chose more consistently when they interacted with pie charts in a gamified context than when they were presented with lotteries as numerical probabilities or passive pie charts, $p < .000$.

- Chapter 5, summarises the overall contributions, relates them to the Main Research Question, and provides a discussion on the different methods employed and possible future work. The findings from the first experiment highlighted that participants' inconsistent choices persisted after the visualisation of lotteries. The second experiment revealed the using external graphical representation can assist in rational decision making in lottery options for people who scored lowered in the validated metrics of numeracy and cognitive thinking style. The third experiment demonstrated that using interactive pie charts can assist in rational decision making in risk lottery elicitation tasks. These findings have implications for the future implementation of risk elicitation tasks that involve lotteries to convey risky situations especially for audiences with numerical difficulties or impulsive thinking by providing an improved presentational way of the lotteries in the task. Interactive approaches with graphs were shown to assist on rational decision making for people who scored lower in the validated metrics of numeracy and impulsive thinking confirming empirical evidence which suggested the use of external representations to understand probability problems. Thus, this experiments' results extended the theories supporting the use of external representations to solve probability problems by using specifically pie charts as external representations that assist the rational choice in the Holt and Laury task. This interactive approach, which could be extended through games, may reduce individuals' irrational choices in risk lottery elicitation task and thus allow the accurate estimations of their risk propensity. The latter would help on accurate predictions of individuals' behaviour under various policies in the context of financial investments.

1.5 Summary

Risk preferences are an important part of someone's financial portfolio. People's differences in risk preferences can affect various parts of the general economic level of a country e.g. the price of risk capital (Palson 1996) or the crop insurance markets (Hill and Viceisza 2012). Therefore, there is an emergent need of an accurate risk elicitation task that would be able to produce valid results to predict someone's choices or decisions. Among the variety of issues that risk elicitation tasks were shown to have which involve complexity of procedure, lack of differentiating point, lack of significant predictive value and difficulty to implement outside laboratories, this thesis investigates the issue of complexity as it has been highlighted from experiments with the Holt and Laury task (2002). This task, even though is considered the most popular and with the highest predictive value when is understood, it has also has been criticized of being complex as it has been revealed by studies showing high inconsistent rates in people's choices, due to their irrational choices. Consequently, the predictive value of the task has been violated. This inconsistency is due to the mathematical sophistication that the task requires. A variety of methods have been employed to simplify the Holt and Laury task. However, none of them have managed to eliminate people's inconsistent answers and help them choose rationally. The tendency to utilise gamified approaches to assess and change behaviour the latest years has been obvious, especially in behavioural and cognitive tasks. Numerous examples of games have been used to assess behaviour, personality, cognitive functions and business effectiveness. Despite some limitations that need to be overcome, games can offer an alternative method of assessment in place of conventional pen and paper questionnaires that can simplify complex mathematical problems. Significant evidence exists supporting that using gamified elements can simplify the procedure of Holt and Laury task. A series of three experiments which are presented in the following chapters are looking to answer the high-level research question whether incorporating visual elements and interactive approaches with graphs in a risk elicitation task influence rational decision making. The introduction has provided an overview of the topic area, scope and findings of this thesis. The following chapters detail the background, methodology and results in depth.

2. LITERATURE REVIEW

2.1 Introduction

This chapter provides an overview of literature related to the topics of risk taking, gamified assessments, irrational choices in risk elicitation tasks and visualisation ways of lotteries in risk elicitation tasks. The literature review is divided into five main sections.

The first section presents the definition of risk propensity and its contributing factors and an overview of risk taking on theoretical groundings regarding the importance of assessing someone's risk taking tendencies as well as the type of risk taking examined in this thesis. Additionally, previous risk elicitation tasks are being described outlining their advantages primarily of differentiating point that distinguishes risk neutrals from risk averse and good predictive value and disadvantages notably of complexity issues and difficulty to implement outside of laboratory settings.

The second section provides a definition of the gamification aspect discussed in this thesis. Even though, the gamification aspect is limited for the purpose of this thesis, a brief literature review of in-game assessments is presented in this section, to highlight the benefit of using visuals in an interactive setting. The literature review of games used as assessment tools highlights their benefits of using visuals to simplify procedures and potential of incorporating measurements from pen and paper questionnaires while outlining their shortcomings as well primary of using unnecessary elements that distract the users.

The third section explores the main elements of a lottery risk elicitation task to understand them further. It also presents the factors that influence people's choices in Holt and Laury tasks, the most profound issues with lotteries presented as probabilities and the potential of using visual thinking to understand the underlying process of people's decision-making procedure. It continues with investigating different means of external representations.

The fourth section of the literature review explores the use of graphs for presenting probabilities in Holt and Laury task. It provides a review of the most popular graphs, the advantages and disadvantages and explores the use of visual aids and interactive engagement with them in problem solving. Finally, outlines the studies with Holt and Laury task that replaced the numerical probabilities with visual aids.

The final section summarises the key points of each part of the literature and the formation of the high-level Research Question.

2.2 Risk taking

The following section presents the definition of risk-taking and its contributing factors as well as the importance of assessing it. Following that, a literature review of previous risk elicitation tasks is presented.

2.2.1 Definition of risk propensity, gamification and its contributing factors

Sitkin and Pablo (1992) suggested that the two core inputs to risk taking are risk perception and risk propensity. The definition of risk propensity has important implications for practical insights into the underlying factors that motivate people to engage in risky behaviour. The concept of risk propensity has a variety of definitions depending on the theory that is based on. In the context of this thesis, the terms of risk propensity and its contributing factors are presented.

In this thesis, risk propensity is defined as someone's tendency to choose more often riskier lotteries compared to safer or vice versa. A risky lottery is considered a lottery that includes two possible scenarios where the smaller probability of winning scenario involves a significantly larger pay off compared to the bigger probability of winning scenario. For example, 10% of winning £50 compared to 90% of winning £15. Risk propensity is also defined in the context of decisions related to financial investment choices. Therefore, this definition of risk propensity is considered a domain specific attribute. For example, an individual's risk propensity is not the same in finance and health. This definition has been used in several studies, which supported that risk propensity is domain specific, and when assessed by the appropriate task it can predict behaviour and decisions in this domain. For instance, Holt and Laury (2002) developed a risk elicitation task, which includes two lottery options (one safe and one risky) that people need to choose from, over ten rows (Table 1).

Table 1. Multiple Choice Lottery task (Holt and Laury 2002)

Some materials have been removed due to 3rd party copyright. The unabridged version can be viewed in Lancester Library - Coventry University.

Empirical evidence showed a link between choices in Holt and Laury task and decisions related to insurances (Hill and Viceisza 2012). Other empirical studies have used this definition of risk propensity to relate it with the choices that farmers do regarding crop selections and various investment choices (Bauremeister and Musshooff 2016). However, there are several critiques for the definition of risk propensity, which supports that it is a domain specific attribute. These consider risk propensity to be a personality trait (Mishra and Lalumiere 2011). This definition has been supported by Eysenck and Eysenck, (1985) who argued that risk propensity is a result of biological differences or early child experiences. Hence, people had different risk propensities not according to the specific context but based on their individual characteristics. For example, efforts have been made to link risk propensity with personality characteristics such as sensation seeking (Zukcerman, Eusenck and Eysenck 1979). This support though proved to be problematic for the definition of risk propensity as a personality trait. In both laboratory studies and managerial contexts conducted by Schoemaker (1990) using the same risk elicitation tasks, individuals were shown to be inconsistent risk takers or averse across different domains. Similarly, MacCrimmon and Wehrung (1986) indicated that managers have different risk attitude towards making decisions using money that are personal versus their companies. The expected utility-based risk elicitation tasks have not had such a success in predicting people's choices in a range of situations (Bromiley and Curley 1992). Despite these critiques whether risk propensity is stable across domains or not, this thesis provides a point of departure towards investigating risk propensity for financial investment decision-making.

Hence, in the scope of this thesis, the assessment of this domain-specific risk taking is investigated. Key limitations arising on the assumption that risk propensity is a domain specific attribute, primarily that it is not stable and difficult to assess as it can be affected by many external factors are considered in Chapter 5.

Even though risk propensity is considered domain specific, the perception of risk is affected by several factors, which in the scope of this thesis are numeracy level, cognitive thinking style and education are investigated. In this thesis, education is defined as the completion of full years of schooling. This definition of education has been widely used by studies where the impact of education on farmers in rural Ethiopia is investigated towards risk propensity and compared farmers who completed their minimal educational attainment, which was considered the full years of schooling and those who have not (Knight, Wei and Wolderhanna 2003). The definition of education though has not been interpreted similarly in all studies that included the definition of a variable that can predict a specific behaviour. For instance, Boughera, Gassman and Piet (2011) who assessed a group of farmers in France for their risk-taking tendencies considered them educated as they have attained a college degree. Therefore, for Boughera, Gassman and Piet (2011), the categorization of an individual as educated is based on the individual's completion of a college degree. However, this definition arises some questions regarding differences between educated, uneducated and illiterate. The differences between definitions of educations have also cultural aspects, which derive from the different terminology of education in developed and developmental countries where education is now available (Suárez-Orozco and Qin-Hilliard 2004). Despite the different interpretation about the definitions of education, in the scope of this thesis, someone who is educated is considered anyone who has finished school and higher educated someone who completed college or university studies. Any limitations arising from this definition primarily that this definition is subjective to culture and to the general educational system of each country, would also be discussed in Chapter 5 of this thesis.

Another contributing factor of risk propensity that is discussed in this thesis is the level of one's numerical ability. Research in numeracy suggests that people differ a lot in their abilities with numbers and some of them might be innumerate (Lipkus, Samsa and Rimer 2001). In the scope of this thesis, someone is considered numerate when can complete basic arithmetic calculations that convert percentages to proportions or estimations of probabilities. This thesis defines numeracy level as someone's score in numeracy scale that assesses calculations related to probabilities. This definition has been supported by Lipkus, Samsa and Rimer (2001) who used an 11-item numeracy scale to assess the numerical level of college students. This scale was also used by Peters (2008) showing the connection between numeracy level and decision-making process. However, this definition of numeracy is not supported by every study, and there are multiple numeracy scales that define numeracy level differently. For example, A'Hearn, Batenand and Crayen, (2009) primarily noted that numeracy level is defined as the general ability of calculation skills in all aspects of arithmetic knowledge which differs between eastern and western societies. A'Hearn, Batenand and Crayen, (2009) explored the differences in the numeracy levels of all global regions and long-term growth of the country. In another study, the numeracy level was assessed by the Wide Range Achievement Test understanding, which includes literacy math problems to investigate the patient understanding of food labels in U.S (Rothman et al. 2006). The Wide Range Achievement Test includes both literacy questions and math problems and reflects the numerical training that is acquired in western countries mainly. Nonetheless, the definition of numeracy level and the measuring method differ globally. However, in the scope of this thesis numeracy level is measured from the aspect of risk and chance, which includes the estimation of probabilities, and is based on the conversion of probabilities to proportions and the opposite. Any limitations related to measurement of numeracy level according to the Numeracy Scale metric that is used in this thesis, would be discussed in Chapter 5.

Finally, the concept of rational choices needs to be clarified in the scope of this thesis. Rational choices are defined as the consistent pattern of switching at some point within the ten rows from Option A to Option B in the standardised task of Holt and Laury and not changing choices in each row or choosing just one option for the whole task. This definition has been supported by several studies, which showed a pattern of participants' inconsistent choices in the Holt and Laury task (Jacobson and Petrie 2009; Eckel and Grosman 2009; Dave et al. 2010). Therefore, in this thesis the number of irrational choices is reflected as an inconsistent pattern and a noise in data. Noise in data, in the context of this thesis, is defined as the participants' random choices in the Holt and Laury task which indicate that participants either change lotteries in each row or choose only one Option between the two lotteries over the ten rows (Table 1). These random and irrational choices are reflected in the data as noise. However, there is also research that defines people's irrational choices as errors of judgement or choice (Tversky and Kahneman 1981). Tversky and Kahneman support that these errors of judgements are not necessarily irrational but influenced by the way the choice is framed or the context within is communicated. Even though, this point is valid is out of the scope of this thesis, which focuses on people's rationality to choose the options in the Holt and Laury task consistent with the methodological procedure of the task and switch at some point lotteries.

In summary, there are multiple definitions of risk-propensity. This thesis explores people's irrational choices under the definition that risk propensity is a domain specific attribute but it can be affected by people's numeracy level, cognitive thinking style and presentational form. Even though, the concepts of risk propensity, numeracy, rationality and education that are discussed in this thesis have been clarified; it is worth mentioning that these contributing factors are not the only ones influencing risk. There other contributing factors such as delayed gratification, marital status and gender that have shown to influence risk propensity (Romer et al. 2013; Grable 2000; Byrnes, Miller, and Schafer 1999). However, the investigation of those is out of the scope of this thesis, which is focusing on the irrational decision making in risk elicitation tasks where the options are presented in the form of numerical probabilities.

2.2.2 Defining risk taking

Debates exist regarding the definition of a risk-taker. For example, a risk taker is someone who will undertake an activity with an uncertain outcome that is also highly likely to have significant negative consequences for the individual (e.g., placing a very large bet on a roulette wheel) (Arrow 1965). Similarly, for Zuckerman and Kuhlman (2000), someone who is risk seeker would be more likely to engage in activity that might provide a short-term positive experience at the risk of severe and long-term negative repercussions. For example, this is the case of drug addicts because they prefer a moment of "high" despite the possibilities of arrest, overdose, and loss of social relationships or employment job. Sitkin and Pablo (1992), in their theory of risk propensity, portrayed the risk taker as someone who has the tendency of taking risks consistently no matter the context. According to Grable (2008), an individual's risk-taking attitude is defined by the maximum amount of uncertainty that one is willing to accept when making a decision. For instance, risk seekers are those who choose to use illegal drugs when they are aware that there are a lot of probabilities to get addicted. All the above definitions illustrate two sides of risk taking: one that involves uncertainty, which is often expressed as probabilities (numbers) and the other that involves the consequences; to what extent it matters for the individual (Hilson and Murray-Webster 2006). What is revealed from these definitions is that in order individuals to evaluate the extend that matter for them; they should be able to assess the uncertainty of the choice which in risk elicitation tasks often is expressed as probabilities. Since, the extend that matters to someone is a subjective point of view; everyone should evaluate the uncertainty accurately. Risk taking two sides involve uncertainty, which is often expressed as probability and the consequence of this probability (to what extend it matters for the individual). The scope of this thesis is focused on the aspect of risk taking's definition that is related to uncertainty and how this is perceived.

Nonetheless, there are a variety of risk taking attitudes among people. For example, if a group of people will be faced with the same investment choice, different decisions will be made; some people will be risk averse; other will be risk neutral while other risk seekers. These are words, which describe levels of attitudes towards uncertainty (Hilson and Murray-Webster 2006). Even though, these labels enable us to classify people's risk tolerance, risk attitude acts in a continuous spectrum (Hilson and Murray-Webster 2006). Perceptions of risk are argued by Hilson and Murray-Webster (2006) to fall in the continuum, as illustrated in Figure 1. However, someone to be classified as a risk seeker or risk averse or risk neutral, it needs to have understood the level of uncertainty correctly. This means that the individual has to evaluate the potential options accurately.

Figure 1. Spectrum of Risk Hillson and Murray-Webster, 2006

In the context of this thesis, risk taking is measured regarding financial investment decision making and more specifically lottery and investment choices, asset valuation, contracts and insurance which have been examined by several researchers in the past (Pratt 1964; Arrow 1965). These decisions are most often assessed by risk elicitation tasks, which involve Multiple Lottery Choice List between safe and risky options that people are asked to choose from. The ability to assess people's risk preferences in such choices is central for the prediction of financial and economic decisions. For example, in countries that depend on the agriculture sector, individuals' (farmers) risk preferences influence decisions related to crop insurance markets (Hill and Viceisza 2012) or pricing of the technology used (Purvis et al. 1995) and therefore the adopted policy of the country. Consequently, measured risk propensity in financial investment decision making choices can offer the potential of predicting future behaviour or options in this context.

For the purposes of this thesis, risk taking is defined by the maximum amount of uncertainty someone is willing to accept when making a decision in the context of financial decision making as it was suggested by Grable (2008), because it highlights the two aspects of risk; the uncertainty and the extent, which matters to the individual. Therefore, the propensity someone has to take a risk, in the scope of this thesis, is measured by the amount of times a person chooses a risky lottery option compared to safe lottery option in the risk elicitation task of Holt and Laury (2002). This thesis focuses on the aspect of uncertainty and how this is perceived in the context of financial decision-making choices. However, since risky financial decisions influence many aspects of individuals' everyday life, over the years, there have been developed a variety of tasks. The next section provides a literature review of the most popular risk elicitation tasks that have been developed so far to highlight the reasons that the Holt and Laury task was selected as risk elicitation task in this thesis.

2.2.3 Previous studies of risk assessments

Numerous research studies have focused on measuring people's propensity towards risk taking. Hence there are many methods that experimentally assess individual's risk attitude. So far, the three principal ways of assessment tools either involve hypothetical self-report questionnaires describing risk-taking mindsets/situations, or hypothetical choice problems, and computerised methods (Rohrmann 2005).

Such questionnaires, which rely on the individual's ability to self-report (Rohrmann 2005), present hypothetical risk-taking scenarios, asking people to rate their willingness to take risks in those situations in a 10-point Likert scale with one completely unwilling (risk averse) and ten completely willing (risk taker) (Charness, Gneezy and Imas 2013). These questionnaires measure risk taking in various contexts such as gambling to smoking i.e. invest one day's income on a sports bet or ask if they would engage in an affair with a married man/woman (DOSPERT Scale, 2006). The most well-known questionnaire that includes a variety of domains is the"domain-specific risk taking scale" (DOSPERT) developed by Weber et al. (2002). The full DOSPERT scale contains 40 items and includes the domains of health, social, ethical, recreational, gambling and investment. The main advantage of this type of risk assessment method is that is simple and understandable by everyone (Charness, Gneezy and Imas 2013). However, self-reported assessments have been questioned as they may be influenced by various distortions. These various distortions include difficulties in self-evaluation; self-esteem that biases objective self-evaluation, the self-representation, which the individual is willing to show to other, and the implicit expectations have been linked with the risk tolerance declaration (Hallahan, Faff & MacKenzie 2004). For instance, a study that measured the validity of self-report questionnaires for risk behaviours showed that these self-reports measures are influenced by cognitive and situational factors (Brener, Billy and Grandy 2003). Finally, questionnaires are not incentivized; hence it is difficult to reflect truthful underlying risk attitudes for financial decisions (Charness, Gneezy and Imas 2013) since there has been strong evidence that supports the association between risk elicitation and incentives (Charness, Gneezy and Imas 2013; Holt and Laury 2002).

Following advantages and challenges of questionnaires, this section reviews the other principal method of risk assessment involves choice problems, which are described as lotteries. In this category, there are two kinds of assessments: simple and complex. The judgment of them as being simple or complex derives from the participants' ability to understand them (Rohrmann 2005; Charness, Gneezy and Imas 2013). A relatively simple risk elicitation method, involving a lottery, was developed by Gneezy and Potters (1997). In this method, the participant gets and X amount and is requested to choose how much money the individual would invest in a risky option and how much would keep (Gneezy and Potters 1997). The choice between the amount of investment and amount of savings is the only choice the participant makes inside the experiment. The main benefit of this method is the simplicity of its use, which demands limited experimental tools and only one question, as well as the ability of the participants to understand it entirely. A disadvantage of this method is that it does not have a differentiating point between risk neutral and risk seeker preferences (Charness, Gneezy and Imas 2013). In addition to the lack of a differentiating point, it needs real monetary payoffs to have reliable results, which impose pragmatic limit on sample size. Another popular and easy elicitation method was developed by Eckel and Grossman (2002). It involves only one choice; participants are asked to choose between two gambling options. Each option has a 50% chance of the participant receiving a low payoff and a 50% chance of the participant receiving a high pay off. An advantage of this method is its simplicity. Dave et al. (2010) demonstrated that this method produced more consistent results regarding participants risk preferences compared to other sophisticated methods for participants with low math abilities. In spite of this, the main issue in this method is that it cannot distinguish between different degrees of risk-seeking behaviour whether someone is risk neutral or risk averse, which is necessary for accurate prediction of someone's behaviour (Rohrmann 2005).

A more complex method of eliciting risk preferences involved multiple choices between gambles. Past researchers have used this type of assessment to evaluate risk preferences in a variety of contexts such as gambling, cognitive ability, time preferences, drugs and alcohol (Anderson 2008; Dohmen et al. 2010; Holt and Laury 2002; Kahneman, Knetsch and Thaler. 1990; Coller and Williams, 1999). Therefore, this is one of the most used risk elicitation methods in economics (Charness, Gneezy and Imas 2013). Participants are asked to choose between one safe (Option A) and one risky lottery (Option B). An early multiple-choice lottery task was used by Binswanger (1981), who presented a multiple lottery choice test in farmers of rural India. However, the method became popular from Holt and Laury's studies (2002). Hence, most of the research communities are referring to this method as the Holt and Laury (HL). The Holt and Laury method involves ten pairs of gambles. These two pairs of gambles include one safe option and one riskier one over ten rows, which represent ten different lotteries (see Table 1). Participants need to choose for each row which option they prefer. The payoffs of the two options remain the same across the ten different lotteries the only thing that changes are the probability associated with each payoff (Table 1). When moving down the rows the probability of the high payoffs increases, therefore the switching point from one option to the other is the measurement of someone's risk preferences. For example, if participants understand the instructions well and they are risk averse, they will choose option A until the 6th row. The outcome of the risk elicitation task whether someone is a risk taker, risk neutral or risk averse relies on the individual's ability to evaluate numerical probabilities, which is subject to influence by the way the person perceives, processes and evaluates the magnitude of the uncertain event described as a percentage. For example, 65% chance of winning £20, could be a possible event for risk takers but some others highly unlikely (Peter 2008). Therefore, this method of assessment has been problematic for people with limited mathematical skills (Dave et al. 2010).

In a research study about risk taking, the biggest percentage of 881 Canadian residents who chose irrationally in the task, were proved to have low mathematical skills when they were assessed by a 31 real-life mathematical problems questionnaire (Cleveland, Harris and McGrill 1982). This irrationality of choices, which is clarified either by switching more than once or making "backwards" choices with one Option and then switching to Option B, confirms the complexity of this method and their low predictive value for risk-taking behaviour.

Nevertheless, a field experiment conducted by Dave et al. (2010) revealed that a complex lottery as Holt and Laury task has better predictive value than easier risk assessments but only for people with high numerical skills (Dave et al. 2010). However, for people with low mathematical skills, risk assessments like Holt and Laury lotteries could create an effect which looks like risk aversion (Dave et al. 2010). One of the main advantages of this method is the ability to distinguish someone's risk preferences level by clearly indicating in which level of risk taking the individual belongs to. Another major advantage is that it can determine participant's inconsistent answers. The way it is structured it can show whether someone did not understand the experimental procedure or choose inconsistently due to hurry. Therefore, these data can be excluded from the analysis providing better predictively value in the task. For the above-mentioned reasons considered by many researchers the 'gold standard' of risk elicitation tasks for people with good numerical skills (Anderson and Mellor 2008; Nielsen and Zeller 2013). According to Holt and Laury, risk aversion is classified based on lottery choices. In the table below the formula where the payoffs were calculated is presented along with the classifications.

Table 2. Risk aversion classifications in Holt and Laury task (2002)

The last category of risk assessment involves computerized methods. A popular method, which involves some game elements, is the Balloon Analogue Risk Task (BART). It measures risk preferences by pumping a balloon up by clicking a button until some threshold when the balloon explodes (Lejeuz et al. 2002). The threshold point is unknown to the participants. This tool was designed to model situations where the more risk taker you are, the less money you will earn and more hazards will be created (Lejeuz et al. 2002). With each pump, the money is increased, but the possibility of losing everything when the balloon is exploded is also getting bigger. In the validation study of the BART tool by Lejeuz et al. (2002) with 86 participants, riskiness on the BART was significantly correlated with the risk scores in other validated self-reported risk assessments (Lejeuz et al. 2002). However, this evidence would not have been sufficient, especially when considering the problems of self-report assessments, if there were no correlation between BART and real-world risk behaviours such as smoking status (Lejeuz et al. 2002). Since, BART was shown to have a significant correlation with real world risk taking behaviours and is a computerised task that shares some gamification elements, such as feedback, progress and visuals, it was used as a validated metric for the pilot study (Experiment 1) to measure the construct validity of the gamified risk elicitation task. However, as a risk elicitation task, BART demands experimentation in a laboratory setting, which is not convenient especially for field experiments, and it is a time-consuming task, which requires incentives to have valid results.

Another computerised task that was developed to assess real-world risk taking in a laboratory setting is Iowa Gambling Task (IGT) (Bechara et al. 1994). Participants are given $2000 to start and are asked to maximise the profit over $100 trials by selecting one card out of four decks. Each time, decks A and B have a profit of $100 on average and decks B and C have a profit of $50. However, A and B are considered as risky and dangerous whereas B and C as safer. The main advantage of this construct is that it can show emotion based decision making (Tumbull et al. 2005). However, it has been found that IGT is complex as a construct, therefore, it has been questioned whether it reveals true risk preferences or random choices out of misunderstanding (Reynaud and Coutoure 2012).

Given the shortcomings of complexity, lack of differentiation between a risk neutral and risk averse individual, limitation in implementing outside of laboratory settings and the need for incentivisation, the potential exists to improve the existing tools to maximise their predictive value. From the review on the most popular methods in each type of risk tool, Holt and Laury (2002) task has the strongest advantages in terms of predicting people's preferences in lottery type choices, determining people's inconsistent choices and differentiating risk neutral from risk averse people. Therefore, Holt and Laury task has the potential, if its complexity is improved to provide accurate predictions for individuals' decision making.

2.2.4 Risk elicitation tasks requirements and considerations

Empirical studies from past risk elicitation task have brought to light specific potentials as well as limitation of such assessment tools. Therefore, according to those findings that were described in depth in Section 2.2.3, this section presents the requirements and the limitations in summary to help the formulation of the hypothesis in Section 2.6.

The requirements of a risk elicitation task are the following:

- It needs to be simple enough for people to understand it (Eckel and Grossman 2002; Dave et al. 2010)
- It needs to include a differentiating point that will distinguish risk neutrals from risk averse (Holt and Laury 2002).
- It has to be easily implemented to use it outside of laboratory settings for wider audiences.

The limitations that need to be taken into consideration before developing a risk elicitation task are the following:

- It has to be used along with monetary incentives to have valid results (Holt and Laury 2002).
- It needs to determine people's irrational choices to exclude them from the analysis (Bauremeister and Musshoff 2016).

The requirements and limitations of a risk elicitation task lead to consideration of gamified approaches to overcome those challenges. Serious games were shown to have certain elements such as 1) engagement outside of laboratory setting and ease of distribution for digital games, 2) ability to visualise information in a novel and interactive game towards simplifying the task 3) ability to incentivize through game-based reward rather than monetary (Ritterfeld, Cody and Vorderer 2009) that can overcome the challenges of risk elicitation tasks. In the scope of this thesis, the potential of simplifying a risk elicitation task with the use of gamified elements has been investigated. The task developed by Holt and Laury includes all the main requirements of risk elicitation tasks. However, its complexity compromises the validity of the assessment outcomes as well as its predictive value, which have been addressed by several empirical studies (Jacobson and Petrie 2009; Charness and Viceisza 2013; Cleveland, Harris and Mcgrill 1983). Therefore, if initially this issue of Holt and Laury's task is addressed with the introduction of gamified elements, then the rest issues concerning incentives and ease of distribution outside the laboratory would be investigated and solved accordingly in a later stage. In Section 2.3, previous in-game assessments are reviewed to highlight the elements of games that can improve risk elicitation tasks.

2.2.5 Summary of risk taking

A variety of definitions of risk taking exist, however, all of them illustrate the two sides of risk, one that involves the uncertainty and the other that is related to the extent that matters to the individual. Since the part of the importance for the individual is a subjective matter, this thesis is focused solely on the uncertainty of risk definition. The concept and the definition of risk are not absolute; it depends on the context. In this thesis, the main focus is on risk-taking in financial context. The risk assessment tools that have been used in current research can be broadly divided into three categories: questionnaires, choice problems (simple or complex) and computerised methods. All of them have advantages and disadvantages. Questionnaires have been criticised because they are self-report and therefore biased by social distortions, but there are easy to implement. Lottery/Choice problems when they are complex they considered to have great predictive value and they can often distinguish a risk neutral from a risk-averse person, but they are considered complex and need incentives to produce truthful predictions. Simple choice problems cannot distinguish a risk neutral from a risk-averse individual, but they are simple to understand. Computerised methods have been linked with risk-associated behaviour such as alcohol abuse, but they are difficult to implement outside of laboratory settings. The most often-used risk elicitation task is the Multiple Price List by Holt and Laury (2002), which belongs to the category of complex choice problems. It involves ten pairs of lotteries, which differentiate the level of someone’s risk attitudes. It is considered a complex tool which produces noisy results as participants tend to choose irrationally in the task (Eckel and Grossman 2005; Charness and Viceisza 2002). Therefore, compromises the validity of the task outcomes. In summary, the insufficient evidence of the so existing risk-taking tools has left space for the further exploration of using other ways of eliciting people’s differences (Rohrmann 2005). Games were shown to have certain elements primarily in using visuals to express otherwise difficult to understand text that could improve an important flaw of passed risk elicitation tasks which is the complexity. By simplifying the experimental procedure of Holt and Laury task with the use of visuals in a gamified context, it opens a new path for also addressing the rest of the requirements. If the Holt and Laury task becomes easier, then the need for incentives can be replaced with rewards in a game and the limitation of task implementation outside of laboratory contexts can be distributed through the use of smartphone games. In the following section, a literature review on in-game assessments presents several examples of games used as assessment tools.

2.3 Gamified assessment tools

This section provides a high-level overview of the use of gamified approaches as assessment tools focused on their strengths relevant to the challenges faced by risk elicitation tasks as identified in Sections 2.2.3. In the context of this thesis, by term gamification is meant the application of specific game elements into a risk assessment task. The game elements

2.3.1 Overview of gamified assessment tools

To address the challenges outlined in Section 2.2.3 regarding risk elicitation tasks, here a high-level overview of games used as assessment tools is reviewed. The outcomes of this literature review will further formulate the hypothesis in Section 2.6.

When players interact with an environment during game-play, certain variables change, for example, finding a clue or an object may increase the player's inventory of goods. In the same way, someone could monitor several variables related to personality or behaviour of the player at different levels of granularity inside the games. For example, someone's creativity or teamwork or thinking skills could be monitored (Shute and Ke 2012). Traditionally, play-based assessment refers to investigating how people play to determine their cognitive development since games can keep track of people's decisions and choices (Michale and Chen, 2005). Games can become effective in assessing people's learning level or cognitive and behavioural skills as they take us away from conventional methods of assessment, which may include questionnaires that can be biased as a result of various distortions such as self-bias and socially accepted answers. According to what Shute and Ventura (2009) describes as a stealth assessment, the player may be unaware of the assessment tool inside the game. This way, the player will be completely concentrated on the game.

In this section, several examples of games used as assessment tools are reviewed, outlining their main strengths primarily in engaging users as well as their potential to simplify and present meaningful the procedures while it portrays their weaknesses related to the implementation of elements that sometimes distract the user.

An example of a game used as an assessment tool is Immune Attack, which was designed to assess students learning of the immune system. Inside the game, the player needs to perform different tasks such as training macrophages to identify enemies or if a blood vessel is infected and attacks from bacteria (Froschauer et al. 2010). A game assessing cultural training in Afghanistan is The Living World (Zielke et al. 2009). This game involves the navigation of cultural moves in the game space as well as the achieved attitude of the locals toward the player. The game assesses the choice of the player, and it provides a score according to that. Feedback is an integrated part of the game, as the player needs to understand why a response is most appropriate. This game element points out an important advantage of games, which is the sense of value that replaces the monetary incentives. In the majority of cases, it was shown that feedback and game rewards can replace the value of monetary incentives, which is valuable for the collection of large samples in experimental procedures (Easley and Ghosh 2013).

The games as mentioned earlier has incorporated measurements deriving from questionnaires but inside the gameplay assessing the individual in a less invasive way. As it was recently shown, games can become effective as an assessment tool of personality traits for a large sample of subjects (Brown et al. 2014). This could be usefully in many areas, for example in the recruiting process of businesses. A serious game like that is where the player acts in a story, makes managerial decisions, some of them even immoral (but very profitable for the company), which impact the virtual company. The measurements are: seek and control of other, someone's propensity to distrust others, engagement in amoral manipulation and a few others. All these measurements exist already in reliable self-assessed tests (Dahling, Whitaker and Levy, 2009). Considering that there have been identified a few disadvantages with the self-assessment questionnaires (van Lankveld et al. 2011), for example, candidates tend to give answers that are socially desirable. According to van Lankveld et al. (2011), a game can reveal a personality trait unbiased from the. This is not achieved by asking the participant to answer questions but by adopting the Five Factor Model of Personality questionnaire within the game play and store automatically player's behaviour. Results from another study with a game that revealed a significant correlation between play style and personality were conducted with a sample of 13,376 individuals (Wooters et al. 2013). In this study 175 game variables were correlated with 100 personalities scores and five personality dimensions with significant results (Wooters et al. 2013). For example, the game variable unlocks score per second was correlated significantly with conscientiousness and extraversion ($r = 0.42$). In this game, it was revealed that the use of visuals simplifies procedures and problems that otherwise would be difficult to be represented inside a questionnaire. Those games that assess personality traits were shown to hold the advantage of using game-play and visuals to ‘put’ the player in a similar position to reveal his true preferences/tendencies.

SimVenture is another type of business game where the player needs to manage a company with four major issues: production, organisation, sales, market and finance (King and Newman 2009; Stainton, Johnson and Borodzicz 2010). Therefore, the player is being assessed against one’s decisions on these matters, and at the end of it, the major issues are highlighted. PIXELearning’s Enterprise Game is a similar game where the player is being assessed for his performance against the competitors in a marketing strategy project (Belotti et al. 2013). These games are useful for recruitment purposes or even in company evaluations. The Great Brain Experiment involves some short games used as assessments, for impulsivity, inhibition, risk taking, memory and happiness. Even though not all results have been published yet, there are some promising outcomes in the inhibition research, where it was found that reactive control deteriorates by age and this occurs faster in men that women (Smittenaar et al. 2015). This was examined through a large-scale sample of male and female participants who ranged in education, age and depression symptoms (Smittenaar et al. 2015). Considering the limitations of this study that it was conducted through the use of smartphones, therefore, it would most probably use by Western societies and the cross-sectional approach which may have affected the results. A huge opportunity arises; the potential of engaging the public in large scales in participating in research (Smittenaar et al. 2015; Teki, Kumar and Griffiths 2016). This assessment tool also highlighted the potential of using games and distributing them through the smartphone to collect data outside of laboratory settings and reach groups of people that otherwise would be difficult to attract in the lab. Similarly, one of the Great Brain’s Experiments games was assessing the ability of the player to remember certain items while ignoring distracted others. The findings showed that older participants had greater difficulty into ignoring distractions compared to younger ones (McNab et al. 2015). The results that were published showed that it is possible to reproduce laboratory findings of several personality traits or behaviours despite not being conducted in the controlled environment of a laboratory (McNab et al. 2015). Even though the above examples of games described the benefits of using in-game assessments, several shortcomings need to be considered and are being described in the next section.

There are also some shortcomings of using games as assessment tools that need to be taken into account. One of the main disadvantages is that the in-game elements may distract the person from the performance of the task; hence sometimes the participants try to ‘beat’ the game by circumventing (Schuller et al. 2013). Moreover, the in-game elements that serve purposes of creating a fun context can sometimes distract players from the actual task, and as a result, their performance is not the realistic. Nonetheless, some games could complicate a situation by adding additional elements, the integration of visual aids in games has the potential of simplifying difficult concepts such as mathematical problems (Annetta 2010). An in-game assessment to be effective needs to take into account the target audience, the context and limitations of potentially distractive game mechanisms. For instance, the use of distracting graphics that are not related to the main objective of the assessment can distract players instead of focusing on the task itself. According to Belloti et al. (2013), successfully integration of the assessment in-game involves the proper mechanisms and the conditions to activate them. Hence a pen and paper task can be transferred to a gamified context that will be less complex without changing its initial purpose. For this thesis, the Holt and Laury task is transferred from the conventional questionnaire style of lotteries into a gamified approach where the lotteries were presented as a type of roulette. The mains structure of the Holt and Laury task were not changed apart from the game mechanic of roulette and a housing element, which was changing from a smaller house to a bigger according to losses or gains.

To sum up, empirical studies with games used as assessment tools have been shown to have certain benefits compared to conventional assessment experimental procedures. These are the following:

- Game elements such as feedback or rewards schemes were shown to create the sense of value to the player and replace the need for monetary incentives to have accurate results in assessment tools (Early and Ghosh 2013; Deterding 2014).
- The game may incorporate measurements from questionnaires in a less invasive way (van Lankveld et al. 2011).
- Use of visuals in games, was suggested that can simplify procedures that otherwise would be difficult to describe through pen and paper e.g. present a hypothetical lottery question compared to playing the lottery (Wouters et al. 2013).
- Games were shown to provide the potential of distributing the assessment in large and difficult to find groups of people through smartphones (Brown et al. 2015).

Despite the benefits of in-games assessment tools, there are also several shortcomings that need to be taken into consideration. These are summarised as follows:

- In game, elements may potentially distract the player and complicate an experimental procedure (Annetta 2010).

- Sometimes, participants may try to 'beat' the game by circumventing and as a result they do not perform naturally (Schuller et al. 2013). If the user tries to circumvent, then the assessed behaviour is not accurate.

Any in-game assessment tool needs to take into account the specific assessment requirements, the target group and the limitations. As outlined in Section 2.2.3, there are certain requirements a risk elicitation task need to have and several issues that need improvement. In light of the empirical findings showing in-game assessments to be effective assessment tools and based on the benefits outlined in the previous section, this thesis investigates whether the introduction of gamified elements would improve the complexity of complex Holt and Laury task and influence participant to choose rationally. Therefore, in the next section, the risk elicitation task's components are being deconstructed towards understanding in depth the limitations of the specific type of risk elicitation tasks.

2.4 Irrational choices in risk elicitation tasks

Considering that several studies have shown people choosing irrationally and therefore inconsistently in risk elicitation tasks such as Holt and Laury's Multiple Price List task, this section of the literature deconstructs the main elements of this task linked to the outcomes to investigate how can be more clearly understood. Therefore, this section focuses on the presentation of lotteries in risk elicitation tasks and the factors that influence people to choose rationally. It also describes the issues of probabilities which are the principal way of presenting the lotteries in the Holt and Laury task (2002). This review continues by exploring the visual thinking approach when reasoning to have a deeper insight into the underlying thinking process of people, towards improving risk assessment tools.

2.4.1 Factors that influence people's rational evaluation of choices in lottery tasks

In lottery/choice problem risk elicitation tasks, risk preferences rely on the individual's ability to judge numerical probabilities, which are subject to influence by the way the person perceives, processes and evaluates the magnitude of the uncertain event described as a percentage (i.e. 65% chance of winning £20, could be a possible event for risk takers but for some others highly unlikely) (Peter 2008). So far, one of the principal methods of risk assessment involves choice problems, which are described as lotteries. Participants are asked to choose between one safe (Option A) and one risky lottery (Option B). Both options are presented to the participant as numerical probabilities (Table 1). Research has shown that people found the options in the task difficult because they are not able to comprehend probabilities (Dave et al. 2010). According to Dave et al. (2010), people with low mathematical skills produce a high level of irrational choices and inconsistent with the pay offs in such lotteries which makes it difficult to assess their risk taking and also compromises the predictively value of the task. This produces significant noise among the data, which may reduce the experimental validity of the measure to predict participant's actual choices (Dave et al. 2010). Jacobson and Petrie (2009) discovered that 55% of a sample of 181 Rwandan adults with low educational level made inconsistent choice switching over gains and losses in a complex lottery method. Charness and Viceisza (2012) found out that 75% out of 91 farmers with a low educational level in rural Senegal chose inconsistently (51% switched preferences more than once, and 24% always chose Option A from the two lotteries, Table 1) again in a lottery task. However, from another study with Canadian residents, it was shown that the majority of the sample who chose inconsistently, it was due to low mathematical understanding as it was shown from a 31-real life mathematical questionnaire (Cleveland, Harris and McGrill 1982).

In any event, mathematical abilities or educational level are not the only factors that influence the evaluation of probabilities. Frederick (2005) found that people who are more reflective than others can better deal with probabilities and they make choices consistent with expected value in lottery tasks (Frederick 2005). Therefore, reflective thinkers who also have certain mathematical abilities, specifically a better understanding of how to transform percentages, are not so affected the way the probabilities are presented (Peters and Levin 2008). For example, they can translate the 67% correctly into 33% incorrect, and they are also able to easily convert the percentages into frequencies (e.g., 33% is the same thing as 297 people out of 900). As Kelley (2009) has addressed it, a person who is skilled in numeracy and is a reflective thinker, tends to make "better" choices because the individual can make more accurate and thoughtful evaluations of gains and losses derived from the probabilities (Kelley 2009). Empirical evidence also suggests that even though some can cooperate fairly well with additional ambiguity others, especially people with low numeracy skills and low confidence become frustrated and more risk averse. As it has been argued by Peters (2008) individuals who have low mathematical ability or have low levels of educational attainment often have difficulty in understanding probability values, thus they do not "reveal" their actual risk preferences (Peters 2008). Therefore, the factors which lead to accurate evaluation of probabilities need to be examined, as it was shown that less numerate people do not have all the information that they need, for example only numbers without explanation, context or visuals. Lacking the depth available to the most numerate, people with limited numerical skills tend to get influenced by irrelevant affective sources such as emotions or other's people experiences with the result to of being more risk taker/averse than they are in reality (Peters 2008).

People should be able to comprehend the information and its meaning so they can interpret the information and use it as the basis for making decisions (Peters 2008). In a study conducted by Peters (2008), it was shown that less numerate people misjudged the significance of an uncertain event when it was expressed in percentages compared to that expressed in the frequency format, for example of every 100 patients similar to Mr Jones, 10 are estimated to commit an act of violence" versus "10% of patients similar to Mr Jones commit an act of violence. Another study conducted by Black, Nease and Tosteson (1995) identified that 38% out of 62% college students made serious mistakes in estimating numerically the likelihood of contracting breast cancer. The last two studies indicate that wrong evaluation of probabilities may not necessary come from low educational attainment but low numerical skills.

On the grounds of the above evidence, education, mathematical understanding and cognitive thinking style could play a role in the accurate evaluation of probabilities. However, this is only one side of the problem, which concerns the individual's abilities, a question remains whether the presentational form of probabilities in those tasks could influence rational decision making. Therefore, the challenge of probabilities' presentation is further reviewed in the following section, before this thesis explores how this issue can be improved in risk elicitation tasks to assist people with low numerical and cognitive thinking style difficulties to decide rationally.

2.4.2 The challenge of presentation of probabilities and the use of visuals

As mentioned in the previous section, other than mathematical abilities, educational level and cognitive style, there are also representational issues with probabilities (Kahneman and Tversky 1972). One representational issue is the concept of subjective probabilities, referred to as people's tendency to deviate in some way from objective probabilities, and the other is the perception of them, which might not correspond to the actual probabilities (Kahneman and Tversky 1972). In line to this, people tend to misjudge percentages as smaller or higher because of the way they are presented than they are which affects their choices. It has been argued that the effects of the probability's presentation depend not only on the "type" of format but also in the context where it is used (Visschers et al. 2009). Hence, people get confused by the contextual phrasing of probabilities, for instance when one is referred only to the negative percentage i.e. 40% death rate, or the format type (i.e. 40% or 40 out of 100). Visschers et al.

(2009) support, in their review of empirical findings on communication about probability information, that the format of probabilities influences more individuals who use heuristics to make their decisions. Examples of these include rules of thumbs, stereotyping, common sense and intuitive judgment rather than those who process information systematically. These heuristics could create systematic bias and disastrous judgments, Evans (2007) in his study with 180 students, observed that people's thinking process while solving problems with probabilities was related to the format in which the data was presented. It was found that students, who solved the probability problems described in frequencies successfully used logical reasoning while those who did very well with the percentage probability problem used probabilistic reasoning which involves a combination of probability theory and deductive logic (Huerta et al. 2011). To that end, there is a clear relationship between mathematical skills and presentation of probabilities. However, there is a solution to the misinterpretation of risk information and the lack of mathematical training; "every piece of statistical information needs a representation that is a form" (Gigerenzer and Edwards 2003). Representation of probabilities is defined as either images or objects or anything that can symbolise or reflect the probability of an event (Confrey and Smith, 1991).

Galesic and Garcia-Retamero (2010) developed a graph literacy scale to predict who can benefit from visual presentations. They showed that people with low numeracy skills understand better the statistical information about health issues with visual displays; whereas, people who lack graph literacy could benefit from the numerical presentations rather than then graphs (Gaissmaier et al. 2012). Garcia- Retamero and Dhami (2011), in their study about the comprehension of risk information using icon arrays in non-native speakers, showed that visual presentations are a helpful method to reduce inaccurate perceptions of risk issue (Garcia-Retamero and Dhami 2011). Despite the benefits of visual displays, there are potential limitations, for instance, people might focus more on the pattern of data rather than the precise values. Hence some graphs are better suited for certain tasks, for example, line graphs are better for trends over time (Feldman-Stewart et al. 2000). Findings have shown that risks regarding health are better perceived through vertical bars, horizontal bars and ovals than pie charts which lead to less accurate estimates of probabilities (Feldman and Stewart et al. 2000). There is evidence that visual formats can assist in reducing several biases that mathematical presentations can cause, such as framing effects (Garcia-Retamero and Cokely 2011), the influence of other's people experience (anecdotal information) (Fagerlin, Zikmund-Fisher and Ubel 2005) and the denominator neglect (Garcia-Retamero and Cokely 2011) and also help in understanding more complex concepts like incremental risks especially in finance (Zikmund-

Fischer et al. 2008). People usually pay more attention to the number of times a target event has happened (“numerators”) than the overall opportunities for it to happen (“denominators”). For example, people tend to rate as higher the likelihood of cancer killing 1,286 people out of 10,000 compared to 24, 12 out of 100 people but this tendency is minimized when it is described visually (Garcia-Retamero and Dhami 2011).

Overall, visual graphics could be used as a powerful tool for conveying statistical information especially for people with significant graphical literacy along with those who have difficulty in interpreting numbers, but they should be used with some caution and be pilot tested for understanding and avoid misleading images for example graphs with misleading scales). The findings from research draw the path for risk communicators to investigate deeper on the value of visual representation or even more to consider a combination of numerical presentation and graphs. The studies reviewed in this section, highlight that people with different educational background comprehend differently the various formats of probabilities presentations. This conclusion is significant for risk assessments as it leads the way towards considering a different methodological approach for people that have difficulty in understanding the numerical presentation of probabilities.

2.4.3 Visual thinking and use of external representations

To address the challenges outlined in the previous section regarding the evaluation of lotteries presented as probabilities in risk elicitation tasks, there is a need to investigate another methodology of reasoning them. Since Section 2.3 has addressed the potential of games to use visuals to describe a situation or a problem, hereby in the following sub-sections, it is investigated the benefits of using external representations to identify the most appropriate format of visuals for people with low numerical skills. In this section, the benefits of external visual representation as a form of better understanding of probabilities are presented.

It has been supported that visual thinking plays an important role in problem-solving in mathematical problems (Schreimber 2004). Visual thinking, in psychology, is considered the thinking that processes images, stimulus, forms or patterns to come to reach a conclusion. Although the imaginary mental process is widely believed that plays a role in mathematical reasoning, there is little evidence to support this argument. This may be because it is difficult to explore mental imaginary in rigorous ways. On the contrary, there are quite a lot of research regarding the external visual representations people use, as it has been argued that they can facilitate and improve the problem-solving procedure of mathematical and logical cases

(Tversky 2001; Stylianou and Silver 2004). External representations are defined as the 'physical symbols or dimensions (e.g. dimensions of a graph) and as the external relations embedded in the physical configurations (e.g. relation of written digits, graphs, diagrammes, etc)' that could be used from an individual to assist in the problem-solving procedure (Zhang and Wang 2009). For instance, when people describe a problem using a pie chart to represent the proportions mentioned there or they use mathematical calculations to find the solution, the pie chart and the mathematical formulas are considered external representations (Zahmer and Corter 2002). It has long been supported that external representations play an important role in problem solving with probabilities especially when such problems are complicated (Russell; Zahmer and Corter 2002). In an experiment conducted by Zahmer and Corter (2002) was found that people who used external representation had a higher rate of success in the problem-solving process compared to those who did not use.

When investigating external representations, there should be a division into two categories. One category of external representations refers to the success of visual presentations that are provided by an external person to help the decision process. These external representations involve various diagrams that are combined with text describing systems (Mayer and Gallini 1990) or actual physical models of scientific systems (Penner et al. 1997). Most of the studies regarding such external visualisations concluded that whenever the experimenter provided with visuals, the problem-solving procedure was successful. The other category involves the external representations an individual uses to reason a probability problem. These are also helpful to enhance the understanding of the probabilities.

A classification scheme of the external representations that problem solvers use, is proposed by Corter and Zahmer (2002). This scheme involves "…schematic diagrams, pictures (iconic) representations and certain forms of spatial organisation and tabulation of problem information". There is an important distinction between schematic and pictorial (iconic) visual representation of probabilities. The schematic representations represent the relationships that are described in the problem while the pictorial reflect the physical appearance of the elements included in the problem. When there are complex mathematical problems involving probabilities, there are some calculations involving numbers. It has been argued that mathematical inscriptions also have visual and symbolic content. In probabilities, there are certain schematic concepts, which are usually used to represent concepts to solve problems, for example, Venn diagrams, outcome trees (Corter and Zahmer 2007, Zahnmer and Corter 2002).

A different methodological approach was introduced towards reasoning probabilities. The methodological approach of people using external visual representation to reason probabilities, it could provide answers related to people's underlying thinking process when they fill in the choices of a risk elicitation task. Therefore, instead of investigating the most appropriate format of presenting the lotteries, investigating people's external representations while choosing in a risk elicitation task can indicate a different methodological approach that could help people to choose rationally.

2.4.4 Previous research on external representations

Empirical evidence suggests that using external representation to reason probabilities help individuals to successfully solve problems (Corter and Zahmer 2002). Towards this direction the following examples highlight the benefits of this methodological approach. The majority of previous research regarding visual representation has been conducted with students, as it is the main group that concerns the area of understanding probabilities. All of the studies suggest that using visual representation along with the problem can help in the problem-solving procedure (Hall, Bailey and Tillman, 1997; Kaufmann, 1990; Mayer, 1989; Mayer and Gallini, 1990; Molitor, Ballstaedt and Mandl, 1989; Santos-Trigo, 1996; Tversky, 2001). The examples of visual representation used in these studies varied from diagrams that were presented along with text to scientific systems. Despite the variety of external representations, common finding to all of these studies was the enhancement of problem-solving procedure and the development of student's understanding towards the problem.

Notably, people do not always get the benefit of using external visual representation while solving problems. Research has shown that individual's differences can group problem solvers into different categories according to their tendency to use visual representations or not. One group involves the people who use verbal representation, and the other group involves people who use primarily spatial/visual representations (Hegarty and Kozhnenikov 1999). According to Hegarty and Kozhnenikov (1999), the visual group can be further split into two categories into an object, and spatial visualizers, with the spatial visualisers, have greater success rates in mathematical and scientific tasks.

Previous work regarding probability problem solving showed that people chose to use external visual representation while solving problems after they had been taught certain visuals in an introductory statistic course (Zahner and Corter 2002; Russel 2000). The types of external visualisations include at least: graphs, tree diagrams, contingency tables, Venn diagrams and pictures. Additionally, formulas and mathematical symbols could also be added to the list as they involve visual and spatial relationships (Presmeg 1986).

According to research conducted by Zahner and Corter (2002) about external representations, it was identified that visual external representations are mostly used with unfamiliar or difficult problems. This evidence supports the theory of Tversky (2001) that there are some potential factors which explain the usefulness of visual representations to aid a problem's solution such as supporting the memory, attracting the attention, recording information, facilitating inference and discovery. Another possible explanation is that by using various representation of the same problem, it leads to a fuller understanding and depth in processing (Mayer, 1989, 2001; Mayer & Gallini, 1990). Schwartz and Black (1996) argued that external representations help individuals to create a mental model of the described situation. This evidence provides support that when people use their external representations and interact with the problem visually, they have a higher likelihood of solving it.

2.4.5 Categories of visual representation

The visual representations can be split up into two basic categories; external and internal visual representation. The internal visual representations involve any mathematical ideas or pictures or schemata developed by the individual through his experience whereas mathematical manifestations, images or pictures, which help the individual to understand concepts that involve mathematical problems, are the external representations. According to previous research the three basic types of external visualisations are non-diagrammatic, schematic diagrams and iconic (Goldin and Shteingold, 2001; Ho and Lowrie, 2014; Pape and Tchoshanov, 2001).

Table 3. Types of external visual representation (Goldin and Shteingold, 2001; Ho and Lowrie, 2014; Pape and Tchoshanov, 2001).

The first category, the non-diagrammatic includes reorganisation and outcome listings. The second category includes schematic diagrams such as Venn diagrams, contingency tables and the third category involves pictures. The schematic category is when the individual uses diagrams to represent the problem to understand it better before calculating the result. Individuals most often use the last category, the picture as the iconic representation of the problem to understand it before start interpreting it (Zahner and Corter 2010). According to Hegarty and Kozhevnikov (1999), the use of schematic representations leads to a higher rate of success in problem solving compared to iconic representation.

2.4.6 Summary of factors that influence people's evaluation of the lottery option and visual thinking

This part of the literature presented factors that influence people choosing inconsistently in risk elicitation tasks like Holt and Laury. According to empirical evidence on the factors that influence people's rational choice in risk lottery elicitation tasks, numeracy levels, cognitive thinking style and education are among the most profound factors (Dave et al. 2010; Cleveland, Harris and McGrill 2006; Charness and Visceiza 2013; Frederick 2005; Peters and Levin 2008). In addition to numerical skills, educational level and cognitive thinking style there are also representational issues with probabilities. The way probabilities are presented to individuals, affects their evaluation especially for those who use rules of thumb make decisions. Evidence has shown that people's thinking process while solving problems with probabilities is related to the format that the probability is presented (Huerta 2007). However, the misinterpretation of probabilities could be solved by using visual representational formats that could be understood by people with low numerical skills. From a study investigating the effectiveness of a graph literacy scale, it was found that less numerate people successfully understood the probability information when they were presented in the format of a graph (Galesic and Garcia-Retamero 2010). Nonetheless, though the use of visuals sometimes can also distract and if not the appropriate graph is used may mislead the individual (Fieldman-Stewart et al. 2000). The contradicting findings whether visuals help or cause problems such as distraction from the important information to people with low numerical skills in understanding probabilities, lead to the need for a different methodology approach in investigating appropriate probability presentation for this target group.

Research on visual thinking has shown to play an important role in problem solving (Schreimber 2004). Visual thinking is divided between mental imagery and external representations. In the scope of this thesis, only the aspect of external visual representations is reviewed because this thesis adopts the position of improving the technique, rather than investigating comprehensively the underlying thought/cognitive process. Therefore, it takes this process as a constant, acknowledging variance through individual difference in its empirical design and large-cohort nature. The external representations people use to solve a probability problem were shown to enhance people's understanding. Between the different external representations that an individual can use such as graphs, pictures, words or mathematical calculations, graphs were mostly used with the unfamiliar and complex problems (Zahner and Corter 2002). Evidence shows that there are some potential factors which explain the effectiveness of visual using as an external inscription to solve a problem, such as supporting the memory and attracting the attention and recording information (Tversky 2001). The categories of visual representation that exist are non-diagrammatic, schematic diagrams and iconic. Hegarty and Kozhenikov (1999) showed that people who use schematic diagrams, which involve the graphs, are more likely to answer correctly in a probability problem compared to those that use pictures.

In this section, the factors that influence people to choose inconsistently in lottery choice risk elicitation tasks were outlined. A major part of them is considered to be the presentational format of probabilities. Using external representations to reason a probability has shown to have a significant effect in previous studies. Therefore, a different methodology approach of using external representation while choosing the options in a risk lottery elicitation task can be reviewed to investigate whether it can influence participants' rational choices.

2.4.7 Visualisation of numerical lotteries

The previous section reviewed the potential factors that influence people in choosing rational in complex choice problems which include, in the scope of this thesis, an individual's numerical level, education level, cognitive thinking style as well as the presentational format of probabilities in lotteries. Empirical evidence has shown the appropriate way of presenting the lotteries can help people to choose rationally in lottery choice risk elicitation tasks especially for people with numerical skills. Research on risk assessments, which involve lotteries, has experimented with various probability presentation formats to achieve rational and consistent choices and a better understanding of the probabilities amongst people (Johnson et al. 2007; Bauremeister 2016; Ihli et al. 2013). Literature suggests that the accurate assessment of quantitative values from visuals referents need the individual differentiation for experts and non-experts (Cleveland, Harris and McGrill 1982; 1983). Therefore, in the following sections, it is reviewed the communication of probabilities with graphs following up with an overview of the most popular graphs used for presentation of probabilities. Finally, previous graphical and visual representation of lotteries in Holt and Laury are presented to provide a full context of the problem deriving from the inappropriate use of graphs, which is linked to people's inconsistent choices. Drawing on findings from previous studies with lotteries, the following chapter section reviews the presentation of probabilities with graphs to explore different visual presentation of lotteries in risk elicitation tasks.

2.5 Communication of probabilities with graphs

Communication of probabilities in any area from health to finance has been categorised as primarily aiming to achieve one or more of three goals: 1) to increase understanding, 2) to change the behaviour or to 3) resolve conflicts (Stone et al. 2003). This research, as has been mentioned before, targets improvement of understanding, because it is the main issue regarding the Holt and Laury task. Empirical evidence has shown that people have difficulty in understanding Holt and Laury task, hence they tend to do irrational choices (Dave et al. 2010). Researchers who undertake the same goal, use outcome measures e.g. accuracy or consistency of reasoning using graphs to determine that the difference in presentational form can help people understand the probability better. The different research methodologies include asking from people to:

a) Estimate a proportion represented in a graph,

b) Interpret a part of a graph, as well as to choose between the same risky and safe options in visual and in textual form.

In the context of the 2nd methodology, regarding interpretation from graphs, it has been argued that people's perception of risk is strongly influenced by the design of the graph, hence when people's choices deviate a lot when are presented with graphs compared to the textual from, then any graphical elements inducing that variation (including framing or axis distortion) are considered unethical (Stone et al. 2003). For example, when people are presented with graphs that the numerator (e.g. no of people who have developed the disease of interest) of risk ratio is emphasised, they tend to change their behaviour positively (Nuovo, Melnikow and Chang 2002). The audience seems to become more risk averse when risk information between the relationships of part-to-whole which something is determined of are not presented in the graphs because the numerator draws attention. Therefore, graphs must present the part-to-whole as well as the relative risk along with the absolute, to provide the information completed. If this is not represented in the graph, then the risk tendency seems to change even for the most educated people (Ancker et al. 2006). If in a graph, the low end of a risk scale is highlighted to emphasise low probability, then the overall perception of risk likelihood decreases. In other words, the low probability draws too much attention, creating false risk perceptions (Woloshin et al. 1999). For instance, if there is a chance of 5% flooding if it is raining continuously over a period of a month and it is emphasised on a graph, it is possible that it will cause more anxiety to people than intended. This is a very common problem in environmental risk information. Therefore, a way of presenting unfamiliar concepts is to point out on the graph when action is needed. Furthermore, Shapiro et al. (1977) qualitative study found that women from a low educational background perceived a risk of breast cancer as bigger when it was shown as a part-to-whole human icon display rather than when it was shown as a part-to-whole bar graph. Another study compared the magnifier with the standard horizontal risk scale (Gurmankin et al. 2005). The results showed that the magnifier scale enabled appropriately low estimates for very rare events but also revealed that it significantly lowered risk estimates for more common events (Gurmankin et al. 2005). This effect was noticed when the participants estimated the risk of different health events without having seen the numbers about the magnitudes of those risks (Gurmankin et al. 2005).

To continue the investigation towards the presentation of probabilities in Holt and Laury task, the following section reviews the most popular graphs used to represent probabilities. This helps in understanding which graph is the most appropriate for presenting the lotteries in Holt and Laury task to prompt individuals in using it while choosing from the options in the task. Following that, the review continues by presenting visual and graphical representation that have been used in empirical research with Holt and Laury task.

2.5.1 Review of graphs

According to Hegarty and Kozhevnikov (1999), using graphs as external representations can assist in probability problem solving. However, there are a variety of graphs that can assist to different kind of probability problems. This section aims to understand the strengths and weakness of each graph to portray the most appropriate for assisting a risk lottery elicitation task.

When the appropriate graph is used to present information, it can summarise information concisely, portray hidden patterns and grasp attention to specific events (Lipkus et al. 1999). Some researchers perceive that graphs are an attractive alternative to numbers as they are visually interesting and can trigger automatic visual perceptions (Cleveland and Mcgrill 1985).

Effective graphics could be adapted to the audience's capabilities, and this can be helpful especially for people with low numeracy (Lipkus and Hollands 1999). For example, people that are less numerate overestimate the personal risk of dying of breast cancer and are less able to identify the magnitude of risk reduction due to mammographs. In that case, the graph could emphasise the risk reduction due to mammographs and change the misperception of less numerate people (Black et al. 1995). Below some of the most prevailing graphs are presented to overview their strengths and weaknesses:

Pie charts are considered a good visual representation of single proportion and are being familiar to the general public (Nelson, Hesse and Croyle 2009). They also provide a comparison between proportions. When Fieldman-Stewart et al. (2000) measured the speed and accuracy of patients and students' judgments in six formats: vertical part-to-whole bar chart, horizontal part-to-whole bar chart, pair of numbers, part-to-whole icon graphic with random arrangement, icon graphic with the icons arranged in a block, and pie chart (Fieldman-Steawart et al. 2000). The results showed that participants were slower least accurate in evaluating the larger of the two quantities with pie charts rather than the other forms. In another study, which compared graphs for conveying risks to physicians, five formats were given to them (Elting et al. 1999). The physicians were asked to see raw data from a fictitious clinical trial and to determine whether the data warranted halting the trial. Among tables of success rates, table of failure rates, bar charts, stacked bar charts and icon arrays, fewer participants noticed the high failure with pie charts and stacked bar graphs (Elting et al. 1999). To sum up, even though pie charts are well known by the general public and appropriate for single proportion representation, they can be misleading for presentation of more than two portions and not easily comprehensible e.g. to compare the size of two quantities.

Another popular graph is the bar chart. Bar charts can be useful to describe the magnitude and present comparisons as well because they display the data in separate columns (Lipkus 2007). Therefore, a bar chart should be used to compare discrete categories e.g. different patients. One of the major advantages is that it shows a scale of categories. However, when reordering the bars can cause misinterpretation. To appreciate how a bar chart can use information that misleads requires understanding some basic design features of bar charts:

1) Part to whole relationships: People's innate perceptual skills help them estimate the proportion object A represents compared to a larger object B. For instance, in a stacked bar where the bar that extends from 0% to 100% involves different colour segments to indicate the proportion differences. Usually, the proportions depicted on the segments are proportional to the quantities. However, there are cases such as a risk scale from 1 to 1

million which proportion segments are not feasible. Therefore, in such cases, the size difference between the various features is not associated with the difference in quantities.

2) Features that trigger basic graphical perception abilities: According to Cleveland and McGrill (1985) who ranked people's visual perception by their accuracy, found that people's accuracy was excellent when people are judging the height or the slope of a bar graph and less good when judging the colour and gray-scale densities (Cleveland and McGrill 1985).
3) Numerical format: The ability to understand risk rations or converting percentages to natural frequencies depends a lot on someone's education, training or numeracy skills (Schwartz et al. 1997). In a study conducted with outpatient clinics, it was found that only 56% of the participants could distinguish the difference between two risks when they were presented in a "1-in-x" format. Hence, the numbers on a bar chart could be clear and include the absolute number of the fact they represent.

According to the basic design features of a bar graph, still there are cases where some of the core information is missing. For example, as seen on the graph (Figure 2), there is no information about the absolute risk.

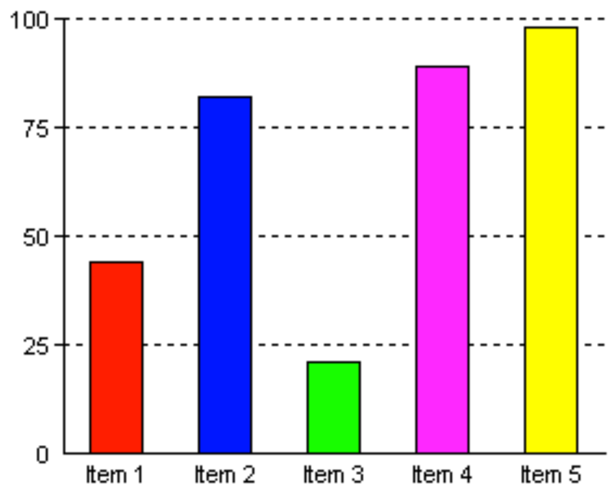


Figure 2. Bar chart

Another way of misleading information is when the bar chart shows negative framing. As seen on the bar chart Figure 3, bars display negative framing in the risks of cancer rather than probabilities of not getting cancer (Spiegelhalter, Pearson and Short 2011). Negative framing is when the information is presented emphasising the losses rather than the gains. Research has shown that when people are presented with negative framing they become more risk takers than when presented with positive framing (Kahneman and Tversky 1981).

Some materials have been removed due to 3rd party copyright. The unabridged version can be viewed in Lancester Library - Coventry University.

Figure 3. Bar chart with negative framing effect showing the risk of getting prostate cancer, the right column is based on lifestyle information and the Harvard school of Public's Health Desease risk website. The left bar provides a qualitative scale. (source: http://www.diseaseriskindex.harvard.edu/update/)

Another type of graph is icon array, which presents the risk at a discrete level of measurement in a group of individual icons. The risk can be described as dots or as figures as seen on Figure 4.

Some materials have been removed due to 3rd party copyright. The unabridged version can be viewed in Lancester Library - Coventry University.

Figure 4. Icon array of medicines treating urgency incontinence (source: https://www.ahrq.gov/)

According to Gigerenzer and Hoffrage (1995), people tend to do better in probability problems when the information is presented at the discrete level as in icon arrays compared to percentages or proportions. In a focus group with women, participants had a preference for icon arrays with smaller denominators because they perceived them as simpler and also highlighted that larger denominators tend to portray risk as smaller (Schapira, Nattinger and McHorney 2001). In another study with a sample of women with low income, participants preferred to see an individual risk estimate presented on a bar chart rather than as an array icon or a percentage, or a bar chart showing relative risks for women in different risk categories (Royak-Schaler et al. 2004). Schapira et al. (1997), in their qualitative study, found that women considered human figures more meaningful compared to face displays and easier to understand compared to bar charts (Shapira et al. 1997). In another study by Royak-Schaler et al. (2004), it was found that African-American women (of low income) preferred part to whole bar chart compared to part to whole icon arrays. Nevertheless, in this study, the bar chart had labels such as ‘high risk’, ‘low risk’ or ‘average risk’ whereas the icon array did not have any labesls. The latter potentially created some difficulty for viewers to identify the risk in the icon array (Royak-Schaler et al. 2004). Some other studies suggest that arrays of icon provide a good presentation of probabilities especially in the case of disease risk, better than a bar chart (Zikmund-Fisher, Fagerlin and Ubel 2008). In a study with 1619 women who completed an Internet-based survey about adjuvant therapy decisions for a patient with an oestrogen receptor-positive tumour, participants were randomized to see one out four risk graphics. The outcome measures were comprehension of key statistics, time which required them to complete the task and also graph perception ratings. The people who were presented icon array were more accurate in reporting the risk reduction achievable, answered the question quicker and like the graph more compared to the people who were shown the bar graph (Zikmund-Fisher, Fagerlin and Ubel 2008). Icon-arrays have the disadvantage of not providing information about the magnitude and are also difficult to compare especially for people with low numerical skills (Ancker, Weber and Kukafka 2011). Line graphs can be used to show a trend over a period of time. Line graph is plotted as a series of points, which then joined in straight lines (Figure 5). According to Lipkus et al. (1999), icon arrays are considered appropriate for presenting trends in data, as seen in Figure 5.

Figure 5. Line graph show changes over time. (Source: Vanderbilt University)

In a study by Mazur and Hickman (1993), who investigated how physicians, medical students and patients' treatments choices for an unidentified medical condition would be influenced by three 5-year survival curves which were different in terms of the area under the curve, found that patients' choices were influenced by the location of the intermediate year (year 0) as well as the end point (year 5). Medical students and physicians were mostly influenced by the shape of the curve between the years 4-5. In another study, participants were influenced by the distance between the curves rather than the numerical differences (Zikmund-Fisher, Fagerlin and Ubel 2005). The findings from these studies align with the literature which emphasises how important is to differentiate the graphs when presented to experts and non-experts and present what is most appropriate for each target group.

In summary, pie charts are effective when conveying proportions, but they need be accompanied by brief captions that state the total of the pie chart. Bar graphs are good to provide comparisons between risks. However, the way bar graphs are designed influence a lot the user's perception, hence information about x and y axis should be provided along with the absolute risk of the problem describing. Icon arrays could be useful when human figures are used, but when the part-to-whole information is not provided the denominator might influence risk perception, therefore they should be designed carefully not to inflate the magnitude of the risk difference. Line graphs are useful for changes over time but when used, the differences in audiences should be considered to address and overcome misconceptions.

The literature on graphical perception illustrates that are certain graphs appropriate for specific tasks and audience and can guide what to avoid or highlight. Therefore, the best design for a graph depends a lot on the purpose of risk communication. Some graphs are appropriate to enhance the understanding while others aim to promote behavioural change. For good quantitative judgments, an important feature is that the size of the graphic element should be proportional to the size that represents. For example, part-to-whole bar charts, pie charts and part-to-whole icon areas are able to promote people's proportion judgment and automatic visual processing (Cleveland and McGrill 1985) hence they are useful when people need to notice mathematical proportions. Bar charts are also useful for helping viewers to put the risk in context or compare with others but here as well perceptions could be influenced by the design of the graphics as addressed above. Graphs, which emphasise the risk ratio, tend to promote behavioural change. As mentioned before when part-to-whole information is not available then people are becoming more risk averse when looking at the numerator in bar charts and icon arrays.

Although graphs seem to be a more intuitive method than words, literature has shown that familiarity, expertise and education affect the accurate interpretation of formats. Therefore, a careful selection of a graph should take place relevant to the purpose of its in-home message.

2.5.2 Review of Visual and Graphical representation of lotteries in Holt and Laury

After reviewing the different ways of presenting probabilities, this section presents a variety of methods that have been employed by empirical studies to present the lotteries of Holt and Laury. Visual representation of probabilities was also employed in the popular risk measurements techniques Holt and Laury to reduce complexity, since it has been shown that participants who do not have the background or the training in numerical presentations, they understand better the risk information when it is accompanied by visual aids (Carlsson Johansson-Stenman and Martinsson 2004; Corso, Hammitt and Graham 2001). Therefore, several experiments with Holt and Laury task used a form of graphical representation to present the lottery options between the risky and the safe choice (Eckel, Warnick and Johnson 2005; Boughera, Gassmann and Piet 2011; Ihli, Chiputwa and Musshoff 2013; Bauremeister and Musshoff 2016; Habib et al. 2016; van Den Berg, Fort and Burger 2009; Akay et al. 2012). More details regarding those experiments are explained further to portray a whole picture of whether graphs could be effective and help people to understand such procedure easier.

Several studies attempted to improve understanding of the Holt and Laury task by presenting the probabilities in the lottery options visually. Eckel, Warnick and Johnson (2005) used an adaptive measure of Holt and Laury task where three out of eleven gambles displayed text only style -exactly as it is in Holt and Laury task - and the rest eight out of eleven gambles were presented as wheels. The measure has an algorithm which implements participant's next choice based on past choices. The researchers used this adaptive method to correlate it with the conventional Holt and Laury pen and paper task to investigate whether was any difference in 109 participants' risk attitude and consistency in two sessions. The results showed no correlation between the two instruments; people's choices in the adaptive instrument were not correlated with their choices in the textual and the authors claimed that was due to the high inconsistency rate of decisions in Holt and Laury textual method (Eckel, Warnick and Johnson 2005). Nevertheless, the pie charts did not lead to intuitive answers and did not make the procedure easier. Even though the numerical probabilities were replaced by pie charts, the size of the pie chart did not change depending on the amount of pay offs making it difficult for the audience to understand intuitively which pie chart will give him the better pay off. Additionally, the researchers investigated also how participants' risk preferences in the adaptive instrument would change after a period of two weeks. The exposure to the adaptive instrument again over a period of time revealed that people became more risk averse in their choices. The latter finding also agrees with the empirical research, which showed that people became risk averse after being exposed in a similar experience of risky options in the past and had lost (Guiso, Sapienza and Zingales 2013). Therefore, similar experiments for risk assessment should avoid being repeated after a period of time since people would have learned the task and perform differently.

Figure 6. Screenshot of a typical decision in Eckel, Warnick and Johnson (2005) study

Boughera, Gassmann and Piet (2011) illustrated the probabilities in Holt and Laury by using a turning wheel (similar to the show "Wheel of Fortune") to thirty well-educated French farmers with the purpose of making the method easier to comprehend to investigate the participant's risk preferences in expected utility and cumulative theory. Even though the sample was small, they found that farmers were more risk averse in the expected utility framework. This study, however, did not report any tests looking whether the different representational form influenced the consistency of participants' choices. Since the sample was considered well educated regarding the farmers who had a diploma (SD = 0. 47 and M= 0.70).

Figure 7. Screenshot of the turning wheel used by Boughera, Grassman and Piet (2011)

Another group of researchers (Ihli, Chiputwa and Musshoff 2013) replaced the numerical and textual representation of probabilities in Holt and Laury and of a task from Brick, Visser and Burns (BVB) (2012) with pictures of bags filled with different colour balls that were representing different prizes (Figure 8). Their aim was to compare risk preferences between the two elicitation tasks and determine the reasons affecting those preferences or inconsistent choices.

Figure 8. Screenshots of the bags filled with eight coloured balls (Ihli, Chiputwa and Musshoff 2013; Bauremeister and Musshoff 2016)

The researchers collected a number of socioeconomics variables, of which those related to this thesis are the age, gender and education defined as the years of formal schooling and a probability test score which was defined as the ability of the person to score in numbers using a task by Charness and Viceisza (2012). The experimental procedure was also supported by posters of graphical examples and by real bags with coloured balls to represent probabilities. To further determine that the participants fully understood the instructions, they had to answer some control questions. Therefore, the instructions of the measure were understood well enough because of the oral visual and written explanations. Only 19 participants out of 332 choose inconsistently regarding the times they switched options (those who chose only Option

B or switched more than one considered inconsistent). The low inconsistency rate compared to other studies in the field indicates that the instructions, as well as the two different methods of probabilities representations, helped people to understand the procedure. It is difficult to draw any conclusions about the visual display of the probabilities because there was also the natural representation of them. As per the author's suggestion to investigate how the original Holt and Laury design with textual probabilities compared to the adapted Holt and Laury with the bags of coloured balls with regard people's inconsistent choices, Bauremeister and Musshoff (2016) conducted an experiment with 307 students. Therefore, 158 students took the visual Holt and Laury (Figure 8) and the rest 149 the textual. The results showed that there is a big difference of people's inconsistency rate between the two designs (10.83%) since in the group of textual form the inconsistency rate was 23.49% and in the group with the visual form 12.66% ($p<0.013$). The inconsistency rate was determined from a number of times a participant switched options with participants who chose more than once a different option to consider inconsistent (Galarza 2009). This means that participants behaved more consistently with the visual format in Holt and Laury. Worth mentioning that the researchers also investigated whether people's attitudes differed between different formats by measuring the CRRA, which is the constant relative risk aversion. The differences between the CRRA and the standard deviations of the two presentational forms revealed that people's risk preferences did not change between the two displayed formats. For next step, it was suggested to determine the best graphical option for tasks like Holt and Laury MPL methods and also for which group a visual display might be beneficial (Bauremeister and Musshoff 2016).

Habib et al. (2016) employed four variations of visual displayed Holt and Laury compared with textual Holt and Laury. The variations were regarding the information provided along with the visuals e.g. full graphs with only the probabilities, or with only the prize, or with probabilities and prize or with no text. Forty-seven participants were exposed to each of the different display. The hypothesis was that the participants would be engaged differently in each display. Comparing the CRRA coefficients between the four visual displays the textual in a within-subject design, it was revealed that graphics tend to make people more risk-averse (Habib et al. 2016). However, as the participants were exposed in a trial before the experimental procedure and they were also shown all the formats it is possible that they learned and transferred insights from one treatment to the other without letting us conclude whether the visual format made is easier for them or reduced their inconsistent answers.

Some materials have been removed due to 3rd party copyright. The unabridged version can be viewed in Lancester Library - Coventry University.

Figure 9. Graphical display with text prize and the probability (Habib et al. 2016)

In addition to graphical representations, some researchers, similar to Ihli, Chiputwa and Musshoff (2013), experimented with real visual aids to make the Holt and Laury procedure easier to understand for illiterate people. Therefore, in an experiment conducted by Akay et al. (2012) between 92 Ethiopian farmers, one of the poorest regions of the world, who were illiterate, to investigate their ambiguity aversion, they were given posters as visual aids for the instructions that were also given to them verbally in their local language. The probabilities were demonstrated to them using balls and dice as the researchers considered them necessary for their sample (Akay et al. 2012). The participants were asked to choose a risky and safe option between a white and yellow ball, and the procedure took place in a private room for each participant separately (Akay et al. 2012). As a control group, the experimenters used 69 students from a Dutch university in a similar procedure, which is often used as a benchmark in similar studies (Trautmann, Veider and Wakker 2011, experiment 4). The result showed that the farmers exhibited more risk aversion compared to Dutch students, the median coefficient of relative risk aversion, which is calculated with the formula $\rho = cA(c) = -cu''(c) / u'(c)$ where u(c) is the constant absolute risk aversion showed that for the Ethiopian sample was $\rho = 0.73$, which is significantly larger than the median of $\rho = 0.34$ in the Dutch student sample (Mann–Whitney U test, $z = 4.391$, $p< .01$). This difference could be explained by socioeconomic factors. As far as the experimental procedure seemed straightforward for illiterate people, it can be difficult to perform, as it demands a lot of time and the intervention of researchers who may affect the participant's judgment (experimenter-bias). However, this procedure highlights the benefit of using visual aids to simplify the procedure, as there were no inconsistent participant choices.

Similar procedure was followed in a study in Peru and Nicaragua where participants were asked to imagine that they have to choose between a project A of success or failure and Project B with safer profits instead of just giving them a textual Holt and Laury because the majority of them were illiterate (van Den Berg, Fort and Burger 2009). The study was examining how natural hazards experiences affect risk aversion. The two-possible pay-offs of project A were associated with the toss of a coin so as participants would understand the probability of each project. Additionally, there was another measure of risk aversion, asking participants to answer two hypothetical questions about how much money they would be willing to pay for a lottery ticket. The above experiments were conducted in the form of two games to make it easier for people with low education. The researchers found that farmers who have experienced natural hazards are substantially risk averse (van Den Berg, Fort and Burger 2009). Even though, no inconsistent results were reported from the researchers, the correlation rates between coefficients of risk aversion and certainty equivalents are not significant suggesting that one or even both tasks did not measure risk aversion accurately. Therefore, there is no sufficient evidence of what exactly helped participants to understand.

Without a doubt, posters, examples or showing the natural objects with bags filled with balls enhance people's understanding for the lotteries in the task, but they are time-consuming and sometimes not compatible with online testing. Bauremeister and Musshoff (2016) managed to reduce participants' irrational choices by 10.83% with the use of bags filled with colourful balls in place of numerical probabilities in Holt and Laury task. Nevertheless, Hegarty and Kozhevnikov (1999) identified that the use of graphs in solving probability problems lead to higher rate of success compared to the use of pictures. This finding aligns with study's results presented in Chapter 5, which showed that participants who used graphs, more specifically pie charts, had a bigger chance in evaluating correctly the probability in a Holt and Laury example. Pie charts have proved to be familiar to the general public, which makes them easier to comprehend (Nelson, Hesse and Croyle 2009). Nonetheless, sometimes pie charts were shown to be problematic for people. As it was identified by Fieldman-Stewart et al. (2000) and Willkinson and Wills (2008), people have difficulty in estimating the angles or to be accurate about the proportions that pie charts involve. For instance, in a study conducted by Fieldman-Stewart et al. (2000), that assessed accuracy and speed, of patients and students in six formats (included pie charts), participants showed least accurate in estimating the quantities with pie charts compared to other formats. Undoubtedly though pie charts were used by several studies with Holt and Laury task (Eckel, Grassman and Johnson 2005; Habib et al. 2016) showing a great potential of using them instead of the numerical format to reduce participants' inconsistent answers. However, the problems with them lead us to investigate whether prompting participants to use pie charts as external representations would assist in their rational decision making.

2.5.3 Summary of graphical presentational ways of probabilities and lotteries

This section has reviewed the use of graphs to communicate probabilities and by outlining the variety of graphs used to present risk information and especially in experiments with lottery risk elicitation task by Holt and Laury. The empirical evidence presented in this section are related to the need of a new way of probability presentation in the risk elicitation task of Holt and Laury which could involve prompting individuals to use graphical external representations to reason their choices to the task.

Empirical evidence has shown that people with low numerical skills and impulsive thinking style tend to choose irrationally in lottery choice tasks such as Holt and Laury when probabilities are presented with numbers (Dave et al. 2010; Frederick 2005). The methodological approach improving the representational format of lotteries involves the investigation of consistency of reasoning using a graph to identify which difference in presentational format can help people understand the probability of the lottery better. Empirical evidence has shown that people's perception of risk is strongly influenced by the design of the graph (Stone et al. 2003). To that end, before, rushing into using any kind of graph to represent lotteries in Holt and Laury task, a systematic review can outline the strengths and weaknesses of the popular graphs used in similar tasks. From the systematic review of the graphs, it was shown that pie charts along with bar graphs with brief captions that state the whole information, for example, the total of the pie chart, are necessary. Drawing on these findings, a review of the variety of visual and graphical representation of lotteries in Holt and Laury task takes place. The latter review leads to the conclusion that pie charts have the greatest potential of reducing people inconsistent choices in Holt and Laury. However, in some case pie charts were shown to mislead people as they have difficulty in estimating the angles to be accurate about the proportions (Fieldman-Stewart et al. 2000; Willkinson and Wills 2008). Therefore, as the literature review presented in Section 2.4.3 and 2.4.4 suggest that reasoning with the use of graphs probabilities provides a high likelihood of successfully solving the problem, then engagement of people to draw graphs in the context of a gamified lottery task can have better results in influencing people's rational.

2.6 Formation of High Level Research Question

This section summarises the key points of the literature review to formulate the main high - level Research Question. Assessing people's risk propensity towards risky decisions related to investment choices is valuable towards understanding and predicting such decisions. As mentioned in the first section of the literature review, decisions related to allocation of assets and investment or insurance choices could be predicted by an accurate risk elicitation task. Even though, a variety of risk taking and risk propensity exists, this thesis focuses on the amount of uncertainty as presented in the form of numerical probabilities someone is willing to accept. Nevertheless, the risk elicitation tasks that have been developed were shown to have certain weaknesses that compromised their effectiveness. These weaknesses involve the complexity of their procedure, lack of differentiation point between risk neutrals and risk averse, difficulty to implement outside of laboratory settings and the need for incentives to have truthful results. Among all the risk elicitation tasks related to such decisions, Holt and Laury choice task, when it is understood, is considered to have the most significant predictive value compared to other risk elicitation tasks. However, Holt and Laury's task major weakness is that the experimental procedure is overly complex as the lotteries are presented in the form of numerical probabilities which produce a high level of people irrational choices and therefore noisy outcomes. To address the problem of participants' irrational choices in the risk lottery elicitation task of Holt and Laury that compromise its predictive value, this thesis investigates the impact of using visual elements and graphs so people can understand the lotteries and choose rationally. In-game assessments that have been developed so far, have shown certain benefits that can improve a standard pen and paper task. Among these benefits that are mentioned in Section 2.3, using visuals that allow simplification of tasks indicate that a gamified task of Holt and Laury with visualised lotteries could potentially have better results in rational decision making. The rest benefits that games could offer such as replacement of incentives and distribution through smartphones could be examined in a later stage after the issue of complexity has been investigated.

This thesis is focused on the benefit of incorporating visuals and more specifically graphs to improve a risk elicitation task by reducing participants' irrational choices. In addition to using visuals inside a game to simplify the complexity of Holt and Laury's task, empirical evidence on visual thinking has shown that when people use external representations to reason probabilities, have greater likelihood of successfully solving it (Zahner and Corter 2002). From the variety of external representations that people use, graphs were shown to help people more in the problem-solving procedure (Hegarty and Kozhevnikov 1999). Therefore, an interactive approach that prompts participants to use graphs while solving the risk elicitation task of Holt and Laury could help their rational decision making. According to the empirical evidence from the literature review presented in this chapter, the high-level research question is formulated as follows: Can incorporating visual elements and interactive approaches with graphs in a risk elicitation task influence rational decision making? The following chapters would present the methodology and the methods employed for three experimental studies that aim to answer this research question.

3. Research Methodology

3.1 Introduction

The previous chapter contemplated on the literature review of the topics relevant to this thesis, which led to the high-level research question: Can visual elements and graphs incorporated in a risk elicitation task influence rational decision making?

This section presents the overall research philosophy and strategy used to address this question. In doing so, it outlines the methodology and methods of three independent studies. The first of these experiments examines whether the introduction of gamified elements to a risk elicitation task helps participants to choose rationally. The second experiment examines the factors that influence participants' performance on lottery risk elicitation tasks to choose rationally between Holt and Laury task options. The third experiment explores the effect of using an interactive approach in the Holt and Laury task that could influence participants' rational choice. The following sections describe the process of designing the experimental instruments, which include a game, the interactive adaptations of Holt and Laury task and the standard pen and paper Holt and Laury task.

Therefore, in detail, the objectives of this chapter are the following:

- Present and justify each study's research strategy.
- Describe the experimental design.
- Outline the dependent and independent variables in each study.
- Describe the data collection, the sample size and related issues.
- Discuss the analytical methods used.
- Present the ethical considerations regarding this research study.

3.2 Research Philosophy and Strategy

The focus of this research is to improve people's performance on lottery risk elicitation task with the use of interactive visual and graphical approaches. Therefore, the high-level research question requires a research methodology, which involves experimental studies that investigate different aspects which can help with people's performance by employing interactive and gamified approaches while observing the outcome.

This thesis, therefore, employed an empirical approach. According to Locke (1948), the empirical approach involves a systematic attempt to reach knowledge by direct or indirect observation. Peters et al. (2006) employed empiricism approach and conducted a series of four studies that examined how the ability to comprehend probability number affect judgment on decision-making tasks. Similarly, Holt and Laury (2002) used an empirical approach to identify whether scaling up the pay offs from the lottery task would increase participants risk aversion. The empirical evidence which includes the collection of observations can be analysed both qualitative and quantitative depending on research question asked. Through quantifying the empirical evidence, the researcher can answer the research questions, which should be clearly stated and answered with all the collected data. Then the researcher is trying to describe the interactions between human and the instrument while observing a specific entity. Based on past empirical evidence the hypothesis that is generated leads to an experiment. Depending on the results of the experiment whether the theory which the hypothesis was based will be supported or not, hypothesis may need to be modified and subject to more testing. In the same manner, this research started with a generated hypothesis that lead to a pilot study (experiment 1) which results did not allow either to reject or accept the null hypothesis because of the noisy data, therefore a new hypothesis needed to be formulated and be tested again which lead to experiment 2. Based on results of experiment 2, the hypothesis had to be modified again and tested in experiment 3. Therefore, three quantitative experiments were designed to collect empirical evidence that would answer the high research question whether incorporating visual elements and graphs in a risk elicitation task would influence rational decision making. These three experiments were looking to answer the high-level research question by answering three research sub-questions.

1st research sub-question:

Does the introduction of game-based elements into a risk lottery task help rational choice in a risk elicitation task?

2nd research sub-question:

What are the factors that influence rational decision making in a risk elicitation task?

3rd research sub-question:

Does the use of interactive pie charts influence rational choice in a risk elicitation task?

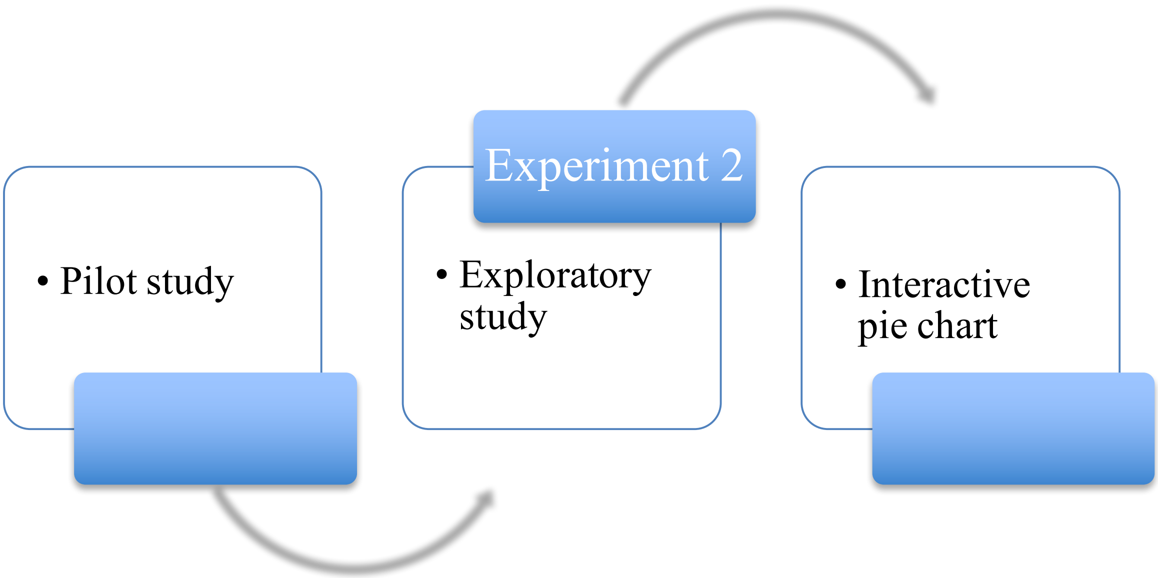


Figure 10. Research methodology

Experiment 1, which was designed as a pilot study, investigated whether the use of gamified elements, such as visuals, can assist in rational decision-making which would reflect the understanding of the experimental procedure and the simplification of the task. Experiment 1 was the study that would indicate whether the noise in the data would persist after the visualisation of the lotteries and the gamification of the risk elicitation task. However, its results lead to regeneration of hypothesis which lead to experiment 2. Experiment 2 investigated the underlying factors that influence individuals' rational decision making. This experiment aimed to understand the most profound factors to explore how to assist them. This is linked to the potential factors that were shown from previous empirical evidence (Dave et al. 2010; Frederick 2005; Zahner and Corter 2002). Building on the results of experiment 2, where using graphs as external representation showed to have a significant effect on rational decision making in the Holt and Laury option, experiment 3 examined whether the use of interactive pie charts influence rational decision-making. This experiment aimed to extend the empirical finding suggesting that using external representations assist in rational decision making (Zahner and Corter 2002) and using graphs is more appropriate for probability problem solving (Hegarty and Kozhevnikov 1999). The findings are reported in the following chapters, and the relation with the high-level research question is discussed further in Chapter 5.

Even though this thesis employed quantitative methods other research approaches in this area have adopted qualitative methodologies with empirical studies (Zahner and Corter 2002), these have strengths in providing richer data on narratives and cases of the use of external representations in reasoning and judgment probability problems. However, this thesis adopts a similar approach to Peters et al. (2006) and Bauremeister and Musshoff (2016) that they used an empirical approach finding the answer to their main research question through a series of experimentation that answers different angles of the same problem.

3.3 Organisation of the study

The methodology employed for these three quantitative experiments is identical however the methods of each procedure differ. Therefore, the methods used for each experiment are presented separately.

Data in **Experiment 1** *(Pilot study)* were collected using a quantitative method of analysing observations through experiments. One game was developed to assess risk preferences via lotteries based on the standard pen and paper Holt and Laury (2002). Participants' risk choices in the game were correlated with participants' choices in a validated computerised risk task, Balloon Analogue Risk Task (Lejuez 2002), that was used to establish the construct validity of the game. Before that, questionnaires were given to a sample of participants to establish game's usability satisfaction to ensure that the game was playable and usability issues would not affect participants' performance on it. The results from the construct validity study showed noise among data, which indicated participants' irrational and inconsistent choices. Results revealed no significant correlations between the choices in the two tasks.

In **Experiment 2** *(Exploratory study of factors that influence participants' rational choices)* a quantitative method of analysing observations was also used. Experiment 2 shows different factors that influence people's choices in lottery options. Numeracy, education, cognitive thinking style (automatic or reflective) and external representations were investigated against the evaluation of a Holt and Laury option. In this experiment, it was examined whether the use of external representations that participants would help them to reason the Holt and Laury option rationally. The results showed that participants with low numerical skills and automatic thinking style have greater likelihood of reasoning rationally the choices in the Holt and Laury task by using external representations, more specifically graphs.

In **Experiment 3** *(The interactive pie chart approach),* a quantitative method of analysing the observation was employed. An interactive pie chart approach, similar to the approach of using external representations, for the Holt and Laury task was compared against the conventional method of Holt and Laury and a passive visual aid Holt and Laury. This experiment showed that the interactive visual approach improved participants' performance and reduce their rational choices significantly compared to numerical and passive (where the user could not manipulate the visual) visual lotteries in the task.

3.3.1 Experiment 1

The purpose of this pilot study was to examine if a correlation existed between the outcomes of an established risk-elicitation task, and a game-based incarnation, to test the validity of a game as a construct for eliciting the risk propensity of an individual. Game ("Risk House") was derived from a lottery-based risk elicitation task (Holt and Laury 2002). Additionally, the experiment investigated whether game based elements would help participants reducing their irrational choices.

It was essential to ensure that the game did not have any usability barriers, as these may lead to influenced responses. Hence, the primary data capture method, the game, was assessed for its usability before the construct validity study. Usability research, in general, is a core component of the system development process and should be given significant consideration to provide helpful services (Rose et al. 2005). The usability satisfaction assessment of any software developed gives the users the opportunity to test and give feedback on the production of a system that can potentially use. Therefore, a usability satisfaction questionnaire was given to the participants to fill in, after they had played the game.

Following that, the game was correlated with other validated assessment, named the Balloon Analogue Risk Task (BART), a computerised method that measures the same aspect of risk taking. The rationale for using a quantitative approach was to determine whether individuals' risk preferences would be similar in the game-assessment tool and in the validated risk elicitation task. To do that, people's risk choices in each task were compared, and statistical methods were applied to determine whether an association exists. Hence, a quantitative approach was the most appropriate to establish the correlation of the constructs. The description of the method used to ascertain the validity of the construct is described in Section 3.3.1.2.

3.3.1.1 Method for usability satisfaction of the game

The usability of the game needed to be investigated to avoid having outcomes that would be affected by usability issues. Therefore, to ensure that usability level of the game was in a good level, the initial prototype of the Risk House game was evaluated to assess usability using the standardised Questionnaire for User Interaction Satisfaction metric. The reason for choosing this questionnaire for assessing usability satisfaction in this study is that previous tools developed to evaluate the human-computer interaction had several validation and reliability problems (Ives, Olson, & Baroudi, 1983). The overall reliability of QUIS is substantive, Cronbach's alpha of 0.96 (Chin, Diehl and Norman 1988), which is greater than alpha=0.7 that Lewis has suggested as the appropriate (Lewis 1995). The mean correlations of construct validity between each item (within the six sections of the questionnaire) and another satisfaction scale ranged between .49 and .61 (SD= .09 - .12) (Harper and Norman 1997). The latter shows a good agreement between the constructs, thus good construct validity.
The results of the QUIS showed a good overall usability and highlighted a few areas of improvement that are presented further in the section. The method described below presents the recruitment of participants, description of the materials used, the procedure along with the variables manipulated and the analysis performed.

Participants

Participants were recruited via emails to academic staff and graduate students in Coventry University where this PhD is undertaken. All of the participants were part of the School of Computer and Engineering of Coventry University, for convenient reasons ten Computer Scientists took part in the usability evaluation of the two games. The Computer Scientists were between the ages of 18 to 33 (M age = 24.8 years, SD = 5.13, 5 men and 5 women). All of them participated volunteering. Thus, there were no incentives for participation in the usability study, and this justifies the small number of participants.

Materials

Questionnaire for User Interaction Satisfaction: The questionnaire that was used to assess usability satisfaction of the game was the Questionnaire for User Interaction Satisfaction (QUIS) designed by Chin, Diehl and Norman (1988). The QUIS was designed to assess user's subjective satisfaction on a variety of aspects of a tool's interface (Harper and Norman 1993). The QUIS questionnaire consists of six sections (Karoulis, Sylaiou and White 2006). Participants are asked to rate each item within these 6 sections in a 9-point Likert scale. Section 1 includes questions that focus on the backgrounds of users, to establish the computer systems they are most familiar with. Section 2 of the questionnaire focuses on the overall reactions of the users to the system. Participants expressed how they felt while playing the game regarding semantic bipolar differentiated expressions (terrible to wonderful, difficult to easy, frustrating to satisfying, inadequate power to adequate power, dull to stimulating, rigid to flexible). In Section 3 the GUI of the game was tested such as the readability of displayed characters, windows layout, etc. Section 4 of the questionnaire, the participants, shared their opinion regarding how easy or encouraging was to learn the Interface of the game. In Section 5 of the questionnaire, the questions were about the system speed and reliability. The participants answered accordingly. Section 6 of the questionnaire was about the multimedia content in the games. The multimedia content includes the images inside the virtual application, the movies, the sound and the nature of the colours. This content enhances the user's experience and adds depth to the game.

Design of Risk House game: The Risk House game was designed to assess participants' risk propensity thus the design had to be simplistic enough to avoid the distraction of the participants (Anetta 2010). The core design of the game aimed to reflect a stealth assessment, as it has been described in Section 2.3.1 of Chapter 2, and create an environment for the player that would make him unaware of the assessment inside the game. The component of aesthetics and graphics defined the overall format and the mechanics (rewards) of the game. Risk House game's aesthetic was based on the 'Great Brain Experiment' which was described in Section 2.3.1. The 'Great Brain Experiment' which involved a risk propensity assessment with a colourful cartoonish style that is framed by catchy happy music. The scene of the game play is very simplistic with just apples falling from a tree and the score. Similarly, Risk House game is contextualised in a cartoonish colourful style with catchy music. The visual design includes the two rotating wheels which are considered as roulettes in the game and the element of a building which depends on participants' score. Therefore, the core of the game's graphics are the rotating wheels that the participants could interact with (click and rotate them). As it has been described in Section 2.3.1, another significant element of effective video game design is the system of goal and objectives which is common to all in-game assessments. The system of goal in the Risk House game is reflected on the score and the visual of tent which changes to a palace as long as the player progresses in the game. The whole design of Risk House game is built around the dynamics of the pen and paper Holt and Laury task which involve the lottery choices. These lottery choices are transferred to the game in the form of visual roulettes, since considerable motivation exists which supports that games can incorporate measurements from pen and paper questionnaires in a less invasive way (van Lankveld et al. 2011). Since risk propensity is measured in the Holt and Laury by participants' switching points between the lotteries, this mechanic should be presented in a similar manner in the game to avoid invalid predictions of risk propensity. However, a game element was needed to justify the purpose of the game. The game element used was the upgrade from a tent to a palace eventually.

The Risk House Game: The results reflected a positive overall consensus on usability amongst ten experts that filled in the QUIS questionnaire. From the QUIS questionnaires the lowest Likert-mean is 5.7, whereas the highest mean value is 7.65 with the maximum of 10, M = 6.86 of maximum 10 and SD= 0.62. Summary data of the descriptive statistics of each section (apart from the Section 1 which includes the demographic questionnaire) are presented in Table 4. The overall data for each of the 5 sections of QUIS questionnaire are descriptively presented on Appendix 9.

Table 4. Descriptive Statistics of all QUIS Sections – Risk House Game (Appendix 9)

Mean Values of 2, 3, 4, 5 & 6 Section					
	N	**Min.**	**Max.**	**Mean**	**Std. dev.**
Overall reactions to the game	10	5.2	7.2	6.5	.75
Screen	10	5.7	7.5	6.7	.79

Terminology and system information	10	7.2	8.1	7.6	.37
Learning	10	6.6	8.1	6.5	.59
System capabilities	10	6.6	8.1	7.0	.64

In a 9-point scale Likert scale and taking into account the relatively high values people most often give in a questionnaire like surveys (Karoulis, Sylaiou and White 2006), the result of this evaluation showed a good level of usability of the interface but also highlighted some improvements to be made. The categories for each dimension such as multimedia, graphics, and windows/characters layout are related to games and showed a good level of usability. In particular, the interface was considered easy enough: satisfying but a "bit dull" and rigid. There were some difficulties regarding the numbers, as they seemed to be a bit small, so the size of them was increased. The learnability of interface was very easy; this can be explained by the fact that the Risk House game does not demand high gamer skills. Multimedia was satisfying for all the users. Some of the participants' comments for Risk House game were: "System was quick and easy to play" & "Colours/sound good and clear". The overall feedback showed that respondents enjoyed the game and the general usability evaluation had fairly positive results.

To formulate the lottery options in the Risk House game, information was gathered from the risk elicitation tasks developed by Holt and Laury (2002) and Eckel and Grossman (2002) (e.g., Holt and Laury 2002; Eckel and Grossman 2002, 2008).

The Risk House game was designed to assess risk taking within a game play, and the elicitation mechanism is based on an adaptation of the Multiple Price List task by Holt and Laury (2002). In the Multiple Price List task, participants are asked to choose between two options over ten rows. These two options involve one Lottery that is risky, and the other Lottery is safe. Over the ten rows, the level of riskiness and safeness of each Lottery changes.

The switching point that the participant would choose to move from the safer to the riskier Lottery determines the participant's level of risk taking. In the Holt and Laury task, the lotteries are presented with the form of numerical probabilities. In the game, the options are presented in the form of roulette so as to improve participants understanding of them. When players click on the graphs with the numbers these rotate and wherever the mark stops is the amount of points they earn for the round. Players are presented with two lotteries to gain more points. The ten first pair of choices is the example for the player to understand how the game is played. The rest ten pair of choices are the assessment options. The lotteries are divided between one ‘risky’ option, the roulette, where are equal possibilities either to win more or to lose more, whereas the ‘safe’ option involves a number that changes but cannot be a greater loss than the lottery.

A game element used in Risk House game is that by gaining more points, you move from a very small house to a bigger one. It has the intention of representing the emotion of ownership (i.e. you do not risk so easily to lose your house). This choice based game is grounded on the Prospect Theory (Kahneman and Tversky 1979) and the certainty effect. According to the latter people tend to overbalance outcomes that are certain compared to outcomes that are probable. The starting score that the player has is 100, and his house is a slum. From that point, there is the possibility, depending upon the participant’s choice, to either upgrade to a better house or move down to a tent.

Figure 11. Interfaces of Risk House game

3.3.1.2 Method for the construct validity of the game

The second part of this experiment concerns the construct validity of the game. Construct validity is an assessment of the quality of an instrument or experimental design. It also includes the relationship between the construct and measures of other constructs. There is no simple metric that can be used to quantify whether a measure can be determined as a valid construct. Typically, researchers establish constructs validity by correlation studies between the measure and other already validated measures that should be related to it or vary independently of it (Westen and Rosenthal 2003).

A quantitative research approach of analysing the observation was employed for this study, to correlate participants’ performances between the game and a pre-existed construct. The data were gathered according to a specific plan in which the game along with other pre-existed construct was used to collect the needed information. Hence, the quantitative approach is used to measure the construct validity of the game. The method described below presents the procedure of gathering participants, description of the materials used, the procedure along with the variables manipulated and the analysis performed.

Participants

Participants were recruited via emails sent to academic staff and graduate students within Coventry University, UK, as it was easier for them to participate and visit the laboratory in the university's premises. The participants were between the ages of 18 and 33 years (M age= 24, SD = 4.665). They were all invited to take part through e-mails, and they were selected by snowball effect (Ghemawat 1990), mostly by sending them emails and asking to forward it to fellow colleagues. In general, 32 out of 50 participants that were contacted, participated in the experiment. The participants received payment according to performance to the Balloon Analogue Task (BART), which was one of the additional constructs used for this study. BART is a computerised behavioural task that measures risk behaviours by presenting to the participants balloon trials, that they have to inflate. It was developed by Lejuez et al. (2002) to overcome the issues of self-report in risk taking assessments. BART was chosen as it has shown high internal consistency reliability in previous studies. The amount paid was based on some pumps a participant would perform in the BART task. Each pump costs 0.05GBP. Therefore, the payment varied from 3GBP minimum payment to a 20 GBP maximum. The average payment per person was approximately 10 GBP with a few exceptions. There was a payment for this construct as it has been proved to have valid results by incentivising the participants with money (Lejuez et al. 2002), which also is the case for the majority of the risk assessment tasks (Holt and Laury 2002).

Materials

The materials used for this study include the Risk House game described in Section 3.1.3 and an additional measure (BART) that elicits risk preferences. Therefore, the Balloon Analogue Risk Task (BART) was used to determine the games construct validity.

BART: BART is a computerised behavioural task that measures risk behaviours. It was developed by Lejuez et al. (2002) to overcome the issues of self-report in risk taking assessments. Even though the majority of risk elicitation tasks have been criticised and there has been no established risk assessment (Rohrmann 2005), BART was chosen as it has shown high internal consistency reliability in previous studies. The measure, in previous studies, showed test re-test, high internal consistency reliability (r = 0.86). In another study conducted by White, Lejuez and de Wit (2008), BART showed that the mean risk behaviour did not differ across test dates and the test-retest correlations across sessions were high (r = 0.77). Additionally, BART was significantly correlated 'with scores on self-reported risk related instruments and with the self-reported occurrence of real world behaviour' (Lejuez et al. 2002). Moreover, it is a computerised task sharing a few game mechanics such as a reward system and the sense of flow, which is a similar concept to the Risk House game. It also consists a mechanism of sequential risk-taking with feedback; it is very easily manipulated with changes that the experimenter can do in the models, and it provides a level of analysis for the individual risk behaviour and not in a group context. The participants in this task are faced with some balloon-trials on the computer. In each trial, on the computer screen is presented a small simulated balloon with a button below which represents a balloon pump, on the left side of the screen is a button labelled collect and from the right side a money-earned display labelled Total Earned, as shown in Figure 12.

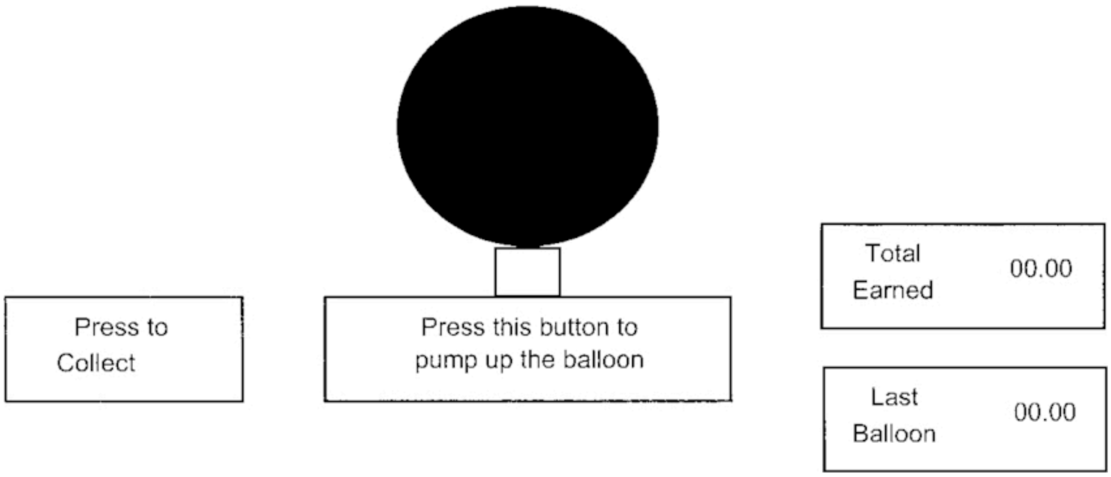


Figure 12. Balloon Analogue Task Screen

When the button-pump is clicked the balloon inflates approximately 0.3 cm in all directions. Each pump causes 2 pennies to be added to a counter until some threshold at which the balloon is overinflated and explodes. Therefore, each pump gives greater risk but also provides the possibility of greater reward. When a balloon explodes, the money that the participant has earned through pumping at this trial are lost. However, at any point through each balloon trial, the participant could press the *Collect* button and transfer the money form a temporary bank to the permanent during which the new amount of money earned will be updated. In our study, the payment was set to two pennies per pump. The maximum payment per person is 20 GBP. The average payment per person was approximately 10 GBP with a few exceptions. After the collection of the money or the explosion another balloon appears on the screen until a total of 20 balloons to be completed. Each balloon has a different probability of exploding and participants are not informed about it. They are just told, "It's up to them how much they want to pump the balloon. Some they can pop very soon but some others not until they fill in the whole screen." There is a maximum number of n pumps and each balloon will explode with a priori probability 1/n. Hence, the first pump, the probability of its explosion will be $1/n$; on the second pump (if the balloon wouldn't explode in the first one), the probability will be $1/(n\text{-}1)$ and so forth. Overall, the probability s_i that an explosion will occur on a pump if the previous balloon had not exploded in the preceding (I - 1) trials is: $s_i = 1/\ n - i + 1$. The numbers that are manipulated in this paradigm includes the number of balloons in a sequence that holds x and n constant. The average number of pumps is an indication of riskiness (the more someone pumps the balloon, the greater the degree of their risk propensity).

Data Collection and Procedures

Participants were tested individually in a quiet room at the Psychology Research Laboratory in the Department of Psychology at Coventry University. The private room included a desk Dell computer, a monitor and a mouse. At the beginning of the experiment the participants were briefed about the research project and described the general details of the study. Immediately after, they were asked to read the Participant Information Sheet to sign the consent form. Then the participants were given the game to play and the additional construct. At the end of the session, participants were given the amount they gained from the BART task and the debrief sheet.

Ethical considerations

According to Beauchamp (2008), which summarises ethical considerations and guidelines to experiments that involve participants in behavioural tasks, are three basic principles. These are the following:

- Respect for persons
- Beneficence
- Justice

The principal of respect for others indicates that individuals should be treated as autonomous people. For that reason, in this study, it was ensured that participants were fully informed about the research objectives and informed consent was obtained before the experiment started (See Appendix 1). The participants also had opportunities to ask questions to the investigator. The second and third ethical principles that are considered are the principles of beneficence and justice. These two refer to the duty of the investigator to maximise the possible benefits and to minimise possible risks that might result from this study, as well as to be fair to the participants. Another ethical principle that we would consider in this study is also the principle of autonomy, which refers to self-determination and the right to full disclosure. Therefore, the participants were fully informed about the procedure, nature and they had the opportunity to decide whether they wanted or not to participate in the study. It was also ensured that all data were treated with appropriate confidentiality and anonymity. In our studies, the investigator ensured that confidentiality was implanted in all procedures. The main researcher kept all the documentation locked in a safe desk. These documents will be destroyed after five years since the end of the study.

Data analysis

The data analysis was performed in SPSS software and the statistical method used was correlation. Correlation describes the relationship between two variables. Therefore, a correlation analysis was performed between participants' risky choices in the Risk House game and their choices in BART task. From Risk House Game, the outcome variables were the number of risky choices and were correlated to the number of explosions a participant would perform in BART task. Then a correlation analysis was performed to indicate the relation between the outcome variables and therefore the constructs. Additionally, the number of participant irrational choices was counted.

3.3.2 Experiment 2

The purpose of this experiment was to examine the factors that influence people's irrational evaluation of the options in a lottery risk task and whether using external representations would help them to evaluate the lottery options rationally. Participants' inconsistent choices in risk elicitation tasks were identified at various empirical studies with Holt and Laury tasks, as it was discussed in the literature review (Dave et al. 2010; Eckel and Grossman 2002). The similar inconsistent pattern of choices was also noticed in Experiment 1, where the numerical lotteries of Holt and Laury task were replaced with visual roulette, highlighting that the visualisation of the lotteries did not help participants to understand them. Therefore, further investigation is required to explore the factors that influence participants' performance on the task and investigate whether using external reasoning can help them to understand the lotteries of the task.

Empirical research showed a few indications of potential factors that influence people's choices in Holt and Laury tasks, these include: the educational level, the numeracy level, the cognitive thinking style (reflective or automatic) and the way probabilities are presented (Frederick 2005, Charness and Viceisza 2012; Dave et al. 2010; Peter et al. 2008). Consequently, an exploratory research was conducted to understand people's underlying thinking process and which one of the above potential factors plays the most significant role in influencing people's lottery choices. Participants were also asked to use external representations to reason their choices such as pictures, graphs, numerical calculations or text. According to the findings from literature that is presented in Chapter 2, Section 2.4, it was suggested that external representations help people to solve probability problems successfully. Therefore, a hypothesis was formulated that participants who have difficulty in understanding numerical probabilities, when using external representations can choose rationally in this task.

3.3.2.1 Method of the exploratory study

The 2nd Experiment involves an exploratory study that aims to answer the research sub-question: What are the factors that influence participants to choose rationally in a risk lottery task? A quantitative research addressed the above question.

Participants

Sixty people participated in the study. According to the parameters, mentioned earlier, the participants were divided into two groups depending on their educational level: One group had completed a university degree-level or higher qualification and included 35 participants. The other group had not completed a university degree-level qualification and consisted of 25 participants. Group differences are described further as follows:

Group 1: Thirty-five participants (21 male, 14 female) ranging in age from 22 to 53 years old (M age 29.5, age range: 22-53) volunteered to participate in the experiment and had completed a degree level or higher qualification. The participants were invited through snowball effect via the network of Hildebrand Technology Ltd and Coventry University that this PhD was supported from. All the participants signed informed consent.

Group 2**:** Twenty-five students from Leicestershire College in the United Kingdom (4 females, 21 males), ranging in age 18-37 volunteered to participate in the experiment (M age 20.64) and had not completed any degree-level qualification. The male participants outweighed significantly the number of females, and this created some limitations in our study. In general, participant ages were predominantly between 18-20, except two participants who were 27 and 36. All of them were assigned to the same experimental task as Group 1, and informed consent was obtained from all the participants of this group as well. Participants were students in the Game Design and Web Design courses at a remedial programme in Leicestershire College. None of the participants had attended a course at a university level.

Table 5. Participants' profile according to sex and highest degree awarded (Appendix 9)

	Group 1 *N*	**Group 2** *N*
Sex		
Male	21	21
Female	14	4
Highest degree awarded		
High school	-	25
BSc	14	-
MSc/MA	14	-
PhD	7	-

Materials

The materials used for this study include a numeracy test, a cognitive reflection test and two-probability problems. One probability problem involved one of the lottery options of Holt and Laury task. The other problem involved a conditional probability problem giving it as an example to participants to understand the procedure of reasoning a probability problem with the use of pictures, graphs, words or numbers.

Numeracy test: The Lipkus Numeracy scale developed by Lipkus, Samsa and Rimer (2001) was employed for this study. It was selected among others as a numeracy assessment tool for this study because it has been the most used in research and has been used to assess basic arithmetic skill in the variety of groups (Peters et al. 2006; Schapira et al. 2012). It is also short, involves only 11 items, and consists of very basic probability questions, which was needed for this study. These 11 items assess how well people can transform probabilities to percentages, percentages into probabilities and perform simple mathematical operations using percentages or probabilities. For example:

Which of the following percentages represents the biggest risk of getting a disease?

1% 10% 5%

The possible total sum scores range 0 to 11, where higher scores indicate better numerical skills compared to lower scores. The Lipkus Numeracy scale has moderate internal consistency, with Cronbach's alpha scores ranging from 0.70 to 0.75 in three separate studies (Lipkus et al. 2001) which make it a good measure of numeracy for this study.

Cognitive Reflection Test: The Cognitive Reflection Test (CRT) (Frederick 2005) is a three-item test, which is designed to assess individual's ability to suppress an impulsive wrong answer in place of a more deliberative cognitively processed correct answer. The decision to choose the CRT as a cognitive thinking measure was because it provides a theoretical foundation for understanding and reasoning part of people's heterogeneity in decision making as it has been addressed by Moritz, Hill and Donohue (2013). CRT is also an easy task (when explained people understand it easily), and also it reveals a reflective thinker over an impulsive since the most intuitive answer is the wrong one, so the individual needs to reflect a bit before finding the solution, which fits our study purposes.

The way to score this measure is by adding up the correct answers. Participants who scored 0 and 1 out of 3 were classified as low reflective thinkers, and those who scored 2 or 3 out of 3 were classified as high reflective thinkers according (Frederick 2005). In our study, the three questions were presented in a different context than Frederick's. The context was slightly changed because Frederick's cognitive reflection test is very well known and we wanted to avoid people knowing the answers in advance. The change in the context did not affect the purpose of each question at all.

Table 6. Frederick's CRT vs this study

Frederick's CRT	This study
A bat and a ball cost $1.10 in total. The bat costs a dollar more than the ball. How much does the ball cost? ____ cents	a) A sandwich and chips cost £1.10 in total. The sandwich costs £1 more than the chips. How much do the chips cost?
If it takes 5 machines 5 min to make 5 widgets, how long would it take 100 machines to make 100 widgets? ____ min	b) If it takes five printers five minutes to print five books, how long does it take 100 printers to print 100 books?
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	c) On a hill, there is a bunch of lily flowers. Every day, the bunch doubles in size. If it takes 48 days for the bunch of lilies to cover the entire hill, how long would it take for the bunch to cover half of the hill?

Conditional probability problem – *example:* This involves the description of a conditional probability event (Jacobson and Petrie 2009). "For a woman at the age of 40 who participates in a routine screening test the probability of breast cancer is 1%. If a woman does not have breast cancer, the probability is 10% that she will still have a positive mammogram. Imagine a woman from this age group with a positive mammogram. What is the probability that she has breast cancer?"

This probability problem can be solved using Baye's theorem, and the answer is approximately 0.78%. However, when asked in 100 physicians, 95% estimated to be around 75% (Eddy 1982). According to Eddy (1982), the example of mammography problem illustrates the complexity of the medical decisions and how the probabilities could be misinterpreted. The reason for choosing this probability problem as an example to familiarise the participants with reasoning procedures is because it has been used in a number of studies to investigate other ways of presenting such complex probability problems (Hoffrage and Gigerenzer 1998; Gigerenzer and Hoffrage 1995; Gigerenzer and Edwards 2003). Hence, it was the ideal training example for this exploratory study that investigates interactive weather reasoning would help people choose 'correctly'.

Therefore, the participants after reading the problem, they were asked to guess the probability answer. To convey their estimation of probability, participants described it as high, selecting a number between 70% - 90% or extremely low selecting numbers from 1% and lower. Following their guess, they were asked to visualise the problem in a way that it would make it easier to understand. Participants were encouraged to use any reasoning process they would prefer from numbers and words to images to describe their way of probability presentation.

Holt and Laury lottery option: This probability problem is one lottery option from the Multiple Price List developed by Holt and Laury (Holt and Laury 2002). The Holt and Laury standard version comprises of two options in ten different rows. The probabilities for the higher amounts are 10% and 90% for the lower amounts, and the probabilities change from row to row while the payoff remains the same. Hence, the expectation values of the two options changes in each row. In the first four rows, the expected value for option A is safer, and option B is riskier. Form the fifth row, and below the risky option, B has higher expected value. If participants are consistent with their choices, they change after some point from option A to option B. The time that they switch over, determine their risk attitude. All rows are presented at once to the participants, and they are asked in each row to decide which option they would prefer. For this experiment, the row that differentiates risk takers from risk neutrals was chosen (5th row) as a measure of people's reasoning.

This measure was not used as a risk assessment; hence it was only used one option out of 10. It was used to identify whether any of the other parameters such as education, numeracy, cognitive thinking style or external representations would be able to predict the rational evaluation of option. Therefore, the two lottery options along with the question to participants were formed as follows:

Could you please indicate which one of the two lotteries is the safest?

Lottery A: 40% of £40 and 60% of £32 ***or***

Lottery B: 40% of £77 and 60% of £2.

Data Collection – Procedure

The participants were given the participant information sheet informing them about the study and the informed consent form to sign. After the informed consents had been obtained,

participants were interviewed filled in in the surveys using pen and paper. On average, each participant needed thirty minutes to complete all of the questionnaires.
When participants were provided with the conditional probability problem, they were instructed to answer by using either numbers or visuals or whatever it was more appropriate to them. When they spent more than the normal time reading the problem, they were asked: “Do you understand the problem?” and then they were encouraged to think the way that makes them understand the data to find the answer. If a person replied that visuals would help them to understand the problem better, he was encouraged to draw. The intervention from the interviewer was only in the conditional probability problem. After they had completed the conditional problem, they were given the Holt and Laury option where they were asked to repeat the procedure. After completing the questionnaires, participants were thanked for their participation. Participants did not receive compensation for their involvement in this study.

Data-Analysis

As it has been supported by Grinnell and Unrau (2010), an empirical research can be considered as ‘ideal’ when independent variables can be manipulated to determine whether a causal relationship between them and dependent variables exist. For this study, the education, the numeracy level, the cognitive style and external representations act as independent variables that form the participants’ underlying thinking process to evaluate a Holt and Laury option. The dependent variable is people’s rational evaluation of probabilities, which is measured by whether they answered correctly in the Holt and Laury question. Figure 13 shows the overview of the variables within this experiment and the relationship between them.

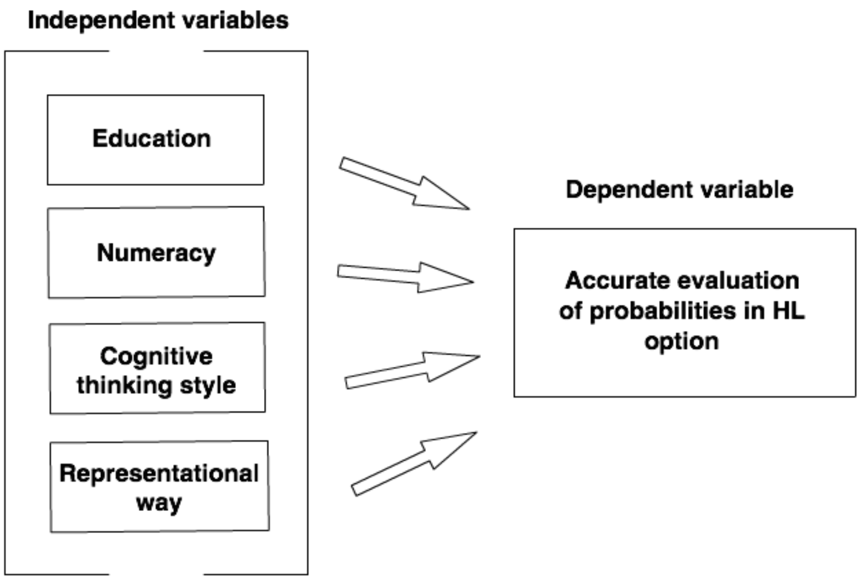


Figure 13. The relationship between independent variables and the dependent variable

A quantitative methodology was employed to analyse the external representations that participants used. Hence, a coding system needed to be used to categorise the reasoning ways. The coding of the external representations was based on the coding of written protocols adapted and adjusted from previous research studied in the area of problem-solving procedures (Corter and Zahner 2007; Russell 2000; Zahner and Corter 2010). Therefore, a table was created that included all the types of representations that were used (if any) by the participants. The identified types were numbers, graphs, pictures, non-diagrammatic, and we added the blank page that it was not included in the coding method that Corter and Zahner (2007) used. Each representation was coded with one according to the above-mentioned types. For instance, if a participant approached the problem solution using numbers, pictures and words then these types were coded with 1 and the rest graphs and the blank page with zero. The mentioned types are defined below.

Types of external representations

Numbers: When there was a mathematical approach to the problem, such as listing numbers, doing calculations or any numerical presentation of the procedure is coded as 1. Note: just the answer in the problem in the form of numbers is not coded as numbers (Figure 15).

Graph: Any use of a classic graphical representation such as Venn diagram, pie charts, bar graphs, line chart, existential, line graph, tree graph or infographics is coded as a graph to represent the approaching way (Figure 16).

Figure 14. Pictures

Figure 15. Numbers

Figure 16. Graph

Figure 17. Non-diagrammatic

Picture: When there is a real-world representation of the situation described in the problem in a pictorial way. For example, in a problem about collecting stamps, any pictorial representation of a stamp would count as picture (Figure 14), or if they use any other picture e.g. a circle, or a square without being in a graph then it is coded as pictures.

Non-diagrammatic: If there is any involvement of words or text representation of the problem and the solution procedure is coded as non-diagrammatic. Also, if there is a verbal description of the problem-solving procedure to the experimenter and there is a note below of the problem in quotations (or not), for example "numbers" is also coded as words.

Blank page: When the participants list the answer without any explanation of the problem - solving procedure (verbal or visual or mathematical) or just referring to the way (e.g. numbers) and when there is nothing on the page is coded as a blank page.

To assess the reliability of coding of participants' reasoning ways, another external examiner was used apart from the principal examiner to compute the inter-reliability.

Ethical Consideration

The three basic principles related to research that involves human participants are: Respect for persons, beneficence and Justice (Beauchamp 2008). The principal of respect for others indicates that individuals should be respected for dignity and fidelity, which is connected closely with anonymity and confidentiality. Therefore, in this study, it was ensured that participants' identity was not possible to be associated with personal responses. Participants got informed about the research and had all the required information to decide as to take part or not. Informed consent was obtained from all participants. Additionally, in respect of the beneficiary principle, the potential risks of causing distress to the participants due to the nature of the question, it was explained to them in the Participant Information Sheet they received. They were also informed that they were free to leave the procedure during the process. An importance part of respect of others takes into account persons with diminished autonomy e.g. prisoners, children, students, etc. that should be protected in the sense that they are not forced to participate. Therefore, in this study where also students (above 18 years old) participated, it was regularly reminded to them that they were free to go whenever they wanted without prior notice. However, there was 0% dropout rate since all of the students that were anticipated participated to the study. Finally, the principle of justice refers to the duty of the investigator to maximise the possible benefits and to minimise possible risks that might result from his study, as well as that he should be fair to the participants. According to Gajjar (2013) who addressed the importance of following ethical guidelines to interpret and report data and outcomes, the researcher followed the recommended guidelines. Hence, data and categories when needed (in the questionnaires) were reviewed by a second coder and were verified.

Data Analysis

The statistical software used to analyse the data was SPSS. A chi square test of independence was used to identify whether participants' educational level influenced the accurate evaluation of the option in Holt and Laury task as it is the most appropriate test to determine if there is a significant relationship between two categorical variables. A chi-square test of independence was performed to examine whether participants who scored higher in the validated numeracy scale would be more likely to answer rationally to the question related to Holt and Laury option. This statistical test would determine whether there is a significant relationship between these two variables. The Numeracy scale scores were categorical (0,1), since the participants were categorised as more numerate those who scored 9-10-11 in the scale and as less numeracy those who scored 8 and below according to Peters et al. (2006). For small sample size, where the number of expected data in each category is less than 5, Fisher's exact test is which is more accurate compared to chi-square test of independence. For this reason, Fisher's exact test was used to explore whether participants' score in the cognitive reflection tasks would be associated with their rational decision making regarding Holt and Laury option. A chi-square test of independence was performed to test the hypothesis whether Group 1 would use different external representations to reason compared to Group 2., a logistic regression was also performed to investigate if using any specific way of external representations would be more likely to predict a rational answer. Logistic regression was the appropriate analysis to conduct when the dependent variable is dichotomous and it can explain the relationship whether the outcome variable (rational choice in Holt and Laury option) – which is coded as 0 and 1 – could be predicted based on the independent variable (external representations) (Appendix 10).

3.3.3 Experiment 3

Following from the results of Experiment 2, where using external representation and more specifically graphs were shown to help participants to reason rationally the Holt and Laury option, Experiment 3 was designed to examine whether by prompting participants to draw graphs (pie charts) to reason the lotteries in the Holt and Laury task, would reduce significantly their irrational choices. The purpose of this experiment was to examine whether an interactive pie chart approach of the lotteries in Holt and Laury task would reduce participants' irrational choices compared to the numerical or passive graphical approach of the lotteries.

The results from the exploratory study revealed that numeracy, cognitive thinking style and the external representational way people use to reason a lottery could influence their performance. Using external representations can help participants to evaluate the probabilities of a lottery rationally. According to the results reported in Chapter 4, it was found that using graphs can predict rational decision making in the Holt and Laury option. In Chapter 4 at the discussion part, there is an interpretation of the outcomes which are further analysed in Chapter 5.

Empirical studies have shown that using pie charts to represent the probabilities in the lottery task of Holt and Laury task has the potential of reducing people's inconsistent answers and increase the accuracy of the tasks' risk predictions (Habib et al. 2016; Eckel, Warnick and Johnson 2005). However, past research has shown that sometimes pie charts were considered problematic for people. On another note, Hegartly and Kozhevnikov (1999) identified that when participants interact with graphs, instead of the passive presentation of them in such problems, they have a higher rate of success. In addition to that, literature on visual thinking supports that interacting with graphs while solving mathematical problems lead to successful solutions (Zahner and Corter, 2002). Therefore, it was hypothesised that interacting with graphs to reason the options in Holt and Laury task will help participants with low numerical skills to understand better the lotteries and be more consistent and rational in their choices compared to numerical or passive graphical display of them. The methodology of the experiment is described in the sections that follow.

3.3.3.1 Method of the interactive pie chart approach

The method of this study aims to answer the research sub-question: Does the use of interactive pie charts influence participants to choose rationally in the risk lottery task? To address this research question, a similar descriptive research design was employed. Descriptive comparison design aims to detect any differences in people's irrational choices between the different display formats. Therefore, a between-subject design was conducted where three groups exist; Participants were randomly assigned to groups to fill in the textual Holt and Laury or either the Holt and Laury displayed with pie charts or the interactive Holt and Laury. Hence, Group 1 received the textual Holt and Laury task; Group 2 received Holt and Laury task displayed with pie charts and Group 3 the interactive Holt and Laury task.

Participants

Students from Faculty of Computer and Engineering in Coventry University, UK, were invited to participate in the experiment via an email list. In total 225 participants filled in the tasks and the questionnaires. Participants in this study included 66 females and 159 males (M age = 29, age range between 18 – 32). There existed three groups of different display formats of Holt and Laury task, and the allocation of the participant in the respective format was randomised. From the beginning of the experiment, it was communicated to participants that when they finish, they would take part in a lottery where one of them would win an Amazon Voucher of £50.

Materials

As described earlier, the materials include the standard Holt and Laury task version, the pie charts Holt and Laury task version, and the interactive pie chart Holt and Laury task version. Hereby a description of all materials is presented:

The Holt and Laury standard version comprises of two options in ten different rows. All rows are presented at once to the participants and they are asked in each row to decide which option they would prefer (see Appendix 3). The payoff amounts for option A are £2 and £1.60 and for option B £3.85 and £0.10. The probabilities for the higher amounts are 10% and 90% for the lower amounts and the probabilities change from row to row when the payoff remain the same. Hence, the expectation value of the two options changes in each row. In the first four rows, the expected value for option A is safer and option B is riskier. Form the fifth row and below the risky option B has higher expected value compared to option B.

If an individual is consistent, he will change after some point to option B, the time that he will switch over determines his risk attitude. The risk-utility function can be calculated by the formula: $U(x) = x^{(1-r)}/(1-r)$, where x is the lottery pay out and r is the risk aversion coefficient for each row (Holt and Laury 2002). The value in the Holt and Laury that represents the safe choices is between 0-3 (HL value <4), for a risk neutral choice the value is 4 and for a risk choice above 4 (HL value>4). A negative CRRA value shows a risk seeking behaviour while a positive CRRA value shows a risk averse one.

The other material involves the **Holt and Laury task displayed with pie charts** (see Appendix 4) where every option between the two lotteries is represented in a pie chart, and there is a text describing the proportions about the payoffs on the pie chart instead of a completely textual

Holt and Laury. The pie charts would be displayed similarly to the full graphical display with the payoff and the probability in text next to the graph, as it was presented in Eckel and Grossman study (2009). Next, to the pie chart, the relevant payoff is displayed textually.

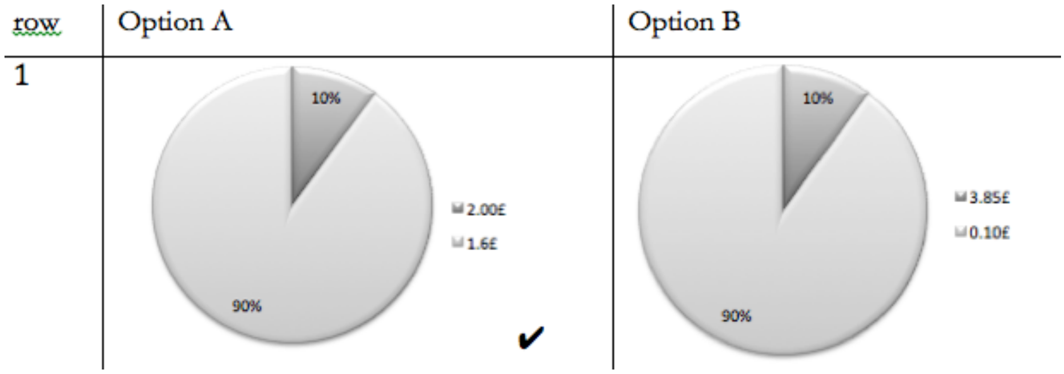


Figure 18. Holt and Laury displayed with pie charts

The stimuli developed for this study, involves the pie charts divided in 10 pieces with the proportions of each probability option of the lottery presented textually above them in order for people to shadow the pieces on the pie chart according to the probabilities outlined above and the mark also the pay offs of those pieces. After they drew the pie charts they would choose one lottery (Figure 19).

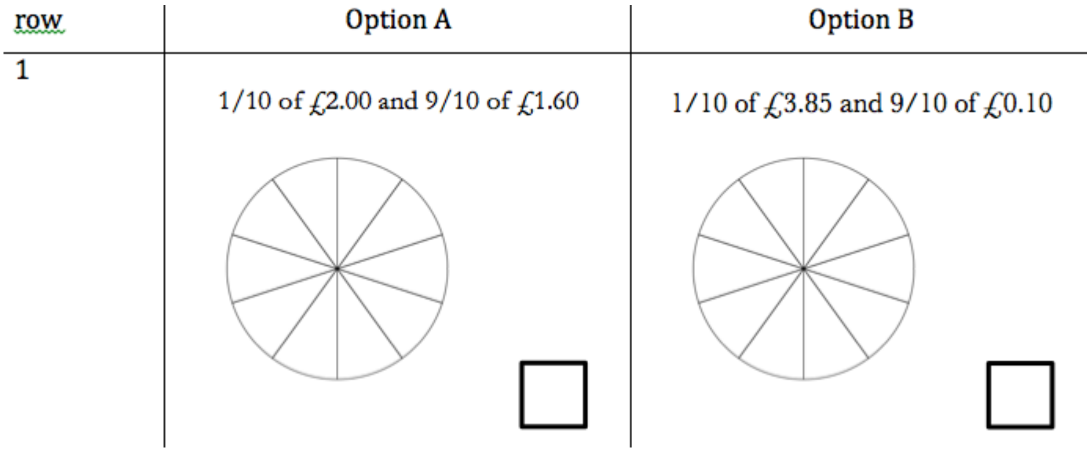


Figure 19. The empty circles where the participants will draw in themselves the proportions for each pie chart and they choose the one they prefer.

The numeracy test used is the Lipkus Numeracy scale (Lipkus, Samsa and Rimer, 2001) that was used in Experiment 2 and described in Section 3.3.2.1.

A couple of **demographic** questions regarding participant's age and gender and a **self-report of difficulty task** were also involved in the experimental procedure. The question

was formed as follows: On a scale of 1-5 with one being very easy and five being very difficult, how difficult was this lottery task for you?
The participants were asked to answer in a 5-pont Likert scale where one was very easy, two was barely difficult, three was somewhat difficult, four was mostly difficult, and five was very difficult.

Procedure

The experiment took part in a lecturing room in the Faculty of Computer and Engineering and lasted approximately 30 minutes. It was divided into three stages. In the first stage, participants were given the participant information sheet informing them about the study and the informed consent form to sign. In the second stage, after the informed consents were obtained, participants were given an envelope that included one of the three display formats of Holt and Laury task with instructions and an example, a numerical test and demographic questions about their age and gender. The participants were asked to follow the order that the brochures were inside the envelope. Thus, the Holt and Laury task had to be filled in first, then the numerical test and last the demographics. The order was strict so the participants would not feel tired when filling in the Holt and Laury task. On average, each participant needed 25 minutes to complete the experiment. Since the envelopes with the different display formats were mixed, the participants were randomly assigned to fill either the textual Holt and Laury task or Holt and Laury task displayed with pie charts or the interactive Holt and Laury task. Hence, Group 1 received the textual Holt and Laury task, Group 2 the Holt and Laury task displayed with pie charts and Group 3 received the textual Holt and Laury task and the interactive Holt and Laury method. Instructions were presented on the first page of each task along with an example. At the final stage, after all, participants completed the experiment, they were asked to put the documents back to the envelopes. The envelopes were collected, and each participant selected a small note from a lottery ball, all notes included numbers except from one that had the letter A and was referring to an Amazon Voucher of £50.

Ethical considerations

For this study, the three principles related to research that involve human participants were considered (Beauchamp 2008). The principal of respect was ensured by informing the participants about the research study, including any potential risk or benefits through the participant information sheets shared with them. The participants also had opportunities to ask questions to the investigator. An importance part of respect of others takes into account persons with diminished autonomy e.g. prisoners, children, students, etc. that should be protected in the sense that they are not forced to participate. Therefore, in our study where university students (above 18 years old) participated, it was regularly reminded to them that they were free to go whenever they wanted without prior notice. Participants they were reminded throughout each study that their participation was voluntary.

The principles of beneficiary and justice were considered by informing participants about possible risks such as distress from the questions. Another ethical principle that we would consider in our studies is also the principle of autonomy, which refers to self-determination and the right to full disclosure. Therefore, the participants were fully informed about the procedure, nature and they had the opportunity to decide whether they wanted or not to the study. Informed consents were obtained from all participants ensuring that they have enough information and freedom to consent or to refuse to participate. It was also ensured that confidentiality was implanted in all procedures. Any names were not linked to specific responses. The names and any other personal information were not mentioned in the research report in the sense that they could be linked to the specific person. The main researcher kept all the documentation locked in a safe desk. These documents will be destroyed after five years since the end of the study.

Data Analysis

In the proposed study, three measures were employed to measure one dependent variable, three independent variables and three moderator variables. These are outlined as follows:

Independent variable: The different displays in Holt and Laury task, which were correlated with people rational choices via the three Holt and Laury tasks (Appendix 11).

Dependent variable: People's inconsistent answers on Holt and Laury task at the three-different displayed Holt and Laury tasks (Appendix 11).

Moderator variables: In addition to the independent and the dependent variable, there were secondary independent variables or moderators that were considered. These include the numeracy level, the demographic, gender and age. The moderator variables can determine the extent to which the relationship between the variables is influenced by secondary factors.

Mediator: The self-report of difficulty task was considered as a mediator.

Initially, the rates of people irrational choices with each different format were calculated. Participants who chose only option A, or only option B or switched more than once in all display formats, they are classified as irrational and inconsistent. All participants were coded to two groups regarding their consistency rate. For instance, Group 1 included participants that were consistent and Group 2 participants that were inconsistent. Therefore, all the inconsistent answers were summed up to provide the inconsistent rate of each different display format of Holt and Laury task. Then McNemar's test was used to determine whether there were any statistical significant differences between interactive pie chart approach and numerical Holt and Laury, interactive pie chart approach and pie chart approach, and pie chart and numerical Holt and Laury, inconsistency rates. McNemar's statistical test was used, as it was also used by Bauremester and Musshoff (2016) who undertook similar methodological approach to determine inconsistency rates between two risk elicitation tasks that were performed by two different groups. McNemar's is the most appropriate statistical test because it can determine whether there is a significant difference in proportions between participants inconsistent choices in standard Holt and Laury numerical display and pie chart display and interactive pie chart Holt and Laury. A binomial logistic regression was performed to understand whether participants' inconsistent choices in three different display formats could be predicted based on gender, age, difficulty, or numeracy level differences. The latter logistic regression was used to exclude the effects of other variables on participants' choices. Binomial logistic regression is the most appropriate analysis when the dependent variable which in this study is participants' inconsistent choices, is dichotomous (binary). It is a predictive analysis, used to understand the relationship between a binary variable and one or more nominal, ordinal, interval or ratio-level independent variables.

3.4 Summary

This chapter analysed the research philosophy and methodology undertaken as well as the methods of each research study. It explained that this research used an empirical approach, which includes a series of three studies that investigate the interactions between humans and the instrument of a risk elicitation task to assist in the rational decision making that increases the tasks' predictive value. Based on past empirical evidence a high-level hypothesis was formulated whether visuals incorporated in a risk elicitation task would influence rational decision making. Therefore, this thesis involves three studies that observe differences between conditions. These conditions are the following:

- Experiment 1: Whether introduction of gamified elements into the risk elicitation task assisted in rational decision-making.
- Experiment 2: The most profound factors that influence rational decision making in risk elicitation tasks and whether using external representations can help.
- Experiment 3: Whether interactive pie charts used as external representations can influence rational decision-making.

All three experiments were designed to lead to the answer of the high-level research question whether gamified and interactive media could assist in rational decision making of a risk elicitation task.

The methods for the first experiment involve a usability study to ensure that there were no usability issues which affect participants performance and a construct validity study that examines whether the task measure the concept that is supposed to measure. The experiment was investigated whether gamified elements assist in rational choices by the participants. The methods of the second experiment involved a quantitative study that explored the most profound factors that influence rational decision making in the risk elicitation task among numeracy, cognitive thinking style and education while investigating the use of external representations whether they would help participants in rational decision making. The third experiment referred to a quantitative study, which explored whether the use of interactive pie charts would assist in rational decision making in the risk elicitation task.

In each experimental procedure, the participants, the materials, the procedures and the data analysis were established. Additionally, the ethical considerations were presented with a justification that the guidelines were followed. This chapter aimed to justify the methodology and the methods that were implemented to answer the high-level research question whether gamified and interactive media would assist rational decision making in a risk elicitation task.

Next chapter describes the statistical analyses and the results of each experimental study. It also involves a brief conclusion of each study that is further elaborated in Chapter 5 where there is a discussion of the results in relation to empirical findings.

4. IMPLEMENTATION AND RESULTS

4.1 Introduction

This chapter presents the statistical analysis and the results of the three experiments. The methodology of each experiment was presented in Chapter 3. Study results' corresponding to each sub-hypothesis that was tested, are presented and discussed in three sections. The first chapter section describes the results of Experiment 1, outlining first the research sub-question linked to the empirical evidence presented in Chapter 2. Following the presentation of results, there is a brief summary and conclusions related to first experiment results that is further discussed in Chapter 5. The second chapter section presents the research sub-question linked to the empirical evidence presented in Chapter 2, that lead to Experiment 2. Following the results of Experiment 2, there is also a summary of conclusion results. The third chapter section presents the third research sub-question and hypothesis linked to the empirical evidence presented in Chapter 2. Then, the results of Experiment 3 and a brief conclusion and summary follows. Further discussion and reflection on results are presented in Chapter 5.

4.2 Experiment 1

This chapter presents the results of the first experiment that is looking to answer the following research sub-question.

RQ_{01}: Does the introduction of game-based elements into a risk lottery task influence rational decision making in a risk elicitation task? Therefore, this section describes, the construct validity study of a gamified approach, which in the scope of this thesis is defined as the degree to which the Risk House game measures risk propensity. The Risk House game was based on Holt and Laury task, and it was developed to improve participants' understanding of the lotteries and reduce their irrational choices. This section briefly presents the empirical evidence that leads to the testable hypothesis.

In the review of risk elicitation tasks, presented in Chapter 2 Section 2.2.4 (Risk elicitation tasks requirements and considerations) the main elements that considered important for a risk assessment tool were summarized including the needs for simple and accurate tasks. In relation to those important elements of risk assessment tools and the empirical evidence in Section 2.3.1 (Overview of gamified assessment tools), where specific benefits of games were highlighted including the use of visuals that simplify complicated procedures, the research sub-question for this experiment was formulated. In the scope of this thesis only the aspect of visualisation in games that simplify the procedure to reduce participants' irrational choices was examined. The rest of the game elements such as no need for incentives and distribution through smartphone are out of the scope of this thesis.

To examine whether the introduction of game-based elements into a risk elicitation task compromise the construct validity of the task, the risk indication of the gamified approach was correlated with the risk indication of preferences in a validated risk construct. This correlation between the two tasks examines the game's construct validity as an assessment tool. Therefore, the hypothesis was formulated according to the evidence from the literature, and this is the following:

Null Hypothesis (H_0): The choices of participants in gamified risk assessment risk assessment, Risk House game, would not significantly correlate with their choices at the Balloon Analogue Risk Task.

H_{01}: The choices of participants in gamified risk assessment, Risk House game, would be significantly correlated with their risky choices at the Balloon Analogue Risk Task.

The variables that evaluate this hypothesis are participants' choices in the game (Risk choices) and participants' choices in BART (Pumps in total and explosions) (Appendix 9). The correlation is a statistical technique which can describe the strength and the direction of the relationship between the two variables. In this study, the validation of game is measured by the significance of relationship between participants' choices in the game and the task. The alternative hypothesis states that there will be a positive correlation between the two variables: if one variable increases, the other tends to increase. In this study, a positive correlation would show that if someone is risky in the game, he will be risky in BART as well. For this reason, correlation is the most appropriate statistical test for the study.

The objectives of this section are as follows:

- Analysis of the results.

- Conclusions to discuss briefly the limitations of the studies and the next steps.

4.2.1 Introduction of Gamified approach results

The validation of Risk House game represents the main step of establishing the gamified approach as a behavioural task that can be used for risk assessment. Validity is defined as "the degree to which measures what it claims or purports to be measuring." (Forth et al. 1996). In the scope of this thesis and experiment, construct validity is examined to determine that the Risk House game measures individual's risk propensity. Therefore, the following section presents the results of the correlation between the riskiness score on the game and the riskiness on the validated computerised behavioural task called BART. To make precise predictions about people's behaviour in similar circumstances, researchers and policy makers need experimentally validated risk elicitation tasks. Experimental validation of a tool is referred to the examination of the effectiveness of the measures (Cronbach and Meehl 1995). This is achieved by statistical analysis between responses to different test items (Cronbach and Meehl 1995). However, there are not adequate evidence for most of the measures developed so far that they have been validated. Thus, these tools produce problems and inaccurate results. These measures to be reliable should be reproducible with a small measurement error (Wolbert and Riedl 2013) while they have to show the exact level of someone's risk aversion, being validated about other measures (Rohrmann 2005).

This experiment's contribution is mainly methodological, as it investigates another method of assessment that could help participants to choose rationally in the task. It is hypothesised that the game will be correlated significantly with the validated task. To test the hypothesis outlined previously, Section 3.3.1.2 of Chapter 3 presents the method used to measure the games construct validity. Participants' performance on the validated risk elicitation task BART was correlated with their performance on Risk House game. The results are presented in the following section.

4.2.3 Results

As it was described in 3.3.1.2 Section of Chapter 3, 32 participants played the Risk House game and performed the BART task. Participants' performances in both constructs were correlated to identify whether a significant association exists between their choices in each task.

The number of risky choices from the Risk House game and the number of adjusted pumps in the BART game was collected and analysed to calculate the correlation and identify whether a relationship could be observed between two assessments. The variables examined from the BART task were the adjusted number of pumps in total as the index of riskiness, and some explosions. Both of them produced almost identical results. Correlation analysis indicated that there is no significant relation between the number of explosions and the numbers of risky choices that the participants made in the Risk House Game, $r (32) = .120, p< .511$. Similarly, there were no significant results in the correlation between the pumps in total in BART and the risky choices from the Risk House Game, $r (32) = .158, p < .389$. These results are presented on the scatterplots (Figure 20 and Figure 21), where a visually observable trend can be seen that may indicate a relationship between two variables (Figure 20: Risky choices and pumps in total, Figure 21: Risky choices and explosions), however the correlation is not statistical significant at $p < 0.05$ level.

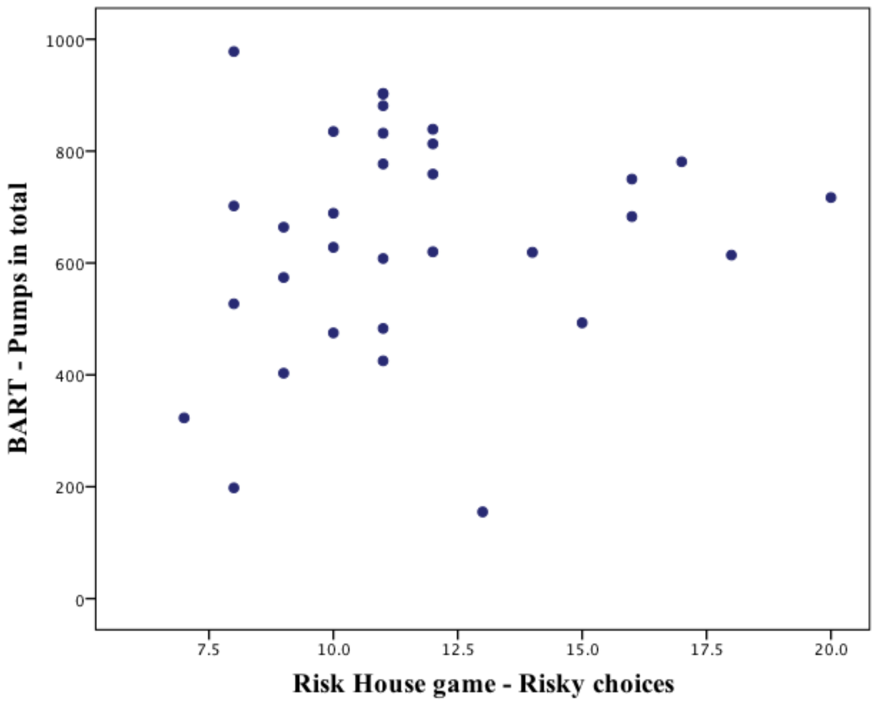


Figure 20. Correlation between the pumps in total in the BART task and participants' risky choices in the game (Appendix 9)

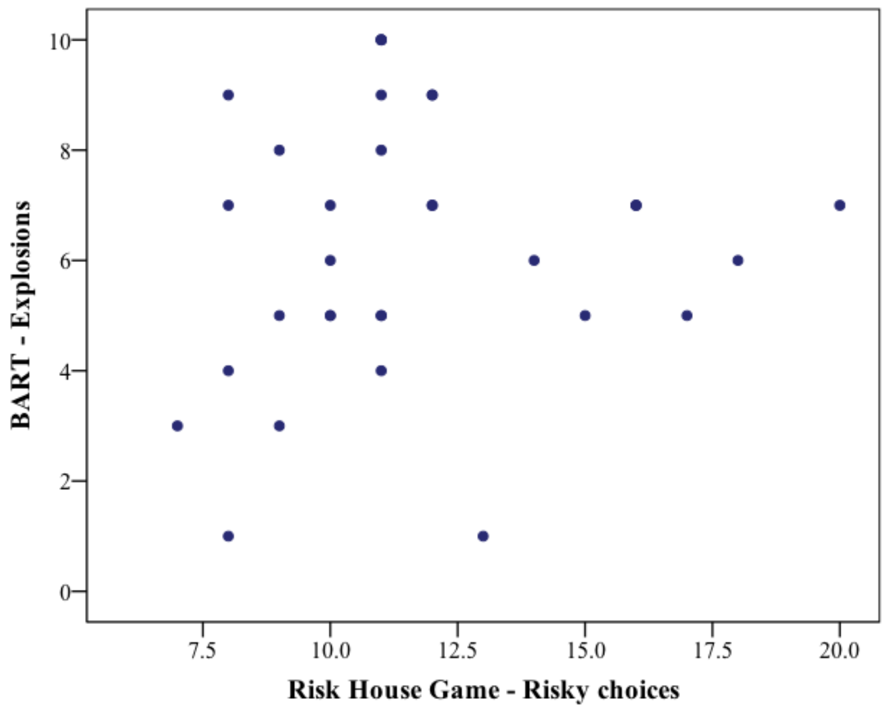


Figure 21. Correlation between the explosions in the BART task and participants' risky choices in the game (Appendix 9)

Additionally, it was noticed that there was a high rate of inconsistency in people's choices in the Risk House game. The consistency level of people's choices in the Risk House game was monitored. It was identified that 16 out of 32 participants chose irrationally in the game, 7 chose only roulette A and 9 moves from roulette A to roulette B constantly.

4.2.4 Summary and Conclusions

This chapter presented the construct validity study of a gamified approach called Risk House game, developed to elicit risk propensity. The usability satisfaction study that was presented in 3.3.1.1 Section in 3rd Chapter showed that the gamified approach was usable by non-gamers. It also highlighted a few areas of improvement on the size of the buttons and "click" areas that were improved before the construct validation study. The construct validity study, which examined the relations between a behavioural task and the gamified approach that both assess risk preferences, showed no significant results. Specifically, the hypothesis that the choices of participants in gamified risk assessment, Risk House game, would be consistent was rejects. However, the hypothesis that participants' choices in the task would correlate significantly with their choices in BART could not be rejected either the null hypothesis approved due to the noisy data in the game. The noisy results in Risk House game produced by people's inconsistent choices showed that the visualisation of the numerical lotteries did not help people to understand them and therefore choose rationally. As it has been demonstrated by previous empirical studies, the pen and paper Holt and Laury tasks produce a lot of noisy results (Dave et al. 2010). As such, people's inconsistent choices demand further exploration, beyond implementation of the task into gamified context. An exploratory study could inform the design of the gamified approach to achieve maximum performance. To that end, the next section describes an exploratory study that investigates the factors, which influence people's choices in such options and an alternative way of people's reasoning process. Further discussion and association of the results from Experiment 1 with the high-level research question, is presented in Chapter 5.

4.3 Experiment 2

This chapter presents the results of the second experiment, what factors influence rational decision making in risk elicitation task. The behavioural experiment that examined the construct validity of the gamified approach, Risk House game, revealed high irrationality rate on people's choices. This finding aligns with findings from empirical studies with Holt and Laury task that showed high inconsistency in people's choices (Hill and Viceisza 2012). Therefore, a step back is required to identify the factors, which influence people's inconsistent choices in these tasks that will assist in the development of a gamified approach. This chapter presents the results of the exploratory study. This section briefly presents the empirical evidence that builds up the testable hypothesis.

Empirical research, examining people's irrational choices in risk elicitation tasks, highlights factors that may affect people's inability to choose rationally, as it has been shown in Chapter 2, Section 2.4.1 (Factors that influence people' rational evaluation of choices in lottery tasks). According to the evidence presented in Section 2.4.1 in Chapter 2 and evidence presented in Section 2.4.4 regarding the previous use of external representations, the research sub-question was formulated as follows.

RQ_{02}: What factors influence participants' rational choice in a risk lottery task?

To answer this research question, an exploratory study was conducted investigating the effect of numeracy level, educational level, cognitive thinking style and external representations on evaluating two lottery options from the Holt and Laury task. According to the evidence from literature, our hypothesis was formulated as follows:

Null Hypothesis H_0: Participants' education, score in the validated numeracy scale, cognitive reflection task and use of external representation would not influence their rational decision making in the Holt and Laury option.

H_{02}: Participants' education, score in the validated numeracy scale, score in cognitive reflection task and use of external representation would influence their rational decision making in the Holt and Laury option.

The independent variables that evaluate this hypothesis are education, score in numeracy scale, score in the cognitive reflection task and external representations (Appendix 10). The dependent variable is participants' decision in Holt and Laury option (Appendix 10).

This section describes the results of the exploratory study.

This chapter's objectives are as follows:

- Results from the computation of inter-rater reliability
- Analysis of the results
- Conclusion to discuss

4.3.1 Computation of Inter-rater reliability for the external representations

Since 60 participants used external representations to reason their choice in the Holt and Laury option, these inscriptions were coded according to the 3.3.2.1 Section of the Research Methods Chapter, data analysis and the coding of written protocols of Corter and Zahner (2007). The principal investigator of the study performed the initial coding of the external representations. However, to assess the reliability of the initial coding, a second rater was trained and coded all participants' representations. The 2nd rater was selected from Coventry University from a different department and was not aware of the research questions, and she had previous experience with research projects.

The instructions of the coding system (see Chapter 3, Section 3.3.2.1, and Types of Representations) were given to the second rater, along with the main researcher's notes from the interviews and examples of each category. When the second rater felt comfortable enough to start the procedure, started categorising the presentations in order from 1 to 60 in an Excel Spread Sheet. The following section presents the results of the computation of inter-reliability.

4.3.2 Results of the Inter- rater reliability

When the coding on the external representations was also completed from the second examiner, Cohen's kappa was run to determine if there was agreement between the two raters' categorisation of them. Table 7 shows the initial Cohen's kappa agreement for each one of the five categories in the Holt and Laury task. Based on the guidelines from Landis and Koch (1977), a kappa of .957 and .977 represents almost absolute agreement between the two examiners.

Table 7. Kappa statistics for the external representations for the Holt and Laury option

Category	Kappa (k)	Interpretation (**Landis and Koch 1977)**
Numbers	*1*	*Very good*
Graphs	*.792*	*Good*
Pictures	*.762*	*Good*
Non-diagrammatic	*.798*	*Good*
BlankPage	*.900*	*Very Good*

The two raters had to discuss the remaining discrepancies and agreed upon the coding. The Cohen's kappa was recalculated as seen in Table 8.

Table 8. Kappa statistics for the external representations for the Holt and Laury option (second round)

Category	Kappa (k)	Interpretation (**Landis and Koch 1977)**
Numbers	*1*	*Very good*
Graphs	*1*	*Good*
Pictures	*.889*	*Good*
Non-diagrammatic	*.865*	*Good*
BlankPage	*.900*	*Very Good*

The resulting consensus coding was used in all analysis reported.

4.3.3 Results

For this study, 25 participants from Group 2 and 35 participants from Group 1 were invited to fill in a numeracy scale, a cognitive reflection test, one example probability problem and a Holt and Laury lottery option. Therefore, the independent variables were the education, numeracy, cognitive thinking style and external representational way. The dependent variable is the rational evaluation of the probabilities in the Holt and Laury option. To test the hypothesis that participants who had more years of schooling would be more likely to answer rationally in the Holt and Laury option a chi-square test of independence was performed. The chi square test of independence showed that participants who graduated from university were not more likely to answer rationally in the Holt and Laury option compared to participants who had not graduated from university, X^2 (1, N=60) = .429, p = .513. To determine whether participants who scored lower in the validated numeracy scale, were more likely to answer irrationally to the Holt and Laury option, a chi square test of independence was performed. Numeracy performance was divided into two groups, those participants who were scored higher (9-10-11 correct) and those that scored less (2-8 items correct). This dichotomous split has previously performed by Peters et al. (2008). Therefore, the data were binary (0 for most numerate and 1 for less numerate). The chi square was statistically significant X^2 (1, N= 60) = 4.176, p = .041, showing that people who scored higher in the validated numeracy scale have a greater chance of answering rationally in the Holt and Laury option compared to those who scored lower. A chi square test of independence was also performed to find out whether gender is associated with answering rationally or not in the Holt and Laury option. The results showed that there is no gender association with the rational answers the participants gave, X^2 (1, N=60) = .019, p = .890.

To test the hypothesis that those who answered correctly in the cognitive reflective task have a greater chance of answering rationally in Holt and Laury option, Fisher's Exact Test was performed. It revealed that participants who answer correctly more problems in the cognitive reflection task were more likely to answer rationally in the option, $p<.05$.

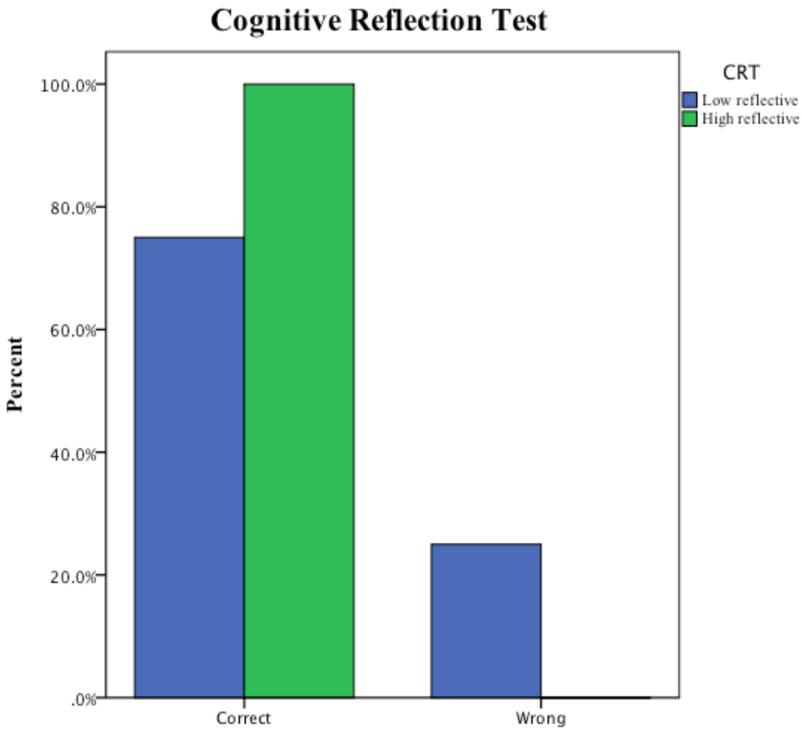


Figure 22. The percentages of high and low scorers in the CRT and the answers they gave in the Holt and Laury option (Appendix 10)

To test the hypothesis that participants who had graduated from university, would be more likely to use different external representations compared to the rest participants, a chi square test of independence was performed which showed that there is no difference among the external representations both groups used, X^2 (4, N= 60) = 3.642, p = .457.

Additionally, to determine whether using any from the external presentations would be able to predict whether participants would answer rationally a logistic regression was performed.

The logistic regression model was statistically significant at $p < .02$ according to the model chi square statistic. The model explained 27.9% (Nagelkerke R^2) of the total variance of the correct answers that the participants gave. Nagelkerke R^2 is the power explanation of the model that is calculated to evaluate the goodness of fit of the logistic regression model. Table 9 shows the statistical tests and significance levels of each of the variables. The overall data are presented in Appendix 10.

Table 9. Statistical significance of the independent variables (coefficients) in the model

Coefficients	Wald	Sig. (p<)
Numbers	6.685	.154
Graphs	4.612	.032
Pictures	.000	.999
Non-diagrammatic	.000	.999
Blank page	6.659	.010

Table 9 shows that using graphs predicts stronger rational decision making. Also, the participants who did not use any of the external representations and were just guessing were more likely to give an irrational answer.

4.2.4 Summary and Conclusions

To summarise, the analysis from the exploratory study revealed the following results. Participants' years of schooling did not influence their rational evaluation of the Holt and Laury option. It was found that participants who scored higher in the numeracy scale and who answered more problems correctly in the cognitive reflection task are more likely to evaluate rationally the Holt and Laury option. Finally, participants who used external representations and more specifically graphs were more likely to evaluate the Holt and Laury option rationally. These findings are consistent with empirical evidence suggesting that numeracy and cognitive thinking plays a role in choosing rationally in lottery risk elicitation tasks such as Holt and Laury (Dave et al. 2010; Frederick 2005). Additionally, it was shown that external representations helped participants to answer the question rationally. These findings align with the literature which supports that external representations aid in the decision-making process with probability (Zahner and Corter 2002), was confirmed. Also, an important finding is that graphs were shown to help participants reasoning rationally in the lottery problem and thus it agrees with the empirical evidence, which shows that graphs can improve the consistency of responses (Eckel, Engle-Warnick and Johnson 2005).

An important limitation of this study should be considered. Participants who had less years of schooling and considered lower educated were English, and this does not let us generalize the results to other populations since there is a difference between educational systems and consideration of an educated individual between western and eastern societies. This is further discussed in Chapter 5.

The findings from this study have implications for the development of a gamified approach. According to the results presented in Table 9, participants who used either graphs had a greater likelihood of evaluating the Holt and Laury option rationally. Therefore, further investigation is required to incorporate an interactive graphical approach in the pen and paper Holt and Laury to disprove the null hypothesis that engagement interactively with graphs before reasoning the choices in Holt and Laury would have no significant effect on rational decision making in the task.

4.4 Experiment 3

This section presents the results of the third experiment that investigates whether interacting with pie charts would influence rational decision making in a risk elicitation task. The section briefly presents the empirical evidence justifying the background of the experiment that leads to the testable hypothesis. Empirical evidence has shown that people have difficulty in understanding Holt and Laury task, hence they tend to do irrational choices (Dave et al. 2010). Researchers who investigate people's irrational choices, use outcome measures e.g. accuracy or consistency of reasoning using graphs to determine that the difference in presentational form can help people understand the probability better. Several empirical studies, which involve lotteries, attempted to use other methods instead of numerical probabilities to present the options to reduce people's irrational choices (Bauremeister and Musshoff 2016; Ihli, Chiputwa and Musshoff 2016). The different research methodologies include asking from people either to estimate a proportion represented in a graph or interpret a part of a graph, as well as to choose between the same risky and safe options in visual and in textual form. However, the way the graph is presented to them can influence their choice massively (Stone et al. 2003). For example, when people are presented with graphs that the numerator of people who have developed the disease of interest, of risk ratio is emphasized; they tend to change their behaviour positively (Nuovo, Melnikow and Chang 2002). Therefore, it is important for a graph to involve all the necessary elements that will make people to choose accurately. Among the variety of graphs that are most usually used in probability presentation problems, part-to-whole bar charts, pie charts and part-to-whole icon areas are able to promote people's proportion judgment and automatic visual processing (Cleveland and McGrill 1986). Although graphs seem to be a more intuitive method than numbers, literature has shown that familiarity, expertise and education affects the accurate interpretation of formats. In empirical studies with Holt and Laury task, the most popular graph that has been used was pie chart, showing that there is a potential of reducing people's inconsistent choices (Habib et al. 2016; Eckel, Warnick and Johnson 2005; Boughera, Grassman and Piet 2011). However, as it was identified by Fieldman-Stewart et al. (2000) and Willkinson and Wills (2005), people have difficulty in estimating the angles or to being accurate about the proportions that pie charts involve. Hence, there are still doubts whether pie charts are the appropriate representation for the majority of people. Literature review on visual thinking, supports that when people engaging in using visuals themselves to reason a probability problem, leads to a fuller understanding depth in processing (Logie & Baddeley, 1990; Mayer, 1989, 2001; Mayer & Gallini, 1990). Therefore, they have higher chances of finding the right solution. According to Hegarty and Kozhevnikov (1999), people who use graphs as external representations have greater chance of success compared to those who use pictures as visuals. Similar finding was shown from the 2nd Experiment, as it was presented earlier; participants who used graphs as external representations had greater chance of success in evaluating the Holt and Laury option rationally. Therefore, the research sub-question this chapter is looking to answer is formulated as follows:

RQ_{03}: Does the use of interactive pie charts influence rational decision making in a risk lottery task?

According to findings from literature that were also presented above, the hypothesis may be stated as follows:

Null Hypothesis (H_0): Individuals that score lower in the validated numeracy scale when they interact with the pie charts in Holt and Laury task would make not significantly more rational choices compared to when they are presented the Holt and Laury task with passive pie charts or numerical probabilities.

H_{03}: Individuals that score lower in the validated numeracy scale when they interact with the pie charts in Holt and Laury task will make significantly more rational choices compared to when they are presented the Holt and Laury task with passive pie charts or numerical probabilities.

The dependent variable in this testable hypothesis is participants' rational choices and the independent variables are the three different display formats are presented with (Numerical, passive pie chart and interactive pie chart) (Appendix 11).

This section presents the experimental results of the methods outlined in Section 3.3.3. It aims to answer the high-level research question of this research: Can gamified and interactive influence rational choices in a risk elicitation task? The outcomes of this third experiment and their implication for the design of a gamified approach will be further discussed in 5th Chapter. The objectives of this section are:

- To present the results
- Conclusions to discuss the limitations of the study briefly

4.4.1 Interactive pie chart approach

The method undertaken to identify whether asking participants to interact with pie charts before choosing the options in the Holt and Laury task would reduce their irrational choices compared to the numerical or passive pie chart version of Holt and Laury task is described in Section 3.3.3 of Chapter 3. The method involves a Holt and Laury task displayed with numbers, a Holt and Laury task displayed with pie charts and a Holt and Laury task where participants have to draw the pie charts reflecting the probabilities of the lotteries themselves before choosing. These three displayed formats of Holt and Laury task were randomly distributed to 225 participants; 76 participants filled in the numerical display, 76 the pie chart display and 73 the interactive pie chart Holt and Laury task. The results are presented in the following section.

4.4.2 Results

According to the classification of participants' irrational choices, 35 (46%) participants showed an inconsistent behaviour in Holt and Laury task numerical format, 27 (35.52%) participants in the pie chart format and only 4 (5.47%) participants in the interactive pie chart Holt and Laury task format. McNemar's test for related samples was applied between the interactive pie chart and numerical Holt and Laury format, which revealed a significant difference in the inconsistency rates between both display formats, $p < .00$. McNemar's test for related samples was also applied between interactive pie chart and pie chart Holt and Laury format which showed a significant difference between their inconsistency rates, $p < .00$. McNemar's test for related samples between the pie chart format and the numerical Holt and Laury format showed no significant difference between their inconsistency rates. Additionally, a binary logistic regression was performed to ascertain whether participants' irrational choices in each display format (numerical, pie chart and interactive pie chart) could be predicted based on their age, gender, the level of their perceived difficulty and numeracy score in the validated scale. For the numerical Holt and Laury format, the binary logistic regression was statistically significant at the .00 level according to the model chi-square statistic. The coefficient on the numerical level had a Wald statistic equal to 12.32 which is significant at the .00 level. The coefficient on the difficulty perception had a Wald statistic equal to 8.84 which is significant at the .003 level. The coefficient on the age had a Wald statistic equal to 4.853 which is significant at the .028 level. The gender had not a significant impact on participants' rational choices in the numerical Holt and Laury. These outcomes suggest that numeracy level, perception of difficulty and age can predict participants' choices in the Holt and Laury numerical task. This model predicts 81.4% of the responses correctly. For the pie chart Holt and Laury, the binary logistic regression showed that the model is not statistical significant. This means that any of the independent variables age, perception of difficulty, numeracy level or gender do not have an impact on participants' irrational choices in the pie chart Holt and Laury display. Finally, the binary logistic regression for interactive pie chart Holt and Laury was statistical significant at the level of .025 according to the model chi-square statistic. The coefficient on the perceived difficulty, had a Wald statistic equal to 5.52 which is significant at the .02 level. The rest dependent variables of age, gender and numerical level are not statistically strong predictors of participants' choices in the interactive Holt and Laury format. These results are not surprising, since the inconsistency level of the interactive pie chart Holt and Laury format was so low (5.47%) which shows that the biggest number despite their age, gender or numeracy score in the validated scale chose rationally. The model explained 33.8% of the responses correctly.

Table 10. Participants' perception on the level of difficulty in each of the Holt and Laury displayed formats (Appendix 11).

Level of difficulty				Groups: Numerical HL	Groups: Pie chart HL	Groups: Interactive HL	Total
Very easy	All groups	Consistent	Count	8	4	29	49
		Inconsistent	Count	0	1	0	1
Barely difficult	All groups	Consistent	Count	15	17	29	62
		Inconsistent	Count	3	6	2	10
Somewhat difficult	All groups	Consistent	Count	7	19	4	30
		Inconsistent	Count	7	10	1	18
Mostly difficult	All groups	Consistent	Count	6	3	1	10
		Inconsistent	Count	12	8	1	20
Very difficult	All groups	Consistent	Count	1	4	2	8
		Inconsistent	Count	14	4	1	17

Table 11. Participants' age groups in three display formats (Appendix 11)

		Age							
		18-22		23-27		28-31		32+	
		All groups		All groups		All groups		All groups	
		Consistent	Inconsistent	Consistent	Inconsistent	Consistent	Inconsistent	Consistent	Inconsistent
		Count	Count	Count	Count	Count	Count	Count	Count
Groups	Numerical HL	25	29	11	5	1	1	3	1
	Pie chart HL	36	23	12	3	0	1	0	0
	Interactive HL	46	5	20	0	0	0	2	0

4.4.3 Summary and Conclusions

In this section, the results of the third experiment investigating whether engaging participants to interact with pie charts, would help them understand the Holt and Laury task and reduce their irrational choices, were presented. In the experiment, 225 participants were given three different displayed formats of Holt and Laury task (numerical, pie chart and interactive pie chart Holt and Laury task). Participants who filled in the interactive pie chart format had significantly less irrational and thus inconsistent choices compared to participants who filled in the pie chart display format or the numerical Holt and Laury task. The McNeman's test showed a statistical difference between inconsistency rates of the interactive pie chart and the numerical or the pie chart Holt and Laury. However, the pie chart and numerical Holt and Laury did not have significant difference in their consistency rates. Age, numeracy scored in the validated scale and the perceived difficulty of the task could predict someone choices in the numerical Holt and Laury. Regarding age this can be explained because the majority of the participants in the numerical Holt and Laury task were from the age group 18-22 (55 out of 76). For the difficulty level, most participants that were consistent perceived the task as very easy or somewhat easy. Regarding the numeracy level, as it has been seen from second experiment and from empirical findings (Dave et al. 2010), individuals' numeracy level has an impact on someone's consistent choices in the Holt and Laury task. For the interactive pie chart Holt and Laury task, only the perceived difficulty was a strong predictor of individuals' choices in the task. This outcome highlights that most of the participants (61 in total) perceived the interactive pie chart display of Holt and Laury task as very easy.

A few limitations arise from the above study. The participants were all students of Computer Science, and they were used to pie charts as a way of data visualisation. The majority of them were between the ages 18-22, therefore this does not allow us to generalise this finding for other age groups. To summarise though, the findings from this study showed evidence that interactive engagement with graphs can have a significant effect on reducing people's irrational choices, which can improve the predictive ability of the Holt and Laury task and also extend the implementation of the risk task in a gamified version which would improve other shortcomings of it, for examples the need for incentives. Future research steps, as well as a further discussion of the results and the limitations derived from this study, are presented in the next chapter.

5. DISCUSSION AND CONCLUSIONS

5.1 Introduction

This chapter summarises the main findings of the research undertaken in the context of this thesis. It also provides the links between the findings and the empirical evidence. A reflection of the general methods is presented. This reflection concludes that the benefits gained from the applied methodology lead to the establishment of an appropriate interactive presentational way of lotteries in a risk elicitation task that can reduce participant irrational choices which was the purpose of this thesis. The studies' limitations are discussed and the areas of possible research future steps are highlighted.

5.2 Results and links to the empirical evidence

To answer the main research question regarding whether incorporating visual elements and graphs can influence rational decision making in a risk elicitation task, three experimental studies were undertaken. As follows, each study's results that lead us to this thesis contribution, which involves the use of interactive approaches with pie charts that reduce participant's inconsistent choices are presented. The results and the links with the empirical evidence are presented as follows.

5.2.1 Experiment 1

The first experiment investigated whether the introduction of game-based elements, more specifically visuals, into a risk elicitation task would influence rational decision making in a risk lottery task. Based on requirements described in 2.2.3 regarding risk elicitation task, a simplified gamified Holt and Laury task was developed that included visuals to simplify the presentation of the numerical lotteries. The experiment showed that there was high variance arising from inconsistency in the choices of participants; 16 participants out of 32 chose irrationally in the game. Additionally, the correlation between the game and the validated task showed no association between participants' choices in the two constructs $r (32) = .120$, $p <. 511$. The variance of participants' choices in the game does not allow the alternative hypothesis that there would be a significant association between participants' choices in the two constructs to be accepted or the null hypothesis to disprove. However, the results do highlight the issues with participants' irrational choices in risk elicitation tasks and show that, despite usability (as revealed from a validated usability satisfaction questionnaire, M = 6.86 of maximum 10 and SD= 0.62), the inconsistency persisted in the Risk House game. Even though the sample size of this pilot study does not allow for any generalisations regarding visualisation in games, the persisted inconsistency in participants' choices points toward the direction that further investigation is needed for the type of visual that should be used to display the lotteries Holt and Laury task. This has the clear implication that the problem of high rates of inconsistency in the risk elicitation task of Holt and Laury required further investigation to understand the factors that contribute to this pattern.

As it was described in Section 2.3.1, games provide the potential of using visuals to present procedures or tasks that otherwise were textual within questionnaires (van Lankveld et al. 2011). Hence, behavioural tasks become simplified and more understandable to the audience. Nevertheless, the findings of this experiment showed that implementing the task in a gamified context with visuals did not reduce participants irrational choices Therefore, the fundamental factors that influence participants' rational choices between risk and safe lotteries remained unknown. The latter conclusion is important for future development of in-game assessment tools because it shows that just gamifying a tool, without doing a background investigation of the factors that influence people's reasoning process does not necessary lead to an improved version of the conventional pen and paper task.

5.2.2 Experiment 2

The exploratory study presented in Chapter 3 investigated the factors that influence rational decision-making in a risk elicitation task. This experiment allowed the investigation of the most profound reasons that influence rational decision making in risk elicitation tasks and the exploration of using external representations to reason participants' choices. This exploratory method of external representations could indicate whether using graphs as external representations could influence rational decision making which supports the high-level research question of these studies. The null hypothesis is partially rejected. According to the results, only the part of hypothesis regarding the relationship between education and rational decision making has been approved. The results of the exploratory study showed that someone's score in a validated numeracy scale and cognitive reflection task style as well as the external representations would use to reason an option, have a significant association in evaluating the options rationally in a risk elicitation task that would lead to a consistent behaviour through the task. More specifically, it was shown that participants who scored higher in the validated numeracy scale task, they were more likely to evaluate rationally the option from the lottery task, X^2 (1, N= 60) = 4.176, p = .041. Similarly, participants who scored higher in the cognitive reflection task had higher chance to evaluate the lottery option rationally compared to those that scored lower, as it was showed by the Fisher's Exact Test $p < 0.5$. These findings reject the null hypothesis and agree with literature supporting that people with low numerical skills and automatic thinkers make inconsistent choices in the lottery task, as it is described in Section 2.4.1 (Frederick 2005; Dave et al. 2010). Additionally, the logistic regression performed to identify whether using external representation would predict rational decision making of participants, showed that when participants used graphs, it would predict their rational decision making, $p < .02$. Empirical evidence on challenges of probabilities presentation suggest the use of graphical information that could enable the individuals who have difficulty in mathematical presentations to understand them, as it has been described in the Section 2.4.2. Additionally, visual thinking and use of external representations like graphs were shown to support successful probability problem solving as it is presented in Literature Review Section 2.4.3. This experiment's finding confirms empirical evidence that visual thinking and use of external representations help participants to solve the probability successfully. However, it also showed that using graphs has greater likelihood of answer rationally compared to using images or any other way of external representation, as it has also been supported by Hegarty and Kozhnenikov (1999). This has the clear implications that future implementations of Holt and Laury task should engage participants to use graphs to understand the probability options better and choose rationally. From this study, it was also shown that participants' years of schooling did not influence their rational decision making. However, as participants were students from the United Kingdom this may not hold true for other cultures or audiences. The assumption that educational level could play a role on people's rational choice as it was shown from empirical evidence with Rwandan adults (Jacobson and Petrie 2009) needs further investigation beyond this experimentation as the definition of low education could be different. The educational level defined as the amount of year spent in higher education does not reflect the educational level definition for other countries and cultures. Therefore, as it was found that there were no differences between Groups 1 and Groups 2 that have different educational levels, it may not be true for differences in educational levels defined differently in other audiences or contexts.

A qualitative approach may also have provided more insight into why graphs helped people to understand the lottery of the Holt and Laury task. Even though Hegarty and Kozhnenikov (1999) supported that graphs lead to more successful problem solving compared to images, Wilkinson and Wills (2005) claimed that humans are poor at estimating angles of pie charts, which make them difficult to understand. Nevertheless, this may account for the fact that in this study participants draw the pie charts themselves to reason their choices and did not choose between pictures of pie charts. The finding of this study demonstrates that the use of graphical external representations, is more likely to result in rational choices in risk elicitation tasks. Hence, the data supports the hypothesis that using graphs externally would help people with low numerical scores and automatic thinking style to understand the lottery from the risk elicitation task of Holt and Laury.

5.2.3 Experiment 3

Building on the results from Experiment 2, the last experiment investigated whether engaging users in interacting with graphs would influence rational decision making in the risk elicitation task thus it would reduce participants' inconsistent choices compared to the passive display of the lottery options with numbers or pie charts. This experiment would allow to draw conclusions regarding the high-level research question. After a review of the literature on graphs as presented in Section 2.5.1 and graphical displays of past experiments with Holt and Laury Section 2.5.2, pie charts were used to prompt participants for reasoning their choices in Holt and Laury task. Pie charts were shown to be better suited as they are more well-known from the general public and easily comprehensible to compare the size of two proportions when they are accompanied by labels (Nelson, Hesse and Croyle 2009). Additionally, empirical evidence from previous studies with the Holt and Laury task experimented with pie charts used as visual aids with promising outcomes (Eckel, Warnick and Johnson 2005; Boughera, Gassmann and Piet 2011; Habib et al. 2016). A step further though was implemented in terms of asking participants to draw the pie charts themselves according to the label of probability along with the payoff displayed on the task, to test whether using external representation of pie charts would help them reason rationally. In total 225 participants were assessed for their irrational choices between the three different display formats. The experiment showed that participants irrational choices were reduced significantly when they used pie charts to reason their choices compared to the two other formats; 36 (47.4%) participants showed an inconsistent behaviour in Holt and Laury numerical format, 27 (35.5%) participants in the pie chart format and only 5 (6.8%) participants in the interactive pie chart Holt and Laury task. The inconsistency difference between interactive pie chart format and numerical or pie chart format was statistical significant, $p < .00$. Therefore, the null hypothesis has been rejected since the findings showed that participants with lower numerical skills can have significantly less irrational choices when they are asked to interact with a pie chart while reasoning for Holt and Laury task compared to the passive pie chart display format or the numerical display Holt and Laury. This outcome confirms the empirical evidence suggesting using graphs as external representation leads to successful probability problem solving (Hegarty and Kozhevnikov 1999) and thus rational choices in the Holt and Laury task. In the experiment, it was also shown that participants who scored low in the numeracy scale answered more rationally in the Holt and Laury task when interacted with pie charts than any other format, confirming the need for a graph that accompany complex information for people who do not have the training or the background to understand numerical probabilities (Carlsson, Johansson-Stenman and Martinsson 2004; Corso, Hammitt and Graham 2001). The interactive pie chart display format was also filled in rationally from those that scored also higher in the numeracy scale, as it is indicated from the low consistency rate in the task (6.8%). This has a significant implication for future implementations of the task. This rational choice is linked to a meaningful contribution to the task for three main reasons. First, the task could be used by people with low numerical skills and reflect their accurate risk preferences. Second, when assessing a certain population to predict risk taking behaviour in similar investment choices, e.g. farmers, participants' choices in the task would be accurately predicted.

Third, there would be less noise in the data and there would be less cases (if any) that data would be excluded from the analysis. However, a point that should be considered is that as the individuals are guided to answer consistently, the more consistent their answers the more variance would tend towards zero, thus the validity of the metric might be affected. Therefore, for future step the validity of the metric needs to be examined and reassured. Furthermore, as this approach was only examined with UK University students, there is a limitation of not generalising to other cultures and audiences. A qualitative approach would have investigated more in depth on the underlying factors why interacting with pie charts help people to understand the lotteries better. The study findings differ from the Bauremeister and Musshoff (2016) finding which demonstrated that bags filled with balls reduce participants inconsistent choices, decreasing the percentage of inconsistency from 23.49% with the numerical format to 12.66% with the bags filled with balls. However, the inconsistency rate with interactive pie charts was even more reduced, from 46% with the numerical to 6.8% with the interactive pie charts. Additionally, this experiment compared the interactive pie charts with the passive ones who also showed quite a substantial inconsistent rate difference, from 35.5% with the passive pie charts to 6.8% with the interactive ones. Hence the data supports the hypothesis that participants reduce their inconsistent choices in the Holt and Laury when they interact with pie charts to reason them.

Empirical studies have shown that when individuals use schematic representations (graphs) for probability solving they have a higher likelihood of success (Hegarty and Kozhevnikov 1999). The options in lottery risk elicitation tasks are most often presented as numerical probabilities, which people have difficulty to understand (Peters et al. 2008). Therefore, the engagement of people in 'playing around' with graphs to understand the probabilities described in the lottery options, solves the issue with irrational choices that either have to be excluded from the data analysis or if included provide misleading predictions. This outcome supports the overall research question of this study whether using graphs or visuals can influence rational decision making in risk elicitation tasks by demonstrating that using pie charts to reason the lotteries reduce significantly participants' irrational choices. This interactivity with graphs could be embedded inside a gamified task where participants could engage with filling in the proportions of lotteries with one click and then choose the option for the task they would prefer. This procedure automatically solves two problems that were highlighted by the literature review, engagement with participants outside the laboratory setting and simplifying the task for less numerate participants. Additionally, this finding also highlights the importance of graphical external representations for probability problems that are difficult to understand, for instance for the presentation of health risks. Hence, the data supports the hypothesis that the use of interactive pie charts, which may extend to games, is more likely to result in consistent choices.

5.3 Reflections and limitations

In the previous sections, the summary of the results, contributions and the links with the empirical evidence were introduced, while the next section reviews limitations of each study in the context of the methodology applied. Over the course of this research, a variety of methods approaches were tailored for the specific of this research and the individual experiments. Therefore, some reflections, as well as limitations that were encountered, can provide some potentially useful insights for future work.

5.3.1 Reflections on Experiment 1

A major advantage of the method used for this approach is the usability study, which was undertaken in the initial stages of development of the game. The usability evaluation of the task provided further insight into the preferences of the audiences, which included an increase of numbers size and the click ability of the buttons. It also eliminated any distraction features of the game. The usability evaluation study allowed for refinement of the gamified task to ensure that the game was usable and that usability would not affect participants' performance. Despite this, the inconsistency rate of participant's choices was (50%) enough to conclude that the same issue of the pen and paper task, as it was identified by both Galarza (2009) at 52% rate and Jacobson and Petrie (2009) with an inconsistency rate of 55%, occurred in the gamified task as well. A limitation is the extent of the experimental procedure that lasted almost an hour, which might cause frustration to the participants. The time length was due to the BART because it involves several sessions. However, it was the only risk elicitation task that has been validated (Lejeuz 2002) and was also computerised method as the gamified approach that was correlated with. Another limitation derives from the definition of risk propensity in this thesis and to acknowledge that a quantitative assessment may not reflect a 'true' behaviour but rather predisposed to certain definitions of it, that the riskier choices someone does in a risk elicitation task, the more risk taker the individual is. Similarly, there is an also supported view that considers risk propensity a personality trait and not a behaviour that changes according to the context. Therefore, the risk elicitation task of Holt and Laury could predict an individual's risk propensity in all contexts and not only regarding economic behaviour. However, the limitations arising from not considering risk propensity a personality trait involve to the issue that is affected by external factors which makes it difficult to assess (Nicholson et al. 2002). Finally, the quantitative route lacks the richness of data that might be associated with interviews in a principally qualitative investigation.

5.3.2 Reflections on Experiment 2

The research design of this study had the advantage of investigating presentational formats involving participants' external representations which provided a more in-depth insight into people's underlying thinking processes while solving similar probability problems. This approach varies from others in the respect that most often different presentational formats are presented to participants and are forced to choose one of the available options. Overall asking participants to use their own external representations was proved useful to identify an appropriate formula to improve their performance in the risk elicitation task. A limitation that should be taken into consideration is that the sample was drawn from UK participants, which may differ when applied to other audiences and cultures. The low numeracy level may not represent general low numerical skills rather how these are defined by the numeracy level and scale in this study. The educational level also remains a limitation as it has been defined, in this thesis, as the full years of schooling. This may differ in other audiences and cultures.

5.3.3 Reflections on Experiment 3

The between subject design was employed for the interactive approaches experiment. By applying this method, the comparison between the three different display formats detected the differences in inconsistency rates. This method proved useful, allowing for successful identification of both thresholds and big effects. A possible limitation of the study is that participants were undergraduate students, therefore were familiar with pie charts. However, there is empirical evidence supporting that students and real decision maker's process information the same way (Ashton and Kramer 1980). As it was shown from the second experiment, educational level does not affect people's rational evaluation of lotteries, but numeracy level does. For that matter, a numeracy scale was involved in the experimental procedure. As it was mentioned earlier, the score on a numeracy scale employed for this study defines numeracy level. Thus, it may not reflect the numeracy level of a sample measured by other cultures or defined differently from this thesis scope. Additionally, this interactive approach should be also examined in relation with cognitive reflections tasks to determine whether this interactive approach would help them to reflect more on the task and choose consistently. Since though, in this experiment participants had to notice the pay offs of the lotteries to draw the pie chart, is most likely that chose reflectively. A limitation that should also be considered is whether by forcing a rational probabilistic approach to reason, can make these results to carry over to situations beyond the test. Therefore, a next step requires examining whether the interactivity with pie charts help people to choose rationally according to their risk preferences in other situations. Finally, another consideration would be whether by interacting with pie charts primes rational answers and thus reducing the capacity of the test to account for individuals that may understand probability. However, from the results, it was shown that it did not affect participants with higher numeracy scores in the in the numeracy scale and that it influenced rational decision making in both groups who scored high and low in the numeracy scale

5.4 Future Work

This thesis described a series of experiments that investigated whether gamification and interactive approaches can influence rational decision-making during risk elicitation tasks. An area for future work is to investigate the interactive pie charts with other levels of education as defined by different cultures and even more with illiterate people. This further experimentation could verify whether this presentational way in Holt and Laury is useful for other cultures and audiences. Also, future work can be seen in the area of categorization of people according to their educational level. There is a debate regarding someone's educational level especially due to cultural differences. A potentially future area of work could look at people's differences of the same educational levels in different cultures about their understanding of probabilistic information and numeracy level. This additional study will provide some insight to the categorisation of uneducated and educated people in similar tasks as well as to understand whether such classification is needed for comprehension of probabilistic information. Similarly, investigating other numerical scales or measures in eastern and western countries in association with interactive pie charts can provide a full picture of the presentational format usefulness. Another further step in line with future research would be, following a qualitative approach along with the quantitative experimentation of Holt and Laury displayed interactive pie chart can provide further explanations to why people understand the lotteries when they use pie charts to reason them. Additionally, the examination of the validity of the metric need to be investigated to ensure that this interactive pie approach increases the predictive value as well as the validity of the task. An area for potential future lines of research is with the implementation of gamified elements into a risk elicitation task. Taking into consideration that from the final experiments' results it was found that the complexity of the lottery options is solved by presenting the lotteries in pie charts where participants have to fill in the proportions of the probabilities and payoffs, the next step could be to implement this in a digital gamified context. Then this game could be potentially correlated with other risk assessments and the pen and paper interactive pie chart approach to identify whether the gamified elements compromise the construct validity of the task. Future work is also needed to fully understand the effects on other forms of validity. This study could not only provide guidelines on how to design a gamified task that could be used as an assessment, but it would also offer a valid risk elicitation task that could be used by commercial and research organisations. As discussed in the 2nd and 3rd Sections of the literature review, games have features that can also overcome some other problems of risk elicitation tasks. As discussed in the literature, MPL methods (which include a series of binary choices between lottery options) of risk elicitation tasks need to be incentivised for the participants to reveal preferences truthfully (Charness, Gneezy and Imas 2013). However, gamified elements, which include bar progress, leaderboards and virtual rewards can replace the value of money and generate the sense of incentives in contexts where money is an essential motivator (Easley and Ghosh 2013; Deterding 2014; Lameras et al. 2016). Additionally, a game can serve as the medium for the risk elicitation task to be distributed to a large sample of participants via smartphones or computers (Teki, Kumar and Griffiths 2015).

Another area for additional work can potentially be seen on the use of interactive approach of presenting complex probabilistic information. This method was used by this thesis towards improving people with low numerical skills and automatic thinking style, understanding of probabilities. Therefore, this could be used to any other area where complex probability information is presented to people. For example, in health-related risks where individual needs to understand which type of therapy needs to follow, e.g. a surgery compared to a long-time medication therapy (Peter et al. 2008). A simple application where individuals can interact with a pie chart to understand deeper the advantages and disadvantages of a specific therapy, could help them process, the information and choose the most appropriate method for them.

5.5 Synopsis

To answer the main research question whether the incorporation of visuals and graphs can influence rational decision making during risk elicitation task, that was formulated based on empirical evidence presented in 3 Chapter, three quantitative studies were implemented. Results from the first experiment, the pilot study, with the gamified risk assessment tool revealed that despite a good usability level (as revealed from a validated usability satisfaction questionnaire, $M = 6.86$ of maximum 10 and $SD= 0.62$), high inconsistency rate among participants in the task persisted where 16 participants out of 32 chose irrationally in the game. Results from the second experiment showed that from a sample of 60 participants, people who scored lower in the numeracy validated metric had greater likelihood to choose irrationally, $X^2(1) = 4.176$, $p = .041$. Similarly, participants who scored lower in the cognitive reflection task had greater likelihood of answering irrationally in the lottery option of Holt and Laury task, Fisher's Exact Test $p < 0.5$. These outcomes confirm previous research showing that people who score low in numeracy scales or they score lower in cognitive reflection task tend to have difficulty in choosing consistently in probability related problems (Dave et al. 2010; Peters et al. 2008; Frederick 2005). However, when those groups used graphs when reasoning the lottery options, they have greater likelihood of answering rationally in the lottery option of the Holt and Laury task, $p < .031$. This result contributes to empirical evidence suggesting that external representation help people in rational probability solving (Zahner and Corter 2002). Results from the third experiment with 225 participants revealed that the group who used interactive pie charts to reason the lottery options chose more rationally compared to the group that had either passive pie charts or numerical presentation of the lotteries in Holt and Laury task, $p < .000$. This outcome is linked with evidence showing that when people use graphical external representation then helps them in probability problem solving (Hegarty and Kozhevnikov 1999).

The results of the three experiments extended the empirical evidence on using visual in games to simplify procedures and on using graphs as external representations for successful problem solving. However, there are a few limitations that need to be taken into consideration and examined in further steps related to this research subject. Different definitions of risk propensity, education, numeracy level and rational decision that are not used in these studies need to be investigated in relation to the outcomes that using graphs as external representations help people to reason rationally in risk elicitation tasks. This is necessary to ensure that this finding agrees with the numeracy level defined from other validated metrics.

The dissertation and the experimental methodology presented herein point towards the development of a gamified risk assessment tool that goes beyond the commonly standard risk elicitation tasks by providing evidence to support that interactive media can enhance people's understanding of lottery options in the pen and paper Holt and Laury task which appeal to the benefits of gamified elements for a risk assessment. This thesis considers a variety of future work areas that also involve using the interactive pie chart approach in a gamified context and explore whether gamifications compromise the validity of the task. This way, if validity is ensured, it will open a path of exploring the rest of risk elicitation tasks issues, such as distribution through smartphones and no need of incentives that could be overcome with the use of games. Finally, this work identified future directions for risk elicitation tasks, understanding of complex probabilistic information and suggestion for further clarifications on categorising people according to their educational level towards understanding of probabilities.

6. REFERENCES

A'Hearn, B., Baten, J., and Crayen, D. (2009) 'Quantifying Quantitative Literacy: Age Heaping and the History of Human Capital'. *The Journal of Economic History* 69 (03), 783-808

Akay, A., Martinsson, P., Medhin, H., and Trautmann, S. T. (2012) 'Attitudes Toward Uncertainty among the Poor: An Experiment in Rural Ethiopia'. *Theory and Decision* 73 (3), 453-464

Aklin, W. M., Lejuez, C., Zvolensky, M. J., Kahler, C. W., and Gwadz, M. (2005) 'Evaluation of Behavioral Measures of Risk Taking Propensity with Inner City Adolescents'. *Behaviour Research and Therapy* 43 (2), 215-228

Ancker, J. S., Senathirajah, Y., Kukafka, R., and Starren, J. B. (2006) 'Design Features of Graphs in Health Risk Communication: A Systematic Review'. *Journal of the American Medical Informatics Association* 13 (6), 608-618

Ancker, J. S., Weber, E. U., and Kukafka, R. (2011) 'Effects of Game-Like Interactive Graphics on Risk Perceptions and Decisions'. *Medical Decision Making* 31 (1), 130-142

Andersen, S., Harrison, G.W., Lau, M.I., Rutström, E.E., (2006) 'Elicitation using multiple price list formats.' *Experimental Economics* 9 (4), 383–405

Anderson, L. R. and Mellor, J. M. (2008) 'Predicting Health Behaviors with an Experimental Measure of Risk Preference'. *Journal of Health Economics* 27 (5), 1260-1274

Annetta, L. A. (2010) 'The "I's" have it: A Framework for Serious Educational Game Design.'. *Review of General Psychology* 14 (2), 105

Arrow Kenneth, J. (1965) 'Aspects of the Theory of Risk-Bearing'. *Helsinki, Yrjö Hahnsson Foundation*.

Ashton, R. H. and Kramer, S. S. (1980) 'Students as Surrogates in Behavioral Accounting Research: Some Evidence'. *Journal of Accounting Research*,1-15

Bagozzi, R. P., Yi, Y., and Singh, S. (1991) 'On the use of Structural Equation Models in Experimental Designs: Two Extensions'. *International Journal of Research in Marketing* 8 (2), 125-140

Bauermeister, G. and Mußhoff, O. (eds.) (2016) *56th Annual Conference, Bonn, Germany, September 28-30, 2016*. 'Risk Attitude and Inconsistencies-does the Choice of Display Format and Risk Elicitation Method Influence the Outcomes?': *German Association of Agricultural Economists (GEWISOLA)*

Beauchamp, T. L. (2008) 'The Belmont Report'. *The Oxford Textbook of Clinical Research Ethics,* 21-28

Bechara, A., Damasio, H., Tranel, D., and Damasio, A. R. (2005) 'The Iowa Gambling Task and the Somatic Marker Hypothesis: Some Questions and Answers'. *Trends in Cognitive Sciences* 9 (4), 159-162

Bellotti, F., Kapralos, B., Lee, K., Moreno-Ger, P., and Berta, R. (2013) 'Assessment in and of Serious Games: An Overview'. *Advances in Human-Computer Interaction* 2013, 1

Binswanger, H. P. (1981) 'Attitudes Toward Risk: Theoretical Implications of an Experiment in Rural India'. *The Economic Journal* 91 (364), 867-890

Black, W. C., Nease, R. F., Jr, and Tosteson, A. N. (1995) 'Perceptions of Breast Cancer Risk and Screening Effectiveness in Women Younger than 50 Years of Age'. *Journal of the National Cancer Institute* 87 (10), 720-731

Bougherara, D., Gassmann, X., and Piet, L. (2011) ‘A Structural Estimation of French Farmers’ Risk Preferences: An Artefactual Field Experiment.’ Retrieved from: https://ideas.repec.org/p/rae/wpaper/201106.html

Brener, N. D., Billy, J. O., and Grady, W. R. (2003) 'Assessment of Factors Affecting the Validity of Self-Reported Health-Risk Behavior among Adolescents: Evidence from the Scientific Literature'. *Journal of Adolescent Health* 33 (6), 436-457

Brick, K., Visser, M., and Burns, J. (2012) 'Risk Aversion: Experimental Evidence from South African Fishing Communities'. *American Journal of Agricultural Economics* 94 (1), 133-152

Bromiley, P., & Curley, S. (1992). Individual differences in risk taking. *Risk-taking behavior* 4, 87-132

Brown, H. R., Zeidman, P., Smittenaar, P., Adams, R. A., McNab, F., Rutledge, R. B., and Dolan, R. J. (2014) 'Crowdsourcing for Cognitive science–the Utility of Smartphones'. *PloS One* 9 (7), e100662

Boyle, E. A., Hainey, T., Connolly, T. M., Gray, G., Earp, J., Ott, M., ... & Pereira, J. (2016). An update to the systematic literature review of empirical evidence of the impacts and outcomes of computer games and serious games. *Computers & Education* 94: 178-192.

Buchner, A., Faul, F., and Erdfelder, E. (1992) 'G Power: A Priori, Post-Hoc, and Compromise Power Analyses for the Macintosh'. *Computer Programme.Bonn, Germany: Bonn University*

Buelow, M. T. and Suhr, J. A. (2009) 'Construct Validity of the Iowa Gambling Task'. *Neuropsychology Review* 19 (1), 102-114

Byrnes, J. P., Miller, D. C., and Schafer, W. D. (1999) ‘Gender Differences in Risk Taking: A Meta-Analysis.’ *Psychological Bulletin* 125 (3), 367-383

Campbell, D. T., and Fiske, D. (1959) ‘Convergent and discriminant validation by the multitrait–multimethod matrix.’ *Psychological Bulletin* 56, 81–10

Carlsson, F., Johansson-Stenman, O., and Martinsson, P. (2004) ‘Is transport safety more valuable in the air?’*Journal of Risk and Uncertainty* 28, 147–163

Charness, G. and Viceisza, A. (2012) 'Comprehension and Risk Elicitation in the Field: Evidence from Rural Senegal' UC Santa Barbara: Department of Economics, UCSB. Retrieved from: http://escholarship.org/uc/item/5512d150

Charness, G., Gneezy, U., and Imas, A. (2013) 'Experimental Methods: Eliciting Risk Preferences'. *Journal of Economic Behavior & Organization* 87, 43-51

Chin, J. P., Diehl, V. A. and Norman, K. L. (1988) ‘Development of an instrument measuring user satisfaction of the human-computer interface.’ *Proceedings of SIGCHI '88*, 213-218

Cleveland, W. S. and McGill, R. (1986) ‘An Experiment in Graphical Perception.’ *International Journal of Man-Machine Studies* 25 (5), 491-500

Cleveland, W. S., Harris, C. S., and McGill, R. (1982) 'Judgments of Circle Sizes on Statistical Maps'. *Journal of the American Statistical Association* 77 (379), 541-547

Coller, M. and Williams, M. B. (1999) ‘Eliciting Individual Discount Rates.’ *Experimental Economics* 2 (2), 107-127

Confrey, J. and Smith, E. (eds.) (1991) ’A Framework for Functions: Prototypes, Multiple Representations and Transformations' *Proceedings of the 13th Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education.*: ERIC

Cook, T. D., Campbell, D. T., and Day, A. (1979) *Quasi-Experimentation: Design & Analysis Issues for Field Settings*: Houghton Mifflin Boston.

Corso, P. S., Hammitt, J. K., and Graham, J. D. (2001) 'Valuing Mortality-Risk Reduction: Using Visual Aids to Improve the Validity of Contingent Valuation'. *Journal of Risk and Uncertainty* 23 (2), 165-184

Corter, J. E. and Zahner, D. C. (2007) 'Use of External Visual Representations in Probability Problem Solving'. *Statistics Education Research Journal* 6 (1), 22-50

Cronbach, L. J. and Meehl, P. E. (1955) 'Construct Validity in Psychological Tests'. *Psychological Bulletin* 52 (4), 281

Dahling, J. J., Whitaker, B. G., and Levy, P. E. (2008) 'The Development and Validation of a New Machiavellianism Scale'. *Journal of Management* 35 (2), 219-257

Dave, C., Eckel, C. C., Johnson, C. A., and Rojas, C. (2010) 'Eliciting Risk Preferences: When is Simple Better?'. *Journal of Risk and Uncertainty* 41 (3), 219-243

Deterding, S. (2014) 'Eudaimonic Design, Or: Six Invitations to Rethink Gamification'. In: Rethinking Gamification. Retrieved from: https://ssrn.com/abstract=2466374

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., and Wagner, G. G. (2011) 'Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences'. *Journal of the European Economic Association* 9 (3), 522-550

Easley, D. and Ghosh, A. (eds.) (2013) *Proceedings of the Fourteenth ACM Conference on Electronic Commerce*. 'Incentives, Gamification, and Game Theory: An Economic Approach to Badge Design': ACM

Eckel, C. C. and Grossman, P. J. (2002) 'Sex Differences and Statistical Stereotyping in Attitudes Toward Financial Risk'. *Evolution and Human Behavior* 23 (4), 281-295

Eckel, C., Engle-Warnick, J., and Johnson, C. (2005) *Adaptive Elicitation of Risk Preferences"*. Working paper: McGrill University

Elting, L. S., Martin, C. G., Cantor, S. B., and Rubenstein, E. B. (1999) 'Influence of Data Display Formats on Physician Investigators' Decisions to Stop Clinical Trials: Prospective Trial with Repeated Measures'. *BMJ (Clinical Research Ed.)* 318 (7197), 1527-1531

Evans, J. S. B. (2007) *Hypothetical Thinking: Dual Processes in Reasoning and Judgement*: Psychology Press

Eysenck, S. B., Eysenck, H. J., and Barrett, P. (1985) 'A Revised Version of the Psychoticism Scale'. *Personality and Individual Differences* 6 (1), 21-29

Fagerlin, A., Zikmund-Fisher, B. J., and Ubel, P. A. (2005) 'How Making a Risk Estimate Can Change the Feel of that Risk: Shifting Attitudes Toward Breast Cancer Risk in a General Public Survey'. *Patient Education and Counseling* 57 (3), 294-299

Feldman-Stewart, D., Kocovski, N., McConnell, B. A., Brundage, M. D., and Mackillop, W. J. (2000) 'Perception of Quantitative Information for Treatment Decisions'. *Medical Decision Making: An International Journal of the Society for Medical Decision Making* 20 (2), 228-238

Feldman-Stewart, D., Brundage, M. D., McConnell, B. A., and Mackillop, W. J. (2000) 'Practical Issues in Assisting Shared decision-making'. *Health Expectations* 3 (1), 46-54

Fiske, D. W. (1982) 'Convergent-Discriminant Validation in Measurements and Research Strategies.'. *New Directions for Methodology of Social & Behavioral Science*, 12 (1), 97-92

Forth, A. E., Brown, S. L., Hart, S. D., and Hare, R. D. (1996) 'The Assessment of Psychopathy in Male and Female Noncriminals: Reliability and Validity'. *Personality and Individual Differences* [online] 20 (5), 531-543

Frederick, S. (2005) 'Cognitive Reflection and Decision Making'. *The Journal of Economic Perspectives* 19 (4), 25-42

Froschauer, J., Seidel, I., Gärtner, M., Berger, H., and Merkl, D. (eds.) (2010) *Virtual Systems and Multimedia (VSMM), 2010 16th International Conference on*. 'Design and Evaluation of a Serious Game for Immersive Cultural Training': IEEE

Fulker, D. W., Eysenck, S. B., and Zuckerman, M. (1980) 'A Genetic and Environmental Analysis of Sensation Seeking'. *Journal of Research in Personality* 14 (2), 261-281

Gaissmaier, W., Wegwarth, O., Skopec, D., Müller, A., Broschinski, S., and Politi, M. C. (2012) 'Numbers can be Worth a Thousand Pictures: Individual Differences in Understanding Graphical and Numerical Representations of Health-Related Information.'. *Health* 31 (3), 286

Gajjar, B. (2013) 'Ethical Consideration in Research'. *International Journal for Research in Education* 2 (7)

Galarza, F. (2009) 'Choices Under Risk in Rural Peru'. Retrieved from: https://mpra.ub.uni-muenchen.de/id/eprint/17708

Galizzi, M. M., Miraldo, M., and Stavropoulou, C. (2013) 'Doctor-Patient Differences in Risk Preferences, and their Links to Decision-Making: A Field Experiment'. Retrieved from: http://hdl.handle.net/10044/1/12578

Garcia-Retamero, R. and Cokely, E. T. (2011) 'Effective Communication of Risks to Young Adults: Using Message Framing and Visual Aids to Increase Condom use and STD Screening.'. *Journal of Experimental Psychology: Applied* 17 (3), 270

Garcia-Retamero, R. and Galesic, M. (2010) 'Who Proficts from Visual Aids: Overcoming Challenges in People's Understanding of Risks'. *Social Science & Medicine* 70 (7), 1019-1025

Garcia-Retamero, R. and Dhami, M. K. (2011) 'Pictures Speak Louder than Numbers: On Communicating Medical Risks to Immigrants with Limited non-native Language Proficiency'. *Health Expectations* 14 (1), 46-57

Ghemawat, P. (1990) 'The Snowball Effect'. *International Journal of Industrial Organization* 8 (3), 335-351

Gigerenzer, G. and Edwards, A. (2003) 'Simple Tools for Understanding Risks: From Innumeracy to Insight'. *BMJ (Clinical Research Ed)* 327 (7417), 741-744

Gigerenzer, G. and Hoffrage, U. (1995) 'How to Improve Bayesian Reasoning without Instruction: Frequency Formats.'. *Psychological Review* 102 (4), 684

Gneezy, U. and Potters, J. (1997) 'An Experiment on Risk Taking and Evaluation Periods'. *The Quarterly Journal of Economics* 112 (2), 631-645

Goldin, G. and Shteingold, N. (2001) 'Systems of Representations and the Development of Mathematical Concepts'. *The Roles of Representation in School Mathematics* 2001, 1-23

Grable, J. E. (2000) 'Financial Risk Tolerance and Additional Factors that Affect Risk Taking in Everyday Money Matters'. *Journal of Business and Psychology* 14 (4), 625-630

Greeno, J. G. and Hall, R. P. (1997) 'Practicing Representation'. *Phi Delta Kappan* 78 (5), 361

Grinnell Jr, R. M. and Unrau, Y. A. (2010) *Social Work Research and Evaluation: Foundations of Evidence-Based Practice*. Oxford University Press

Guiso, L. and Paiella, M. (2004) 'The Role of Risk Aversion in Predicting Individual Behaviors' . Retrieved from: https://ssrn.com/abstract=608262

Guiso, L., Sapienza, P., and Zingales, L. (2013) 'The Determinants of Attitudes Toward Strategic Default on Mortgages'. *The Journal of Finance* 68 (4), 1473-1515

Gurmankin, A. D., Domchek, S., Stopfer, J., Fels, C., and Armstrong, K. (2005) 'Patients’ Resistance to Risk Information in Genetic Counseling for BRCA1/2'. *Archives of Internal Medicine*165 (5), 523-529

Habib, S., Friedman, D., Crockett, S., and James, D. (2016) 'List Construction and Lottery Presentation Modulate Multiple Price List Responses'

Hall, V. C., Bailey, J., and Tillman, C. (1997) 'Can Student-Generated Illustrations be Worth Ten Thousand Words?'. *Journal of Educational Psychology* 89 (4), 677

Hallahan, T. A., Faff, R. W., and McKenzie, M. D. (2004) 'An Empirical Investigation of Personal Financial Risk Tolerance'. *Financial Services Review* 13 (1), 57

Harper, B. D. and Norman, K. L. (eds.) (1993) *Proceedings of the 1st Annual Mid-Atlantic Human Factors Conference*. 'Improving User Satisfaction: The Questionnaire for User Interaction Satisfaction Version 5.5'

Harrison, G. W., Lau, M. I., Rutström, E. E., and Tarazona-Gómez, M. (2012) 'Preferences Over Social Risk'. *Oxford Economic Papers*, gps021

Hegarty, M. and Kozhevnikov, M. (1999) 'Types of visual–spatial Representations and Mathematical Problem Solving.'. *Journal of Educational Psychology* 91 (4), 684

Hill, R. V. and Viceisza, A. (2012) 'A Field Experiment on the Impact of Weather Shocks and Insurance on Risky Investment'. *Experimental Economics*15 (2), 341-371

Hillson, D. and Murray-Webster, R. (eds.) (2006) *PMI Global Congress EMEA Proceedings*. 'Managing Risk Attitude using Emotional Literacy'

Ho, S. Y. and Lowrie, T. (2014) 'The Model Method: Students’ Performance and its Effectiveness'. *The Journal of Mathematical Behavior* 35, 87-100

Hoffrage, U. and Gigerenzer, G. (1998) 'Using Natural Frequencies to Improve Diagnostic Inferences.' *Academic Medicine* 73 (5), 538-540

Holt, C. A. and Laury, S. K. (2002) 'Risk Aversion and Incentive Effects'. *American Economic Review* 92 (5), 1644-1655

Holt, C. A. and Laury, S. K. (2005) 'Risk Aversion and Incentive Effects: New Data without Order Effects'. *The American Economic Review* 95 (3), 902-904

Huerta, M. P., Cerdán, F., Lonjedo, M., and Edo, P. (eds.) (2011) *Proceedings of the Seventh Congress of the European Society for Research in Mathematics Education*. 'Assessing Difficulties of Conditional Probability Problems'

Ihli, H. J., Chiputwa, B., and Musshoff, O. (2016) 'Do Changing Probabilities Or Payoffs in Lottery-Choice Experiments Affect Risk Preference Outcomes? Evidence from Rural Uganda'. *Journal of Agricultural and Resource Economics* 41 (2), 324-345

Ives, B., Olson, M. H., and Baroudi, J. J. (1983) 'The Measurement of User Information Satisfaction'. *Communications of the ACM* 26 (10), 785-793

Jacobson, S. and Petrie, R. (2009) 'Learning from Mistakes: What do Inconsistent Choices Over Risk Tell Us?'. *Journal of Risk and Uncertainty* 38 (2), 143-158

James, W. (1907) 'Pragmatism's Conception of Truth'. *The Journal of Philosophy, Psychology and Scientific Methods* 4 (6), 141-155

Kahneman, D. and Tversky, A. (1979) 'Prospect Theory: An Analysis of Decision Under Risk'. *Econometrica: Journal of the Econometric Society* 263-291

Kahneman, D. and Tversky, A. (1981) The Simulation Heuristic. Retrieved from: ADA099504

Kahneman, D., Knetsch, J. L., and Thaler, R. H. (1991) 'Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias'. *The Journal of Economic Perspectives,* 193-206

Karoulis, A., Sylaiou, S., and White, M. (2006) 'Usability Evaluation of a Virtual Museum Interface'. *Informatica* 17 (3), 363-380

Kaufmann, G. (1990) 'Imagery effects on problem solving.' In P. J. Hampson, D. E. Marks, & J. T. E. Richardson (Eds.), Imagery: Current developments 169–197

Kelley, C. M. (2009) 'Cognitive Abilities and Superior Decision Making Under Risk: A Protocol Analysis and Process Model Evaluation Edward T. Cokely Max Planck Institute for Human Development Center for Adaptive Behavior and Cognition'. *Judgment and Decision Making* 4 (1), 20-33

King, M. and Newman, R. (2009) 'Evaluating Business Simulation Software: Approach, Tools and Pedagogy'. *On the Horizon* 17 (4), 368-377

Kirby, K. N., Petry, N. M., and Bickel, W. K. (1999) 'Heroin Addicts have Higher Discount Rates for Delayed Rewards than Non-Drug-using Controls'. *Journal of Experimental Psychology: General* 128 (1), 78

Knight, J., Weir, S., and Woldehanna, T. (2003) 'The Role of Education in Facilitating Risk-Taking and Innovation in Agriculture'. *The Journal of Development Studies* 39 (6), 1-22

Lameras, P., Arnab, S., Dunwell, I., Stewart, C., Clarke, S., and Petridis, P. (2016) 'Essential Features of Serious Games Design in Higher Education: Linking Learning Attributes to Game Mechanics'. *British Journal of Educational Technology*. Retrieved from:https://www.srhe.ac.uk/downloads/LamerasEssential_Features_of_%20Serious_Games%20Design_Short_FINAL.pdf

Landis, J. R. and Koch, G. G. (1977) 'The Measurement of Observer Agreement for Categorical Data'. *Biometrics* 159-174

Lejuez, C. W., Aklin, W. M., Zvolensky, M. J., and Pedulla, C. M. (2003) 'Evaluation of the Balloon Analogue Risk Task (BART) as a Predictor of Adolescent Real-World Risk-Taking Behaviours'. *Journal of Adolescence* 26 (4), 475-479

Lejuez, C. W., Read, J. P., Kahler, C. W., Richards, J. B., Ramsey, S. E., Stuart, G. L., Strong, D. R., and Brown, R. A. (2002) 'Evaluation of a Behavioral Measure of Risk Taking: The Balloon Analogue Risk Task (BART).'. *Journal of Experimental Psychology: Applied* 8 (2), 75

Lewis, J. R. (1995) 'IBM Computer Usability Satisfaction Questionnaires: Psychometric Evaluation and Instructions for Use.' *International Journal of Human-Computer Interaction.* 7, 57-78

Lind, A., Rosenberg, E., Seljom, P., Espegren, K., Fidje, A., and Lindberg, K. (2013) 'Analysis of the EU Renewable Energy Directive by a Techno-Economic Optimisation Model'. *Energy Policy* 60, 364-377

Lipkus, I. M. (2007) 'Numeric, Verbal, and Visual Formats of Conveying Health Risks: Suggested Best Practices and Future Recommendations'. *Medical Decision Making* 27 (5), 696-713

Lipkus, I. M. and Hollands, J. G. (1999) 'The Visual Communication of Risk'. *Journal of the National Cancer Institute.Monographs* 25(25), 149-163

Lipkus, I. M., Samsa, G., and Rimer, B. K. (2001) 'General Performance on a Numeracy Scale among Highly Educated Samples'. *Medical Decision Making* 21 (1), 37-44

Locke, J. (1948) 'An Essay Concerning Human Understanding, 1690.'

Lohmander, P. (2000) 'Optimal Sequential Forestry Decisions Under Risk'. *Annals of Operations Research* 95 (1-4), 217-228

MacCrimmon, K. R. and Wehrung, D. A. (1986) 'Taking Risks: The Management of uncertainty' *Administrative Science Quarterly* 33(4), 638-640

McNab, F. and Dolan, R. J. (2014) 'Dissociating Distractor-Filtering at Encoding and during Maintenance.'. *Journal of Experimental Psychology: Human Perception and Performance* 40 (3), 960

Mayer, R. (1989)'Systemic thinking fostered by illustrations in scientific text. ' *Journal of Educational Psychology* 81(2), 240–246

Mayer, R. E. and Gallini, J. K. (1990) 'When is an Illustration Worth Ten Thousand Words?'. *Journal of Educational Psychology* 82 (4), 715

Mazur, D. J. and Hickam, D. H. (1993) 'Patients' and Physicians' Interpretations of Graphic Data Displays'. *Medical Decision Making: An International Journal of the Society for Medical Decision Making* 13 (1), 59-63

McNab, F., Zeidman, P., Rutledge, R. B., Smittenaar, P., Brown, H. R., Adams, R. A., and Dolan, R. J. (2015) 'Age-Related Changes in Working Memory and the Ability to Ignore Distraction'. *Proceedings of the National Academy of Sciences of the United States of America* 112 (20), 6515-6518

Michael, D. R. and Chen, S. L. (2005) *Serious Games: Games that Educate, Train, and Inform*: Muska & Lipman/Premier-Trade

Mishra, S. and Lalumière, M. L. (2011) 'Individual Differences in Risk-Propensity: Associations between Personality and Behavioral Measures of Risk'. *Personality and Individual Differences* 50 (6), 869-873

Molitor, S., Ballstaedt, S. P., & Mandl, H. (1989) 'Problems in knowledge acquisition from text and pictures.' *Advances in Psychology* (58), 3-35

Moritz, B. B., Hill, A. V., and Donohue, K. L. (2013) 'Individual Differences in the Newsvendor Problem: Behavior and Cognitive Reflection'. *Journal of Operations Management* 31 (1), 72-85

Nelson, D. E., Hesse, B. W., and Croyle, R. T. (2009) 'Making Data Talk: Communicating Public Health Data to the Public, Policy Makers' . Retrieved from: https://www.cancer.gov/publications/health-communication/making-data-talk.pdf

Nielsen, T., Keil, A., and Zeller, M. (2013) 'Assessing Farmers' Risk Preferences and their Determinants in a Marginal Upland Area of Vietnam: A Comparison of Multiple Elicitation Techniques'. *Agricultural Economics* 44 (3), 2

Nicholson, N., Fenton-O'Creevy, M., Soane, E., and Willman, P. (2002) 'Risk Propensity and Personality'. *London Business School, Open University Business School and Said Business School Oxford, London*

Nuovo, J., Melnikow, J., and Chang, D. (2002) 'Reporting Number Needed to Treat and Absolute Risk Reduction in Randomized Controlled Trials'. *Jama* 287 (21), 2813-2814

Pape, S. J. and Tchoshanov, M. A. (2001) 'The Role of Representation (s) in Developing Mathematical Understanding'. *Theory into Practice* 40 (2), 118-127

Penner, D. E., Giles, N. D., Lehrer, R., and Schauble, L. (1997) 'Building Functional Models: Designing an Elbow'. *Journal of Research in Science Teaching* 34 (2), 125-143

Peters, E. (2008) 'Numeracy and the Perception and Communication of Risk'. *Annals of the New York Academy of Sciences* 1128 (1), 1-7

Peters, E. and Levin, I. P. (2008) 'Dissecting the Risky-Choice Framing Effect: Numeracy as an Individual-Difference Factor in Weighting Risky and Riskless Options'. *Judgment and Decision Making* 3 (6), 435

Peters, E., Hibbard, J., Slovic, P., and Dieckmann, N. (2007) 'Numeracy Skill and the Communication, Comprehension, and use of Risk-Benefit Information'. *Health Affairs (Project Hope)* 26 (3), 741-748

Peters, E., Västfjäll, D., Slovic, P., Mertz, C., Mazzocco, K., and Dickert, S. (2006) 'Numeracy and Decision Making'. *Psychological Science* 17 (5), 407-413

Pratt, J. W. (1964) 'Risk Aversion in the Small and in the Large'. *Econometrica: Journal of the Econometric Society,* 122-136

Presmeg, N. C. (1986). Visualization in high school mathematics. For Learning of Mathematics, 63(3), 42–46

Price, T. J. and Wetzstein, M. E. (1999) 'Irreversible Investment Decisions in Perennial Crops with Yield and Price Uncertainty'. *Journal of Agricultural and Resource Economics*, 173-185

Purvis, A., Boggess, W. G., Moss, C. B., and Holt, J. (1995) 'Technology Adoption Decisions Under Irreversibility and Uncertainty: An Ex Ante Appproach'. *American Journal of Agricultural Economics* 77 (3), 541-551

Reynaud, A. and Couture, S. (2012) 'Stability of Risk Preference Measures: Results from a Field Experiment on French Farmers'. *Theory and Decision* 73 (2), 203-221

Ritterfeld, U., Cody, M., and Vorderer, P. (2009) *Serious Games: Mechanisms and Effects*: Routledge

Roberts, W., Fillmore, M. T., and Milich, R. (2011) 'Linking Impulsivity and Inhibitory Control using Manual and Oculomotor Response Inhibition Tasks'. *Acta Psychologica* 138 (3), 419-428

Rohrmann, B. (2005) 'Risk Attitude Scales: Concepts, Questionnaires, Utilizations'. Retrieved from: Http://www.Rohrmannresearch.net/pdfs/rohrmann-Ras-Report.Pdf

Romer, D., Duckworth, A. L., Sznitman, S., and Park, S. (2010) 'Can Adolescents Learn Self-Control? Delay of Gratification in the Development of Control Over Risk Taking'. *Prevention Science* 11 (3), 319-330

Rose, A. F., Schnipper, J. L., Park, E. R., Poon, E. G., Li, Q., and Middleton, B. (2005) 'Using Qualitative Studies to Improve the Usability of an EMR'. *Journal of Biomedical Informatics* 38 (1), 51-60

Rothman, R. L., Housam, R., Weiss, H., Davis, D., Gregory, R., Gebretsadik, T., Shintani, A., and Elasy, T. A. (2006) 'Patient Understanding of Food Labels: The Role of Literacy and Numeracy'. *American Journal of Preventive Medicine* 31 (5), 391-398

Royak-Schaler, R., Blocker, D. E., Yali, A. M., Bynoe, M., Briant, K. J., and Smith, S. (2004) 'Breast and Colorectal Cancer Risk Communication Approaches with Low-Income African-American and Hispanic Women: Implications for Healthcare Providers'. *Journal of the National Medical Association* 96 (5), 598-608

Russell, W. E. (2000) ‘The use of visual devices in probability problem solving. ‘(Doctoral Dissertation, Columbia University, 2000). Dissertation Abstracts International, 61, 1333.

Santos-Trigo, M. (1996) ‘An exploration of strategies used by students to solve problems with multiple ways of solution.’ *Journal of Mathematical Behavior*, 15(3), 263–284

Schapira, M. M., Meade, C., and Nattinger, A. B. (1997) 'Enhanced Decision-Making: The use of a Videotape Decision-Aid for Patients with Prostate Cancer'. *Patient Education and Counseling* 30 (2), 119-127

Schapira, M. M., Nattinger, A. B., & McHorney, C. A. (2001). Frequency or probability? A qualitative study of risk communication formats used in health care. *Medical Decision Making* 21, 459 – 467

Schapira, M. M., Walker, C. M., Cappaert, K. J., Ganschow, P. S., Fletcher, K. E., McGinley, E. L., Del Pozo, S., Schauer, C., Tarima, S., and Jacobs, E. A. (2012) 'The Numeracy Understanding in Medicine Instrument: A Measure of Health Numeracy Developed using Item Response Theory'. *Medical Decision Making* 32 (6), 851-865

Schoemaker, P. J. (2013) *Experiments on Decisions Under Risk: The Expected Utility Hypothesis*: Springer Science & Business Media

Schreiber, C. (2004) 'The Interactive Development of Mathematical inscriptions—a Semiotic Perspective on Pupils Externalization in an Internet Chat about Mathematical Problems'. *Zdm* 36 (6), 185-194

Schuller, B. W., Dunwell, I., Weninger, F., and Paletta, L. (2013) 'Serious Gaming for Behavior Change: The State of Play'. *IEEE Pervasive Computing* 12 (3), 48-55

Schwartz, D. L. and Black, J. B. (1996) 'Shuttling between Depictive Models and Abstract Rules: Induction and Fallback'. *Cognitive Science* 20 (4), 457-497

Schwartz, L. M., Woloshin, S., Black, W. C., and Welch, H. G. (1997) 'The Role of Numeracy in Understanding the Benefit of Screening Mammography'. *Annals of Internal Medicine* 127 (11), 966-972

Shapiro, S. (1977) 'Evidence on Screening for Breast Cancer from a Randomized Trial'. *Cancer* 39 (6), 2772-2782

Shute, V. J. and Ke, F. (2012) 'Games, Learning, and Assessment'. in *Assessment in Game-Based Learning*. ed. by Anon: Springer, 43-58

Shute, V. J., Ventura, M., Bauer, M., and Zapata-Rivera, D. (2009) 'Melding the Power of Serious Games and Embedded Assessment to Monitor and Foster Learning'. *Serious Games: Mechanisms and Effects* 2, 295-321

Sitkin, S. B. and Pablo, A. L. (1992) 'Reconceptualizing the Determinants of Risk Behavior'. *Academy of Management Review* 17 (1), 9-38

Slovic, P. (2001) 'The Risk Game'. *Journal of Hazardous Materials* 86 (1), 17-24
Smittenaar P, Rutledge RB, Zeidman P, Adams RA, Brown H, Lewis G, et al. (2015) Proactive and Reactive Response Inhibition across the Lifespan. PLoS ONE 10(10): e0140383. doi: 10.1371/journal.pone.0140383

Spiegelhalter, D., Pearson, M., and Short, I. (2011) 'Visualizing Uncertainty about the Future'. *Science (New York, N.Y.)* 333 (6048), 1393-1400

Stainton, A. J., Johnson, J. E., and Borodzicz, E. P. (2010) 'Educational Validity of Business Gaming Simulation: A Research Methodology Framework'. *Simulation & Gaming* 41 (5), 705-723

Stone, E. R., Sieck, W. R., Bull, B. E., Yates, J. F., Parks, S. C., and Rush, C. J. (2003) 'Foreground: Background Salience: Explaining the Effects of Graphical Displays on Risk Avoidance'. *Organizational Behavior and Human Decision Processes* 90 (1), 19-36

Stylianou, D. A. and Silver, E. A. (2004) 'The Role of Visual Representations in Advanced Mathematical Problem Solving: An Examination of Expert-Novice Similarities and Differences'. *Mathematical Thinking and Learning* [online] 6 (4), 353-387

Suárez-Orozco, M. M. and Qin-Hilliard, D. (2004) *Globalization: Culture and Education in the New Millennium*: Univ of California Press

Tanaka, T., Camerer, C. F., and Nguyen, Q. (2016) 'Risk and Time Preferences: Linking Experimental and Household Survey Data from Vietnam'. in *Behavioral Economics of Preferences, Choices, and Happiness*. ed. by Anon: Springer, 3-25

Teddlie, C. and Tashakkori, A. (2009) *Foundations of Mixed Methods Research: Integrating Quantitative and Qualitative Approaches in the Social and Behavioral Sciences* Sage Publications Inc

Teki, S., Kumar, S., and Griffiths, T. D. (2016) 'Large-Scale Analysis of Auditory Segregation Behavior Crowdsourced Via a Smartphone App'. *PloS One* [online] 11 (4), e0153916

Trautmann, S. T., Vieider, F. M., and Wakker, P. P. (2011) 'Preference Reversals for Ambiguity Aversion'. *Management Science* 57 (7), 1320-1333

Turnbull, O. H., Evans, C. E., Bunce, A., Carzolio, B., and O'Connor, J. (2005) 'Emotion-Based Learning and Central Executive Resources: An Investigation of Intuition and the Iowa Gambling Task'. *Brain and Cognition* 57 (3), 244-247

Tversky, A. and Kahneman, D. (1992) 'Advances in Prospect Theory: Cumulative Representation of Uncertainty'. *Journal of Risk and Uncertainty* 5 (4), 297-323

Tversky, A. and Kahneman, D. (1985) 'The Framing of Decisions and the Psychology of Choice'. in *Environmental Impact Assessment, Technology Assessment, and Risk Analysis*. ed. by Anon: Springer, 107-129

Tversky, B. (2001). Spatial schemas in depictions. In M. Gattis (Ed.), Spatial schemas and abstract thought (pp. 79–112). Cambridge, MA: MIT Press

Van Den Berg, M., Fort, R., and Burger, K. (eds.) (2009) *International Association of Agricultural Economists Conference, Beijing, China*. 'Natural Hazards and Risk Aversion: Experimental Evidence from Latin America'

van Lankveld, G., Spronck, P., van den Herik, J., and Arntz, A. (eds.) (2011) *Computational Intelligence and Games (CIG), 2011 IEEE Conference on*. 'Games as Personality Profiling Tools': IEEE

Visschers, V. H., Meertens, R. M., Passchier, W. W., and De Vries, N. N. (2009) 'Probability Information in Risk Communication: A Review of the Research Literature'. *Risk Analysis* 29 (2), 267-287

Weber, E. U., Blais, A., and Betz, N. E. (2002) 'A domain‑specific risk‑attitude Scale: Measuring Risk Perceptions and Risk Behaviors'. *Journal of Behavioral Decision Making* 15 (4), 263-290

Westen, D. and Rosenthal, R. (2003) 'Quantifying Construct Validity: Two Simple Measures.'. *Journal of Personality and Social Psychology* 84 (3), 608

White, T. L., Lejuez, C. W., and de Wit, H. (2008) 'Test-Retest Characteristics of the Balloon Analogue Risk Task (BART).' *Experimental and Clinical Psychopharmacology* 16 (6), 565

Wilkinson, L. and Wills, G. (2008) 'Scagnostics Distributions'. *Journal of Computational and Graphical Statistics* 17 (2), 473-491

Winter-Nelson, A. and Amegbeto, K. (1998) 'Option Values to Conservation and Agricultural Price Policy: Application to Terrace Construction in Kenya'. *American Journal of Agricultural Economics* 80 (2), 409-418

Wölbert, E. and Riedl, A. (2013) 'Measuring Time and Risk Preferences: Reliability, Stability, Domain Specificity'. Working paper. Retrieved from: https://ssrn.com/abstract=2302494

Woloshin, S., Schwartz, L. M., Black, W. C., and Welch, H. G. (1999) 'Women's Perceptions of Breast Cancer Risk: How You Ask Matters'. *Medical Decision Making* 19 (3), 221-229

World Bank. (2010). Remittance prices worldwide. Retrieved from http://remittanceprices.worldban.org/documents/RemittancePriceWorldwideAnalysis.pdf

Wouters, P., Van Nimwegen, C., Van Oostendorp, H., and Van Der Spek, Erik D (2013) . ‘A Meta-Analysis of the Cognitive and Motivational Effects of Serious Games. ‘ *Jounal of Educational Psychology* 105 (2), 249 - 265

Zahner, D. and Corter, J. (eds.) (2002) *Annual Meeting of the American Educational Research Association, New Orleans, LA*. 'Clinical Interviewing to Uncover the Cognitive Processes of Probability Problem Solvers'

Zahner, D. and Corter, J. E. (2010) 'The Process of Probability Problem Solving: Use of External Visual Representations'. *Mathematical Thinking and Learning* [online] 12 (2), 177-204

Zhang, J. and Wang, H. (2009) 'An Exploration of the Relations between External Representations and Working Memory'. *PloS One* 4 (8), e6513

Zielke, M. A., Evans, M. J., Dufour, F., Christopher, T. V., Donahue, J. K., Johnson, P., Jennings, E. B., Friedman, B. S., Ounekeo, P. L., and Flores, R. (2009) 'Serious Games for Immersive Cultural Training: Creating a Living World'. *IEEE Computer Graphics and Applications* 29 (2), 49-60

Zikmund‐Fisher, B. J., Fagerlin, A., and Ubel, P. A. (2008) 'Improving Understanding of Adjuvant Therapy Options by using Simpler Risk Graphics'. *Cancer* 113 (12), 3382-3390

Zuckerman, M. and Kuhlman, D. M. (2000) 'Personality and risk‐taking: Common Bisocial Factors'. *Journal of Personality* 68 (6), 999-1029

7. APPENDICES

APPENDIX1: INFORMED CONSENT

<u>Informed Consent Form</u>

Project Title: Gamification and Interactive approaches impact on influencing rational decision making in a risk elicitation task

To collect these data for this research, we need to have your consent. By agreeing below, you confirm that:

Please initial

1. I have read and understood the participant information sheet for the above study and have had the opportunity to ask questions ☐
2. I understand that my participation is voluntary and that I am free to withdraw at any time without giving a reason ☐
3. I understand that all the information I provide will be treated in confidence ☐
4. I understand that I also have the right to change my mind about participating in the study for a short period after the study has concluded ☐
5. I am above 18 years old ☐
6. I know that you are going to use these data for research purposes. ☐
7. I agree to take part in the research project ☐

Date: Name of Researcher: Signature of researcher:

Appendix 2: Numeracy Scale

Could you please answer the following questions?

1) Which of the following percentages represents the biggest risk of getting a disease?

1%   10%   5%

2) Which of the following frequencies represents the biggest risk of getting a disease?

1 in 100   1 in 1000   1 in 10

3) If the chance of getting a disease is 20 out of 100, this would be the same as having a _______% chance of getting the disease

4) If the chance of getting a disease is 10%, how many people would be expected to get the disease out of 100 people

5) Imagine that we roll a fair, six-sided die 1,000 times. Out of 1,000 rolls how many times do you think the die would come up as an even number?

6) If the chance of getting a disease is 10%, how many people would be expected to get the disease out of 1000 people:

7) In the BIG BUCKS LOTTERY, the chances of winning a $10.00 prize are 1%. What is your best guess about how many people would win a $10.00 prize if 1,000 people each buy a single ticket from BIG BUCKS?

8) If Tom's risk of getting a disease is 1% and Peter's risk is double that of Tom's, what is Peter's risk?

9) In the ACME PUBLISHING SWEEPSTAKES, the chance of winning a car is 1 in 1,000. What percent of tickets of ACME PUBLISHING SWEEPSTAKES win a car?

10) If Tom's chance of getting a disease is 1 in 100, and Peter's risk is double that of Tom's, what is Peter's risk?

11) The chance of getting a viral infection is .0005. Out of 10,000 people, about how many of them are expected to get infected?

APPENDIX 3: HOLT AND LAURY TASK

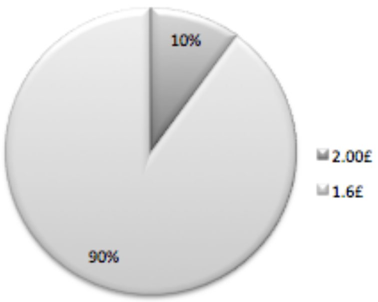
In this test, you will be presented with two lotteries, which are called option A and option B and you can choose between the two options in each row. You can start with row 1 and then go down from row to row. You can tick next to the option you choose in each row.

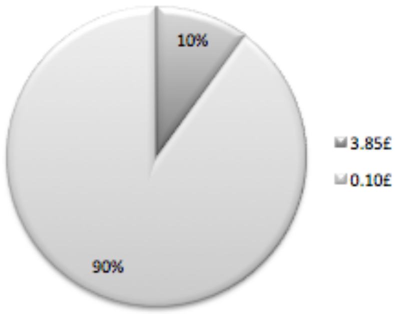
Row	**Option A**	**Option B**
1	1/10 of £2.00, 9/10 of £1.60	1/10 of £3.85, 9/10 of £0.10
2	2/10 of £2.00, 8/10 of £1.60	2/10 of £3.85, 8/10 of £0.10
3	3/10 of £2.00, 7/10 of £1.60	3/10 of £3.85, 7/10 of £0.10
4	4/10 of £2.00, 6/10 of £1.60	4/10 of £3.85, 6/10 of £0.10
5	5/10 of £2.00, 5/10 of £1.60	5/10 of £3.85, 5/10 of £0.10
6	6/10 of £2.00, 4/10 of £1.60	6/10 of £3.85, 4/10 of £0.10
7	7/10 of £2.00, 3/10 of £1.60	7/10 of £3.85, 3/10 of £0.10
8	8/10 of £2.00, 2/10 of £1.60	8/10 of £3.85, 2/10 of £0.10
9	9/10 of £2.00, 1/10 of £1.60	9/10 of £3.85, 1/10 of £0.10
10	10/10 of £2.00, 0/10 of £1.60	10/10 of £3.85, 0/10 of £0.10

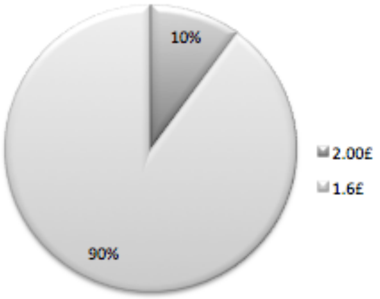
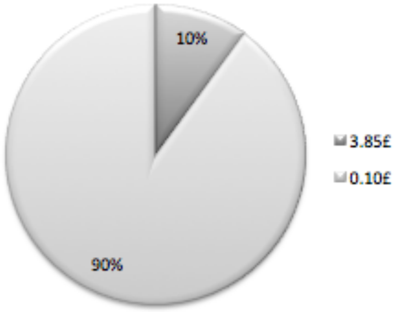
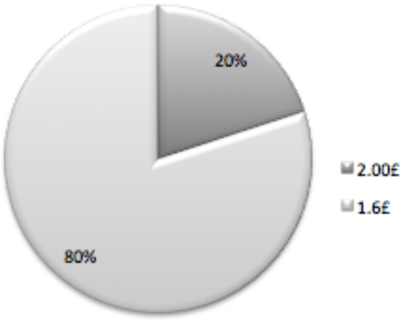
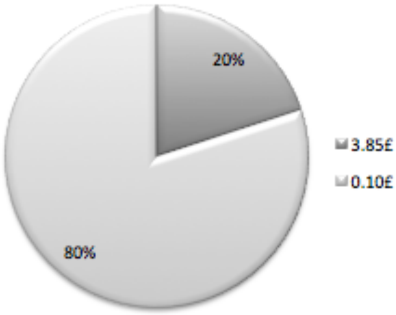
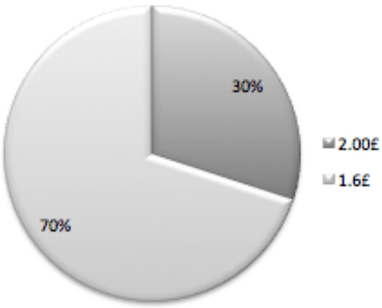
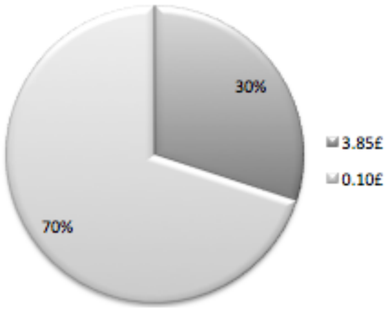
APPENDIX 4: HOLT AND LAURY TASK DISPLAYED WITH PIE CHARTS
(NOT ALL 10 ROWS ARE PRESENTED HERE)

In this test you will be presented with two lotteries represented as pie charts, which are called option A and option B and you can choose between the two options in each row. You can start with row 1 and then go down from row to row. You can tick inside the box of the option you prefer.

Example:

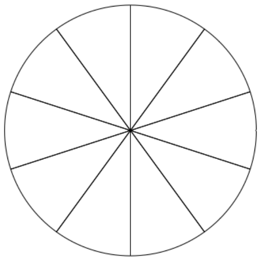
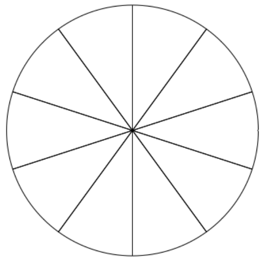
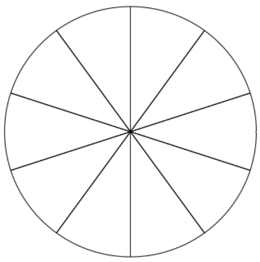
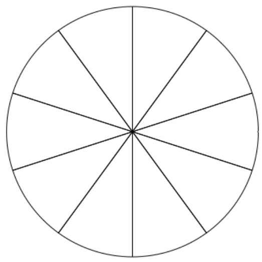
<table>
<tr><td>row</td><td>Option A</td><td>Option B</td></tr>
<tr><td>1</td><td>

✔</td><td>
</td></tr>
</table>

row	Option A	Option B
1		
2		
3		

Appendix 5: Holt and Laury task displayed with interactive pie charts
(Not all 10 rows are presented here)

In this test, you will be presented with two lotteries, which are called option A and option B. Please paint each pie with the proportions written above them e.g. for the proportion 1/10 please shadow one slice of the pie chart. After you shadows each pie of the options then choose which option (A or B) you would prefer in ten rows.

Row	Option A	Option B
1	1/10 of £2.00 and 9/10 of £1.60 	1/10 of £3.85 and 9/10 of £0.10 
2	2/10 of £2.00 and 8/10 of £1.60 	2/10 of £3.85 and 8/10 of £0.10 

APPENDIX 6: ETHICAL APPLICATION

PROJECT SUMMARY OF ETHICAL CERTIFICATE P48719:

The project aims to improve an already existing risk elicitation task developed by Holt and Laury (HL) 2002 by replacing the numerical probabilities in the measure with interactive visual aids. The problem with the standard HL task is that the presentation of probabilities confuses people especially those with low numerical skills and lead them to choose inconsistently in the task producing noisy data and exhibiting wrong risk attitude (Galarza 2009; Andersen et al. 2006). Empirical studies have shown that using either pie charts to represent the probabilities in the lottery task of HL has the potential of reducing people's inconsistent choices and increase the accuracy of the tasks' risk predictions (Habib et al. 2016; Eckel, Warnick and Johnson 2005). However, past research has shown that sometimes pie charts were shown to be problematic for people. On another note, Hegarty and Kozhevnikov (1999) identified that when participants use graphs (instead of passively display) have high rate of success. Therefore, we propose a study that explores whether people's choices will be more consistent if they are asked to draw the pie charts themselves.

APPENDIX 7: ETHICAL CERTIFICATE 1

REGISTRY RESEARCH UNIT

1 ETHICS REVIEW FEEDBACK FORM (P1698)

(Review feedback should be completed within 10 working days)

Name of applicant: Stella Doukianou ....

Faculty/School/Department: [Faculty of Engineering and Computing] Computing & The Digital Environment...

Research project title: Risk and patience assessment

Comments by the reviewer

4. Evaluation of the ethics of the proposal: Data is anonymised and consent forms are now appropriate for the format.
5. Evaluation of the participant information sheet and consent form: .
6. Recommendation: (Please indicate as appropriate and advise on any conditions. If there any conditions, the applicant will be required to resubmit his/her application and this will be sent to the same reviewer).

X	Approved - no conditions attached
	Approved with minor conditions (no need to re-submit)
	Conditional upon the following – please use additional sheets if necessary (please re-submit application)
	Rejected for the following reason(s) – please use other side if necessary
	Not required

Name of reviewer: Anonymous ....

Date: 10/02/2014 ....

APPENDIX 8: ETHICAL CERTIFICATE 2

REGISTRY RESEARCH UNIT

2 ETHICS REVIEW FEEDBACK FORM

(Review feedback should be completed within 10 working days)

Name of applicant: Stella Doukianou ....

Faculty/School/Department: [Faculty of Engineering and Computing] Computing & The Digital Environment..

Research project title: Risk assessment P30737

Comments by the reviewer

7. Evaluation of the ethics of the proposal: Data is anonymised and consent forms are now appropriate for the format.	
8. Evaluation of the participant information sheet and consent form: .	
9. Recommendation: (Please indicate as appropriate and advise on any conditions. If there any conditions, the applicant will be required to resubmit his/her application and this will be sent to the same reviewer).	
X	Approved - no conditions attached
	Approved with minor conditions (no need to re-submit)
	Conditional upon the following – please use additional sheets if necessary (please re-submit application)
	Rejected for the following reason(s) – please use other side if necessary
	Not required

Name of reviewer: Anonymous ....

Date: 09/04/2015 ....

APPENDIX 9: EXPERIMENTAL PLAN AND DATA FOR EXPERIMENT 1

Experimental Plan:

-Summary description of the system: A gamified assessment task Risk House game. The game is a risk elicitation task based on the Holt and Laury lottery task.

-Targeted user and goals of the experiment:
Targeted users are the general public. The goal of the experiment is to test whether the game is user friendly and there would be no usability issues affecting their performance.

Description of the usability study

-Characteristics of the participants*:* Participants required to be Computer Experts above 18 years old.

-Measures: The measures include the game and QUIS questionnaire.
Design of the experiment: In the experiment, all the participants are assigned to play the game and fill in the questionnaire

- Procedures: Invitation to participants for the date of the experiment. After they accepted the invitations, participants visited the room where the experiment would take place. After they read the Participant Information Sheet, were asked to sign the Informed Consent. After the Informed Consent was collected, participants were asked to play the game and then fill in the QUIS questionnaire. When the participants completed filling up the QUIS, the questionnaire was collected. Next temp involved the import of QUIS data into SPSS to identify the usability level of the game and which areas needed further investigation.

QUIS DATA

Data organisation in SPSS: The data were collected from the questionnaires (QUIS) that participants had to fill in and they were imported into SPSS software. Each section included several questions that were named into SPSS Q1, Q2… depending on the number of the question. Each Section (5 in total) was imported into SPSS separately to avoid confusions. At the end means of min, max, mean and SD, were calculated and reported at Table 4.

Age									
P1	**P2**	**P3**	**P4**	**P5**	**P6**	**P7**	**P8**	**P9**	**P10**
18	18	30	18	29	33	28	24	26	24

Overall Reactions to the game					
Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**
5	6	5	7	6	6
7	7	7	6	7	7
7	7	7	7	7	7
7	7	6	7	7	7
7	6	6	6	6	6
7	7	7	7	7	5
7	7	7	7	5	6
5	5	7	7	5	6
7	6	6	7	6	7
8	8	7	7	7	7

Screen					
Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**
6	6	5	7	6	7
7	6	7	7	8	7
7	7	7	8	8	8
7	6	7	8	7	8
6	6	5	5	6	6
7	7	7	7	8	7
7	7	5	8	6	6
6	6	7	6	6	6

Terminology and system information							
Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**	**Q7**	**Q8**
7	7	7	7	8	8	8	8
7	8	7	8	8	8	7	7
8	8	7	7	8	8	8	7
8	7	7	8	8	8	7	7
7	7	7	7	8	8	8	8
7	7	7	7	8	8	8	8
7	8	7	8	8	8	8	7
8	7	7	7	8	8	8	8
7	7	7	7	8	8	8	9
8	8	8	8	8	8	8	8

Learning										
Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11
6	6	6	7	7	7	7	7	7	7	6
8	7	8	8	8	7	7	8	7	7	6
8	7	8	8	8	8	7	7	7	8	8
8	7	8	7	8	7	7	7	7	7	6
6	6	6	7	7	7	7	7	8	7	6
8	8	9	7	7	8	7	7	8	7	6
7	8	8	7	7	7	8	8	8	7	6
6	6	6	7	7	7	7	8	8	7	6
8	7	6	8	7	7	7	8	8	7	6
8	8	8	7	8	9	7	8	7	8	7

System Capabilities										
Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11
6	6	6	7	7	7	7	7	7	7	6
8	7	7	7	8	7	7	7	7	7	6
6	6	8	8	8	8	7	7	7	8	8
8	7	9	7	7	7	7	7	7	7	6
6	6	6	7	7	7	7	7	8	7	6
7	8	7	7	7	8	7	7	7	7	6
7	8	8	7	7	7	8	7	8	7	6
6	6	6	7	7	7	7	7	8	7	6
8	7	6	8	7	7	7	7	8	7	6
8	8	8	7	8	9	7	7	7	8	7

CONSTRUCT VALIDITY

Experimental Plan:

Summary description of the intervention: A gamified assessment task Risk House game. The game is a risk elicitation task based on the Holt and Laury lottery task.

-Targeted user and goals of the experiment:

Targeted users are the general public. The goal of the experiment is to validate that Risk House game can assess risk taking and that there would be no participants' irrational choices.

Description of the pilot study

-Characteristics of the participants: Participants required to be above 18 years old from Coventry University

-Null Hypothesis (H_0): The choices of participants in gamified risk assessment risk assessment, Risk House game, would not significantly correlate with their choices at the Balloon Analogue Risk Task.

H_{01}: The choices of participants in gamified risk assessment, Risk House game, would be significantly correlated with their risky choices at the Balloon Analogue Risk Task.

-Variables: Participants' choices in the game (Risk choices) and participants' choices in BART (Pumps in total and explosions).

-Measures: The measures include participants' risky choices in the game and participants' choices in BART.

Design of the experiment: In the experiment, all the participants are assigned separate times to play the game and BART.

Procedures: Emails were sent to 50 participants. Participants were gathered through snowball effect. From 60 emails in total, 32 participants accepted the invitations and arranged to do the experiment on different timings. Each participant came to the room were the experiment was conducted by appointment. As soon as the participant entered the room, the Participant Information Sheet was given to him and the Informed Consent to sign. After the collection of the Informed Consent, the participant was asked to play the Risk House game and the BART task. The order which each participant played the Risk House game and the BART task was random to avoid order effects. At the end of each participant's session, he received a payment according to performance to BART task and the debrief sheet about the study. The same setting in the experiment was used for each subject to avoid setting effects.

Data organisation in SPSS: Both BART and the Risk House Game, involved backend mechanisms that recorded participants' choices. Therefore, for Risk House game, the number of risky choices was recorded and in BART, the number of explosions and number of pumps for each balloon. These raw data were collected and imported into SPSS without changing anything to calculate the correlation coefficient. Only gender was coded, 0 for males and 1 for females. The way these data were organised in SPSS is shown in the following table.

BART pumps in total	BART Explosions	Game risky choices	Gender	Age
839	9	12	1	22
483	5	11	1	28
903	10	11	1	27
689	6	10	0	22
403	3	9	0	33
902	10	11	0	19
323	3	7	1	22
614	6	18	0	19
475	5	10	1	18
619	6	14	0	26
832	9	11	1	19
702	7	8	1	22
620	7	12	1	21
835	7	10	0	27
425	4	11	0	23
155	1	13	0	20
881	10	11	1	22
978	9	8	1	18
777	8	11	1	24
683	7	16	0	33
781	5	17	1	20
527	4	8	0	32
759	7	12	1	21
608	5	11	0	30
198	1	8	1	31

574	5	9	1	20
628	5	10	0	22
664	8	9	1	32
813	9	12	1	22
717	7	20	1	22
493	5	15	0	20
750	7	16	0	25

APPENDIX 10: EXPERIMENTAL PLAN AND DATA FOR EXPERIMENT 2

Experimental Plan:

Summary description of the system: A choice between two lotteries from the Holt and Laury task that participants would be required to reason it by using external representations.

-Targeted user and goals of the experiment:
Targeted users are university and non-university graduates. The goal of the experiment is to explore which is the most principal influential factor of participants' irrational choices in the Holt and Laury task and whether using external representation would reduce them.

Description of the exploratory study:

-Characteristics of the participants: Participants required to be above 18 years old from general public.

- *Null Hypothesis H_0:* Participants' education, score in the validated numeracy scale, cognitive reflection task and use of external representation would not influence their rational decision making in the Holt and Laury option.
H_{02}: Participants' education, score in the validated numeracy scale, score in cognitive reflection task and use of external representation would influence their rational decision making in the Holt and Laury option.

-Variables*:* Independent are education, score in numeracy scale, score in the cognitive reflection task and external representations. The dependent variable is participants' decision in Holt and Laury option.

-Measures: The measures include their educational level whether they have graduated from university or not, their score in Lipkus Numeracy Scale, their score in the CRT, their choice in the Holt and Laury and the external representation they used for it.

Design of the experiment: In the experiment, all the participants are assigned separate times to fill in the questionnaires.

Procedures: Invitations were sent to 80 participants. Out of 80, 60 participants accepted the invitation. Each participant separately was invited to the room where the experiment was conducted by appointment. As soon as the participant entered the room, the Participant Information Sheet was given to him and the Informed Consent to sign. After the collection of the Informed Consent, the participant was asked to do the following: fill in the Numeracy test, the CRT, the example (probability problem) and the Holt and Laury option. The order that these tasks were given to him (apart from the example problem and Holt and Laury that were given always in that order) was random. All this procedure was observed by the researcher of this study. There were no interventions by the researcher apart from the example problem where the participant had to use external representation to reason his solution. The term external representations and example of them were explained to the participant by the researcher. Following that, the researcher asked the participant if he understood the procedure and prompted him to use anything from visuals to numbers to solve the problem. After he finished the example, the participant was given the Holt and Laury and he was asked to do the same as the example problem, using any external representation to reason his choice. After the completion of all tasks another document was given to him to fill in his age and gender. At the end, the participant received the debrief sheet of the study and the next participant got in the room. The setting of the experimental room remained the same for all the participants.

Coding Schemes				
Education	**Holt and Laury**	**Numerical test**	**CRT**	**External Representations**

0 - University Graduates 1 - Non-University Graduates	0 – Rational 1- Irrational	0 - 9,10,11 Most Numerate 1- 7 < Less Numerate	0 -2,3 Reflective 1- 0,1 Impulsive	1 - Numeracy 2 - Graphs 3- Blank page 4 - Words 5 - Pictures

Data organisation in SPSS: The data collected from this experiment were coded according to the Types of Representations (p.81) with 1 – Numeracy, 2 – Graphs, 3 – Blank Page, 4 – Words, 5 – Pictures and then imported them into SPSS. The table below, shows the categorisation of the data into SPSS.

Education	Holt and Laury	Numerical test	CRT	External Representations	Gender	Age
0	0	0	0	1	2	23
0	0	0	1	4	2	23
0	1	0	0	5	2	24
0	0	0	0	2	2	30
0	0	0	0	2	2	30
0	0	0	1	2	1	47
0	0	0	0	2	1	26
0	0	1	0	2	1	26
0	0	0	1	2	2	28
0	0	0	1	2	2	32
0	0	0	1	2	1	26
0	1	1	0	1	1	27
0	1	1	0	2	2	33
0	0	0	1	2	2	33
0	1	1	0	1	1	24
0	0	0	0	2	1	29
0	0	0	1	2	1	31
0	0	0	1	2	2	23
0	0	0	0	1	2	n/a
0	0	0	1	2	2	26
0	0	1	1	2	2	26
0	0	0	0	4	1	28

0	0	0	1	1	2	37
0	1	1	0	5	1	25
0	0	1	0	3	2	30
0	0	0	1	2	2	40
0	0	0	1	2	1	26
0	0	0	1	2	2	36
0	0	1	0	3	2	33
0	0	0	0	1	1	26
0	0	1	1	2	1	30
0	0	1	1	2	1	30
0	0	0	1	2	2	33
0	0	0	0	2	2	29
0	1	0	0	1	2	53
1	0	0	0	2	2	20
1	1	0	0	2	2	20
1	0	1	0	1	2	21
1	0	0	1	2	2	21
1	0	0	1	5	2	20
1	0	0	1	1	2	20
1	0	1	1	2	1	19
1	0	0	1	2	2	36
1	0	0	1	2	2	20
1	0	0	1	2	2	19
1	1	0	0	2	2	22
1	1	0	0	5	2	21
1	1	1	0	2	2	22
1	0	0	0	1	1	21
1	0	0	1	4	2	21
1	0	1	1	1	1	18
1	0	1	1	4	1	18
1	1	1	0	5	1	18
1	0	0	0	2	2	18
1	0	0	1	1	2	19
1	0	0	1	1	2	18
1	0	0	1	2	2	18
1	0	0	1	1	2	18
1	1	0	0	2	2	27
1	0	0	1	1	2	21

APPENDIX 11: EXPERIMENTAL PLAN AND DATA FOR EXPERIMENT 3

Experimental Plan:

Summary description of the system: A Holt and Laury task where users are prompted to draw a pie chart of the options in each lottery to reason their choices.

-Targeted user and goals of the experiment:
Targeted users are general public and especially people with limited numerical skills. The goal of the experiment is to investigate whether using pie charts to reason choices in the Holt and Laury task would reduce participants irrational choices compared to be presented with numerical display Holt and Laury task or passive pie chart display Holt and Laury task.

Description of the study:

-Characteristics of the participants: Participants required to be above 18 years old.

- ***Null Hypothesis (H_0):*** Individuals that score lower in the validated numeracy scale when they interact with the pie charts in Holt and Laury task would make not significantly more rational choices compared to when they are presented the Holt and Laury task with passive pie charts or numerical probabilities.
H_{03}: Individuals that score lower in the validated numeracy scale when they interact with the pie charts in Holt and Laury task will make significantly more rational choices compared to when they are presented the Holt and Laury task with passive pie charts or numerical probabilities.

- *Variables:* Dependent variable is participants' rational choices and the independent variables are the three different display formats are presented with (Numerical, passive pie chart and interactive pie chart).

-Measures: The measures include their score in Lipkus Numeracy Scale, and whether they filled in the Holt and Laury rationally or not.

Design of the experiment: In the experiment, different participants are randomly assigned to each of the display format of Holt and Laury.

Procedures: Invitations were sent to the Lecturers that have the biggest number of students in the class in the Department of Computing. A Lecturer with 250 students accepted the invitation to run the experiments before his Lecture. The experiment took place in a lecturer class where 225 students before the class started. Before the participants were divided randomly into groups, they were asked to read the Participant Information sheet and then sign the Informed Consent. When the informed consent was collected, they were divided randomly into 3 groups. One group would receive the numerical display Holt and Laury, the second group would have the passive pie chart Holt and Laury and the third group would have the interactive pie chart Holt and Laury. Sealed envelopes with one of the three display formats Holt and Laury, the Numeracy scale and a demographics questionnaire was given to all the participants. Participants were asked to fill in the tasks in the order they were in the envelope; first the Holt and Laury and then the numeracy scale and the demographic questions. As soon as they finished the envelope was collected and the participant could each participant selected a small note from a lottery ball, all notes included numbers except from one that had the letter A and was referring to an Amazon Voucher of £50. There was no intervention from the researcher throughout the experimental procedure.

Data organisation in SPSS: The following table shows the categorisation of the data into SPSS.

Coding Schemes				
Consistency	**Numerical test**	**Age**	**Gender**	**Perceived Difficulty**
0 – Consistent 1 - Inconsistent	0 - 9,10,11 Most Numerate 1- 7 < Less Numerate	0 = 18-22 1 = 23-27 2 = 28-32 3 = 32 +	0 – Male 1 - Female	1 – Very easy 2 – Barely difficult 3- Somewhat difficult 4 – Mostly difficult 5 – Very difficult

NumericalHL	PiechartHL	InteractivePiechartHL	Age_Numerical	Age_Piechart	Age_interactive	Gender_Numeric	Gender_Piechart	Gender_Interacti	Difficulty_Numerical	Difficulty_Piech	Difficulty_Interactive	Numerical_Numerical	Numerical_Piech	Numerical_Interactive	Allgrou	Gender	Difficult	Numericaltest	Age
0	0	0	0	0	0	0	0	1	2	4	2	0	0	0	1	0	2	0	0
1	0	0	3	0	1	0	0	1	5	2	1	1	0	1	1	0	5	1	3
1	1	0	0	0	3	0	0	1	4	5	1	1	1	1	0	0	4	1	0
0	1	0	0	0	1	1	1	1	1	4	5	0	1	0	0	1	1	0	0
1	1	0	0	0	1	0	0	1	2	3	1	0	1	1	0	0	2	0	0
1	0	0	0	0	1	0	1	1	2	3	3	1	0	1	1	0	2	1	0
0	1	0	1	0	0	0	1	0	2	3	1	0	0	0	0	0	2	0	1
0	0	0	0	0	0	0	0	0	1	4	2	0	1	0	0	0	1	0	0
0	1	0	0	1	1	1	1	0	5	4	2	0	1	0	0	1	5	0	0
0	0	0	1	1	1	0	1	0	2	3	2	0	0	0	0	0	2	0	1
0	0	0	0	0	1	0	0	0	3	3	1	0	1	0	0	0	3	0	0
0	0	0	0	1	0	0	1	0	1	3	1	0	1	0	0	0	1	0	0
0	0	0	0	1	0	0	0	0	2	3	1	0	0	0	1	0	2	0	0
1	0	0	0	0	0	0	0	0	4	3	1	1	1	1	0	0	4	1	0
0	0	1	1	0	0	0	0	1	2	3	2	0	1	1	0	0	2	0	1
1	0	0	2	0	1	1	0	0	4	2	1	1	1	0	0	1	4	1	2
0	1	0	0	0	1	0	1	0	3	4	1	0	1	0	0	0	3	0	0
0	1	0	3	2	0	0	0	0	3	4	3	0	0	0	1	0	3	0	3
1	0	0	0	1	0	0	0	0	5	3	2	1	0	1	0	0	5	1	0
1	1	0	0	0	0	0	0	0	5	1	1	1	0	1	0	0	5	1	0
0	1	1	0	0	0	0	0	0	4	4	2	0	1	1	0	0	4	0	0
0	0	0	1	1	1	0	0	0	4	2	2	0	1	1	1	0	4	0	1
0	0	0	1	0	1	0	0	0	2	2	2	0	1	1	0	0	2	0	1

0	1	0	1	0	0	0	0	1	4	3	2	0	1	0	1	0	4	0	1
1	0	0	0	1	0	0	0	0	3	3	1	0	0	1	0	0	3	0	0
0	1	0	0	0	1	0	1	1	3	2	1	0	0	1	1	0	3	0	0
1	1	0	0	0	0	0	0	0	5	3	1	1	1	1	0	0	5	1	0
0	1	0	0	0	0	0	0	0	3	2	1	0	0	1	0	0	3	0	0
0	0	0	1	0	0	0	0	0	-	2	1	0	0	0	0	0	-	0	1
0	1	0	1	1	0	1	0	1	2	3	2	0	1	1	0	1	2	0	1
0	0	0	0	0	0	0	1	0	-	2	2	0	0	0	0	0	-	0	0
1	1	0	0	0	0	1	0	0	3	3	-	1	1	1	1	1	3	1	0
1	0	0	1	1	0	1	0	0	4	2	-	1	0	1	1	1	4	1	1
1	0	1	1	0	0	1	0	0	4	-	5	1	1	0	0	1	4	1	1
1	0	1	1	0	0	1	0	1	5	-	4	1	1	0	1	1	5	1	1
1	0	0	0	0	1	1	0	0	4	3	1	1	0	1	0	1	4	1	0
1	0	0	0	0	0	0	0	0	4	3	2	1	1	0	1	0	4	1	0
0	0	0	0	0	1	0	0	0	2	2	1	0	1	0	0	0	2	0	0
0	0	0	0	1	0	0	0	0	4	3	2	0	1	1	0	0	4	0	0
1	1	0	0	0	0	0	1	0	3	3	2	1	1	0	0	0	3	1	0
0	0	0	0	0	0	0	0	0	2	3	2	0	0	0	0	0	2	0	0
1	1	0	0	0	0	0	1	0	3	2	1	1	0	0	0	0	3	1	0
0	0	0	0	0	0	1	1	1	1	5	2	0	0	1	0	1	1	0	0
1	1	0	0	0	0	0	0	1	4	2	2	1	0	0	0	0	4	1	0
1	1	0	0	0	0	1	0	0	5	4	2	1	0	1	1	1	5	1	-
1	0	0	0	0	0	0	0	1	4	3	2	0	1	1	0	0	4	0	0
1	0	0	0	0	0	0	1	0	4	4	2	0	0	0	1	0	4	0	0
1	0	0	0	0	0	0	0	1	3	2	-	1	0	1	0	0	3	1	0
1	0	0	1	1	0	0	0	0	5	1	1	1	0	0	1	0	5	1	1
1	0	1	0	0	0	1	0	0	4	3	3	1	1	1	1	1	4	1	0
1	1	0	0	0	0	0	0	1	4	3	2	1	0	1	1	0	4	1	0
1	0	0	0	0	0	0	0	0	5	5	2	0	0	1	0	0	5	0	0
1	1	0	1	0	0	1	0	1	3	2	3	1	1	1	1	1	3	1	1
1	0	0	0	1	0	0	0	1	5	3	5	0	0	0	0	0	-	0	0

1	1	0	0	1	3	1	1	1	3	4	4	1	1	0	0	1	3	1	0
1	0	0	0	0	1	0	1	0	5	1	3	1	0	1	1	0	5	1	0
0	0	0	2	0	0	1	0	0	1	2	2	0	0	1	1	1	1	0	2
0	0	0	0	0	1	0	1	1	4	5	2	0	1	1	0	0	4	0	0
1	1	0	0	0	0	0	1	0	5	5	1	1	0	0	1	0	5	1	0
0	0	0	0	0	0	0	0	-	1	2	1	0	0	1	1	0	1	0	0
1	0	0	0	0	1	0	0	1	5	2	1	1	0	0	0	0	5	1	0
1	0	0	0	0	0	0	0	0	5	3	1	0	0	1	0	0	5	0	0
0	0	0	0	0	0	0	0	0	1	2	1	0	0	0	0	0	1	0	0
0	1	0	3	0	1	1	0	0	2	3	2	0	0	1	0	1	2	0	-
0	0	0	0	0	0	0	0	0	-	2	2	0	0	1	0	0	-	0	0
0	0	0	0	0	0	0	1	1	2	2	1	0	0	1	0	0	2	0	0
0	0	0	0	1	0	0	0	0	2	3	2	0	1	1	0	0	2	0	0
1	0	0	0	1	0	0	1	0	5	5	2	1	1	1	1	0	5	1	0
0	0	0	1	0	0	0	1	1	3	3	1	0	0	0	0	0	3	0	1
0	0	0	0	0	1	1	0	1	2	1	1	0	0	1	1	1	2	0	0
0	1	0	1	0	0	0	1	0	3	4	2	1	1	0	0	0	3	1	1
0	0	0	0	0	1	0	0	1	2	2	1	0	1	1	1	0	2	0	0
0	0	0	1	0	0	0	1	0	2	1	2	1	1	1	1	0	2	1	1
0	0	-	0	0		1	0		1	2		0	0		1	1	1	0	0
0	1		3	0		0	0		4	2		0	1		0	0	4	0	-
1	1		0	0		0	0		2	3		1	1		0	0	2	1	0
															0	0	4	0	0
															0	0	2	0	0
															1	0	5	1	0
															1	1	4	1	0
															1	0	3	1	0
															0	1	3	0	-
															1	1	3	0	0
															0	0	4	1	0

															1	1	4	1	1
															0	1	3	0	1
															0	0	3	1	0
															0	1	3	1	1
															0	0	3	0	1
															0	0	3	1	0
															0	0	3	1	0
															0	0	2	1	0
															1	1	4	1	0
															1	0	4	0	2
															0	0	3	0	1
															1	0	1	0	0
															1	0	4	1	0
															0	0	2	1	1
															0	0	2	1	0
															1	0	3	1	0
															0	0	3	0	1
															1	1	2	0	0
															1	0	3	1	0
															1	0	2	0	0
															0	0	2	0	0
															1	0	3	1	1
															0	1	2	0	0
															1	0	3	1	0
															0	0	2	0	1
															0	0	-	1	0
															0	0	-	1	0
															0	0	3	0	0

															0	0	3	1	0
															0	0	2	1	0
															0	0	3	1	1
															1	1	3	1	0
															0	0	3	0	0
															1	1	2	0	0
															0	1	5	0	0
															1	0	2	0	0
															1	0	4	0	0
															0	0	3	1	0
															0	1	4	0	0
															0	0	2	0	0
															0	0	1	0	1
															0	0	3	1	0
															1	0	3	0	0
															0	0	5	0	0
															1	0	2	1	0
															0	0	3	0	1
															1	1	4	1	1
															0	1	1	0	0
															0	0	2	0	0
															0	1	5	1	0
															1	1	5	0	0
															0	0	2	0	0
															0	0	2	0	0
															0	0	3	0	0
															0	0	2	0	0
															1	0	3	0	0

															0	0	2	0	0
															0	1	2	0	0
															0	0	3	1	1
															0	1	5	1	1
															0	1	3	0	0
															0	0	1	0	0
															1	1	4	1	0
															0	0	2	1	0
															0	1	1	1	0
															0	0	2	0	0
															1	0	2	1	0
															1	0	3	1	0
															0	1	2	0	0
															0	1	1	1	1
															0	1	1	1	3
															0	1	5	0	1
															0	1	1	1	1
															0	1	3	1	1
															0	0	1	0	0
															0	0	2	0	0
															0	0	2	0	1
															0	0	2	0	1
															0	0	1	0	1
															0	0	1	0	0
															0	0	1	0	0
															0	0	1	1	0
															1	1	2	1	0
															0	0	1	0	1

															0	0	1	0	1
															0	0	3	0	0
															0	0	2	1	0
															0	0	1	1	0
															0	0	2	1	0
															0	0	2	1	1
															0	0	2	1	1
															0	1	2	0	0
															0	0	1	1	0
															0	1	1	1	1
															0	0	1	1	0
															0	0	1	1	0
															0	0	1	0	0
															0	1	2	1	0
															0	0	2	0	0
															0	0	-	1	0
															0	0	-	1	0
															1	0	5	0	0
															1	1	4	0	0
															0	0	1	1	1
															0	0	2	0	0
															0	0	1	0	1
															0	0	2	1	0
															0	0	2	0	0
															0	0	2	0	-
															0	0	1	0	-
															0	1	2	1	0
															0	1	2	0	0

															0	0	2	1	0
															0	1	2	1	0
															0	0	2	0	0
															0	1	-	1	-
															0	0	1	0	0
															1	0	3	1	0
															0	1	2	1	0
															0	0	2	1	0
															0	1	3	1	0
															0	1	5	0	0
															0	1	4	0	3
															0	0	3	1	1
															0	0	2	1	0
															1	1	2	1	1
															0	0	1	0	0
															0	1	1	1	-
															0	1	1	0	1
															0	0	1	1	0
															0	0	1	0	0
															0	0	2	1	1
															0	0	2	1	0
															0	1	1	1	0
															1	0	2	1	0
															0	0	2	1	0
															0	1	1	0	0
															0	1	1	1	1
															0	0	2	0	0
															0	1	1	1	1

															0	-	2	1	0