Title
Reinforcing Rational Decision Making in a Risk Elicitation task through Visual Reasoning

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 41(0)

Authors
Doukianou, Stella
Daylamani-Zad, Damon
Lameras, Petros
et al.

Publication Date
2019

Peer reviewed

Reinforcing Rational Decision Making in a Risk Elicitation task through Visual Reasoning

Stella Doukianou (S.Doukianou@gre.ac.uk)
Damon Daylamani-Zad (D.D.Zad@gre.ac.uk)
School of Computing and Mathematical Sciences, University of Greenwich,
Park Row, London, SE10 9LS, U.K.

Petros Lameras (PLameras@cad.coventry.ac.uk)
Ian Dunwell (IDunwell@cad.coventry.ac.uk)
School of Computing, Electronics and Maths, Coventry University, Priory St,
Coventry CV1 5FB, U.K.

Abstract

Metrics seeking to predict financial risk-taking behaviors typically exhibit limited validity. This is due to the fluid nature of an individual's risk taking, and the influence of the mode and medium, which presents a decision. This paper presents two experiments that investigate how an existing risk elicitation task's predictive capacity may be enhanced through the application of an interactive model of visual reasoning in a digitized version. In the first experiment, 60 participants demonstrated their reasoning process. In the second experiment, 225 participants were randomly assigned into three groups, with the validated risk elicitation task compared as a control to interactive digital and non-interactive digital stimuli with pie charts. The experiments yielded significant results, highlighting that when participants interact with a graph to reason their choices, it leads to consistent choices. The findings have implications for improvement of the risk task's validity and the deployment of digital interactive assessments beyond laboratory settings.

Keywords: visualization, decision-making, risk-taking, external representations, reasoning

Introduction

The ability to elicit the degree to which an individual or demographic is risk-taker or -averse has a value across various fields. Existing risk elicitation tasks have shown to have predictive capacity; however, in risk elicitation tasks that involve lotteries, a key constraint is participants' limited understanding of the probabilities, representing the risk associated with each choice, in those tasks. If participants do not understand a probability-based question, their answers lack internal consistency. Whilst within these tests internal checks for validity exist, these alone allow only for the exclusion, rather than accommodation, of participants with low numeracy skills that have limited understanding of the probabilities presented in the task. As a result, findings can be skewed towards a subset of a larger sample, limiting validity and predictive capacity.

Towards resolving this issue, this paper presents the findings of two empirical studies that were conducted to investigate the use of visualization reasoning methods as an assistive tool for users to understand the probability described and choose consistently in a risk elicitation task. Experiment 1 sought to explore the most effective reasoning methods used by 60 participants in a risk elicitation task, by asking participants to illustrate their thought-process. Visual reasoning was identified as having the strongest positive effect, among all external representations used by the participants, on the consistency of their choices in the risk elicitation task (Holt & Laury, 2002), showing that it helped them understand the probabilities in the risk task. To confirm whether visual reasoning, which can be defined as using visuals to reason a probability problem, on risk elicitation task can help participants with low numeracy level, Experiment 2 translated the existing risk-elicitation task to two versions; a non-interactive visual digital version and an interactive visual digital version, in addition to the standardised (control) numerical format task. Our discussion and conclusions reflect on the relevance of the findings in terms of increasing the accessibility and meaningfulness of risk-elicitation tasks to less numerate participants by using visual reasoning processes and also the implications for the use of digital and interactive visual media in place of standardised paper based tasks.

Background

Methods of risk-elicitation tasks can be broadly categorized into self-report questionnaires describing hypothetical situations; hypothetical choice problems; or computerized methods (Rohrmann, 2005). A range of moderating variables have been observed to affect the validity of the majority of these methods, leading to inaccurate results and predictions (Andersen et al., 2006; Dave et al., 2010). This can limit the scalability of an elicitation-exercise, and relates to a fundamental challenge in transferability of results to different contexts: any psychometric tool seeking to establish or predict behavior must consider the fluidity of individual characteristics, and furthermore how a slight change may result in a meaningful change in decision-making.

The Multiple Price List (MPL) task belonging to the category of hypothetical choice problems, wherein participants need to choose between a 'safe' or 'risky' bet over ten different lotteries, has demonstrated predictive capability and can be implemented straightforwardly (Andersen et al., 2006; Dave et al., 2010; Rohrmann, 2005). Holt and Laury's variant of the MPL task has been examined in several studies (e.g.Nielsen, Keil, & Zeller, 2013; Dave et al., 2010), demonstrating significant predictive value but only for people with higher numerical skills who understand the probabilities for each of the parameters (Dave et al., 2010). The additional value of Holt Laury's MPL method is its capacity to identify the inconsistency rate in an individual's responses, allowing these choices to be excluded. However, this results in decreasing the validity of the metric.

To increase participants' understanding of the probabilities presented in the Holt and Laury MPL method, researchers have explored the use of visual display formats to represent the lotteries which reflect the losses and gains of the two options, allowing for more consistent and rational choices (Boughera, Gassmann & Piet, 2011; Bauermeister & Mußhoff, 2016; Habib et al., 2016). According to empirical evidence, using visualization tools to illustrate the uncertainty of the variables has a significant merit for people's informed decision making (Deitrick & Edsall, 2006). Integrating visualizations in the reasoning process could help users with low numerical skills to choose meaningfully regarding the risk related choices (Padilla et. al, 2018). Given that numeracy is an observed, reliable predictor of responses in risk elicitation tasks, a goal here is to address demographics or individuals who may have lower numeracy skills, but for whom a predictive mean of assessing their risk-taking or aversion holds value. Additionally, cognitive thinking style has been suggested to be another influential factor in a person's risk related choices (Frederick, 2005). Therefore, lower inconsistency rate achieved by presenting problems, which allows a wider range of individuals to respond consistently, would yield more reliable data. Research has illustrated the capacity of visualization to make probability reasoning more intuitive, and therefore better understood by participants (Hegarty & Kozhevnikov, 1999). In turn, this is shown to improve predictive validity. In the following section, visualization and external representations are briefly discussed.

Visualizations and external representations

Visualizations can become a reasoning process through interaction (Khan, Breslav & Hornbæk, 2018). For problems that involve probabilities, the individual relies on both internal visualization of the problem and the use of external representations such as sketches or diagrams to facilitate the solution. Visualizations serve as cognitive aids in problem solving situations (Khan, Breslav & Hornbæk, 2018; Padilla et al. 2018). However, individual differences, personality traits and cognitive abilities can have a significant effect on the use of a method that can aid the problem solving process (Ziemkiewicz et al., 2012; Gray & Holyoak, 2018). Therefore, visualization may not be the most appropriate external representation to assist in decision making for every individual (Starns et al., 2018). Even more, the type of visualization that can support the decision making for various tasks can differ significantly (Starns et al., 2018).

According to Corter and Zahner (2007) when investigating external representations, there is a division into two categories: the first refers to external visual presentations, which are provided towards influencing or informing a decision-making process (Corter, & Zahner 2007). The second involves understanding the internal representations the individual uses to reason when faced with a decision. The second involves the effects of user-generated visual representations while engaged in decision making and problem-solving activity. These internal representations, which can be defined as visual imagery (Corter & Zahner, 2007), have a core role in the production of knowledge. When externalized, they can provide valuable insight into individual's decision-making and aid successfully with problem solving (Gilbert, 2008). Empirical evidence suggests that using external representation to reason probabilities helps individuals to solve problems successfully (Corter & Zahner, 2007; Zhang 1997). In studies for logical reasoning, it has been argued that graphical representations allow the successful interpretation of abstract concepts (Zhang, 1997).

As the body of research suggests that external representations may be able to aid the reasoning process in probability problems, the following studies investigate the use of external representations to help individuals choose rationally in a risk elicitation task with lotteries. The first experiment is looking to determine whether there is a significant relationship between level of education, numeracy level and cognitive thinking style in rational decision-making in risk elicitation tasks. The second experiment investigates whether an interactive pie chart approach can influence the rational decision making in those tasks.

Experiment 1

Method

Participants Sixty volunteers from the UK participated in the study. The participants were divided into two groups depending on their educational level:

Group 1: Thirty-five participants (21 m, 14 f) with an age range between 22 and 53 (M=29.5) 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 a UK SME in the energy sector, and a British University.

Group 2: Twenty-five students from a further education college in the UK (4 f, 21 m), ranging in age 18-37 (M =20.64) volunteered to participate in the experiment and had not completed any degree-level qualification. The male participants outweighed significantly the number of females, this would be perceived as influencing the design and implementation of our study. All of them were assigned to the same experimental task as Group 1; similar research ethics procedures were also applied to this group. Participants were students in the Game Design and Web Design courses at a remedial programme. None of the participants had attended a course at a university level.

Measures The Lipkus Numeracy scale developed by Lipkus, Samsa and Rimer (2001) was used for this study. It was selected among others as a numeracy assessment tool for this study because it has been used in similar research to assess basic arithmetic skill in the variety of groups (Peters et al., 2006; Schapira et al., 2012). It is a short task, involving only

11 items, and consists of basic probability questions. These 11 items assess how well people can transform probabilities to percentages, percentages into probabilities while also performing simple mathematical operations using percentages or probabilities. The possible total sum scores range 0 to 11, where higher scores indicate better numerical skills compared to lower scores.

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. CRT reveals a reflective thinker over an impulsive as the most intuitive answer of the task is the wrong one. The individual needs to reflect before finding the solution. This measure is scored 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 were classified as high reflective thinkers (Frederick, 2005).

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. The probabilities change from row to row while the payoff remains the same. Hence, the expectation values of the two options change 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, determines 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. This measure was not used as a risk assessment. Rather, 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 and which external representation used by the participants would facilitate consistent choices.

Procedure Participants were given the participant information sheet informing them about the study and the informed consent form to sign. After the informed consents were obtained, participants started filling in the tasks using pen and paper. On average, each participant needed thirty minutes to complete all of the questionnaires.

The participants were given the Holt and Laury task and they were instructed to answer using whatever external representation was more appropriate to them. All of the participants had 20 minutes to complete the Holt and Laury task.

Results The independent variables were the education, numeracy, cognitive thinking style and external representational way. The dependent variable was the rational choice of the probabilities in the Holt and Laury task. In the context of this paper, is reflected as the participants' random choices in the task which indicate that participants either change lotteries in each row or choose only one Option between the two lotteries over the ten rows which has been supported as an inconsistent pattern by other studies (Jacobson & Petrie 2009; Dave et al. 2010). A coding system was used to categorize the external representations that participants used for the Holt and Laury task. The coding of the external representations was based on the coding adapted from previous research studied (Corter & Zahner, 2007; Zahner & Corter, 2010). The identified types were numbers, graphs, pictures, non-diagrammatic (text), and we added the blank page that it was not included in the coding method of Corter and Zahner (2007). Each representation was coded with one according to the above-mentioned types and added to a table. 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. To assess the reliability of the coding, two independent raters coded the responses. Cohen's kappa was run to determine if there was agreement between the two raters'. A Cohen's kappa of .957 and .977 represented almost absolute agreement between the two examiners for each of the five categories in the Holt and Laury task. To test the hypothesis whether participants who graduated from university would be more likely to answer rationally in the Holt and Laury task, a chi-square test of independence was performed. This test showed that participants who graduated from university were not more likely to answer rationally in the Holt and Laury compared to participants who had not graduated from university, X2 (1, N=60) = .429, p = .513.

To determine the relationship between specific variables and the rational decision making in Holt and Laury task, a chi-square test of independence was used. Therefore, to examine whether participants who scored lower in the validated numeracy scale, was associated to their rational choices in the Holt and Laury, a chi-square test of independence was performed. Numeracy performance was divided into two groups, one group with participants who were scored high (9-10-11 correct) and another group with those that scored less (2-8 items correct). Because the distribution of data was highly skewed, as mean numeracy was 8.1 out of 11 (a =.63), a median split was used for analysis, although it was taken under consideration that this split can cause loss of power (Peters et al. 2006; MacCallum, et al., 2002). 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 had a greater chance of choosing rationally in the Holt and Laury 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, to exclude it as a factor which influences the rational decision-making in this task. The results showed that there is no gender association with participants' rational choices, X^2 (1, N=60) = .019, p = .890. To test the hypothesis that participants who had graduated from university, would be more likely to use different external representations compared to the participants who did not attend university, 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. To test the hypothesis that there was an association with correct answers in the CRT and the rational choices in Holt and Laury, Fisher's Exact Test was performed. It revealed that participants who answered correctly more questions in the cognitive reflection task, they were more likely to answer rationally in the Holt and Laury, $p<.05$. The relation between numeracy and CRT was not examined as it was out of the context of the study.

A logistic regression was performed to investigate if using any specific way of external representations would be more likely to predict a rational answer. The logistic regression model was statistically significant at $p <.02$ according to the model chi-square statistic, meaning that the use of external representations can predict the rational choices in Holt and Laury. According to the logistic regression, graphs were shown to predict the rational choices in Holt and Laury and blank page, which included the answers where there was no verbal, mathematical or visual decision making process to reason the choices in the task and resulted in irrational choices (Table 1). The software used for the statistical analysis was SPSS.

Table 1. Statistical significance of the independent variables in the logistic regression whether any of the external representation would be more likely to predict a rational answer.

	B	S.E.	Wald	df	Sig	Exp(B)
Numbers	-1.438	1.478	.947	1	.330	.237
Graphs	3.095	1.431	4.675	1	.031	22.085
Pictures	.888	1.469	.365	1	.546	2.430
Words	.155	1.205	.017	1	.898	1.168
Blankpage	3.280	1.627	4.063	1	.044	26.577
Constant	-1.695	1.362	1.548	1	.213	.184

Experiment 2

Method

Participants In total 225 undergraduate computer science students, from a UK University, completed the tasks and the questionnaires. Participants in this study included 66 females and 159 males (M age = 29, age range 18 – 32). The allocation of the participant in three groups that were divided based on the respective format of Holt and Laury was randomized. Participants entered a lottery to win a £50 Amazon voucher as incentive to take part in the experiment, which was communicated during the introduction session.

Materials The materials used for this study include the Holt and Laury standardized version, the digital Holt and Laury displayed with pie charts and the digital Holt and Laury asking participants to fill in the pie charts before choosing the option.

In the Holt and Laury task displayed with pie charts (Figure 1), every option between the two lotteries is represented with a pie chart, and there is a text describing the proportions about the payoffs on the pie chart. Next, to the pie chart, the relevant payoff was displayed textually. The task was deployed in Unity Game Engine and logged all user actions and choices.

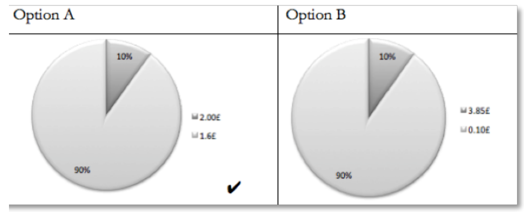

Figure 1. Holt and Laury displayed with non-interactive pie chart.

The stimuli developed for this study (Figure 2), involved 10 "empty" pie charts divided in 10 pieces with the proportions of each probability option of the lottery presented textually above them. Participants had to click on the region of the pieces on the pie chart to "fill them in" with specific colours according to the probabilities outlined in text and mark the pay offs of those pieces. For example, fill in with red one out of ten pies and fill in with blue nine out of ten. This way the participant can see clearly which option is more likely to happen. After the participants filled in the pie charts, they would choose one option. The stimuli was developed in a way that it did not allow the participant to choose the option before "filling in" the pie chart. Visually, the risk-taking task was very simple reflecting the standardised pen and paper HL task to avoid the effect of visual elements on participants' decision making process.

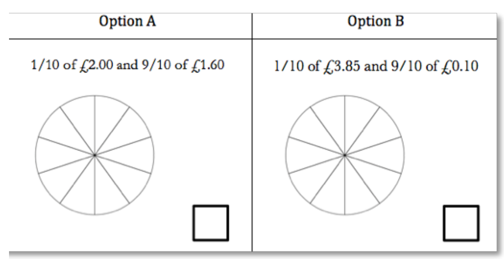

Figure 2. The empty circles where the participants had to "fill in" the parts for each pie chart before choosing the option.

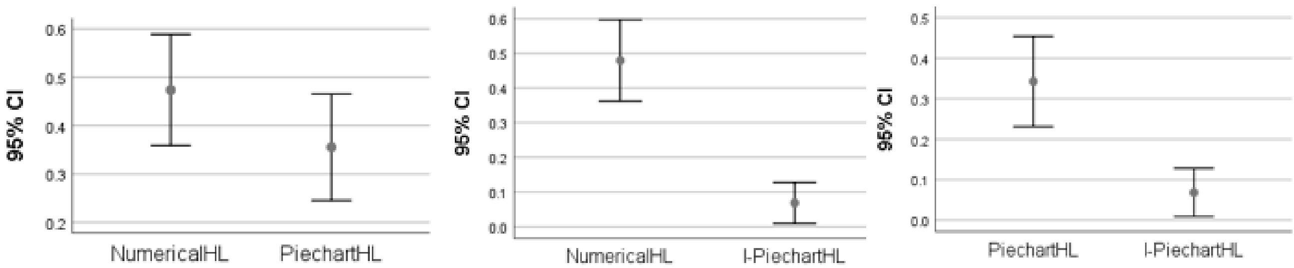

Figure 3. Participants' choices over the three different display formats with 95% Confidence Intervals. Dependent variable was the inconsistency rate in each display format of the Holt and Laury task. The lower mean shows a more consistent rate.

The numeracy test used is the Lipkus Numeracy scale (Lipkus, Samsa, & Rimer, 2001). 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. After they had completed the Holt and Laury tasks. They were asked to answer a question about task's level of difficulty. The question about difficulty level 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 and five was very difficult.

Procedure The experiment took place at the Faculty of Computer and Engineering at Coventry University 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. After the consent forms were obtained, in the second stage, the participants were given the numerical test and the demographic questions. When these were also collected, in the final stage, the participants were directed to five different rooms. Four rooms had computers where the interactive and non-interactive stimuli were set up. In the fifth room, the standardized Holt and Laury task was set up. The allocation of the participants to the rooms was random. The four rooms with the computer could fit 40 people each and the fifth room was a lecture theatre for 200 students' capacity. Participants were instructed that they would have 20 min to fill in the tasks. The participants were randomly assigned to go into one of the five rooms. Finally, 76 participants filled in the numerical display, 76 completed the pie chart display and 73 the interactive pie chart Holt and Laury task. Hence, Group 1 received the textual Holt and Laury task, Group 2 the Holt and Laury digital task displayed with pie charts and Group 3 received the digital interactive Holt and Laury method. The results are presented in the following section. After the completion of the task, 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.

Results According to participants' irrational choices, 35 out of 76 (46%) participants showed an inconsistent behavior in Holt and Laury task numerical format, 27 out of 76 (35.52%) participants in the pie chart format and only 4 out of 73 (5.47%) participants in the interactive pie chart Holt and Laury task format (Figure 4). 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, $p < .00$ (Figure 3). McNemar's test for related samples was also applied between interactive pie chart and pie chart Holt and Laury format w hich showed a significant difference between their inconsistency rates, $p < .00$ (Figure 3). 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 (Figure 3).

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 suggesting that numerical level (Wald statistic equal to 12.2), difficulty perception (Wald statistic equal to 8.8) and the age (Wald statistic equal to 4.7), were shown to be significant at the .00 level. Hence, they can predict participants' choices in the Holt and Laury numerical task. For the pie chart Holt and Laury, the binary logistic regression showed that the model is not statistical significant, $p < .178$. 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 5.42, which is significant at the .02 levels. The rest dependent variables of age, gender and numerical level were not statistically strong predictors of participants' choices in the interactive Holt and Laury format.

Participants took from 4 to 17 minutes to complete the Holt and Laury task using each representation type. The mean time taken to answer the task was 8.7 minutes with the interactive pie chart, 8.3 using the passive pie chart and 8.5 for the numerical. Even though the mean differences show that using the passive pie chart took them slightly less to fill in compared to the interactive and numerical, there is no significant difference between the times spent in each format.

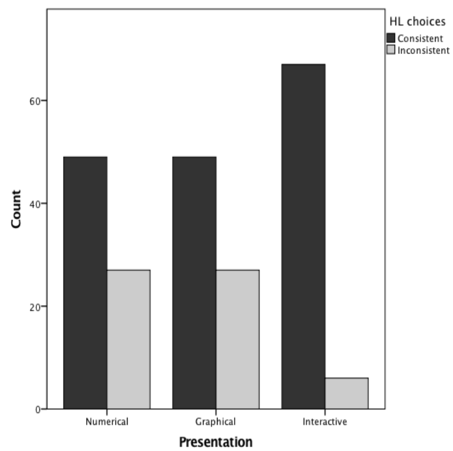

Figure 4. Participants' choices in relation to the presentation method of Holt and Laury.

Discussion

This paper presents the results of two experiments that investigated the methods that would assist in choosing rationally in a risk elicitation task. The findings from the 1st experiment validated empirical evidence that the use of external representations helps participants to understand the lotteries in the risk elicitation task. It also showed that using graphs has greater likelihood of choosing rationally in a risk elicitation task, compared to using images or any other way of external representation, as it has been supported by Hegarty and Kozhnenikov (1999). The study also showed that participants' years of education 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. Even though, it was found that there were no disparities between Groups 1 and Groups 2 with substantially dissimilar educational levels, this may not be true for educational levels defined in qualitatively different ways from other participants, or contexts.

Pie charts were shown to be better suited as they are better known from the general public and easily comprehensible to compare the size of two proportions when they are accompanied by labels (Nelson, Hesse, & Croyle, 2009). A step further though was demarcated 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. This outcome confirms the findings suggesting using graphs as external representation to facilitate successful probability problem solving (Hegarty, & Kozhevnikov, 1999) and thus rational choices in the Holt and Laury task. In the 2nd experiment, it was shown that participants who scored low in the numeracy scale chose more rationally in the Holt and Laury task when interacted with pie charts than any other format, confirming the need for a visual reasoning process to assist problem solvers (Carlsson, Johansson-Stenman, & Martinsson, 2004; Brase 2009). The interactive pie chart format was filled in rationally from those that scored high and low on the numeracy scale, as it was indicated by 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 population with specific characteristics to predict risk-taking behaviour in similar investment choices, 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 studies the validity of the metric needs to be examined and reassured. As this approach was only examined with UK University students, there is a limitation on generalizing to other cultures and audiences. Deploying a qualitative approach in conjunction to the quantitative methodology would enable to investigate more in depth on the underlying factors of why interacting with pie charts help people to understand the lotteries better. Finally, even though the average time spent in each display format of the task did not show any significant differences, future studies, need to explore whether using graphs to reason the choices in the Holt and Laury task, force the individuals with impulsive cognitive style to reflect more on their choices and aid their decision making.

This interactivity with graphs 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, automatically simplifies the task for less numerate people and allows for employing digital mediums, such as mobiles, for experimentation outside of laboratory settings. The data supports the hypothesis that the use of interactive pie charts, is more likely to result in consistent choices. This outcome may extend to other interactive artefacts such as games, simulations or analytics software, for crowdsourcing data for cognitive science of specific groups' (e.g. farmers) outside of the laboratory.

References

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

Bauermeister, G. & Mußhoff, O. (eds.), (2016), "Risk Attitude and Inconsistencies-does the Choice of Display Format and Risk Elicitation Method Influence the Outcomes? "Paper presented at *the 56th Annualbrod Conference, Bonn, Germany, September 28-30, 2016.German Association of Agricultural Economists*

Bougherara, D., Gassmann, X. and Piet, L. (2011), “A Structural Estimation of French Farmers Risk Preferences: An Artefactual Field Experiment”.
Brase, G.L. (2009). Pictorial representations in statistical reasoning. *Applied Cognitive* Psychology,23, 369-381
Carlsson, F., Johansson-Stenman, O. and Martinsson, P. (2004), “Is transport safety more valuable in the air?, *Journal of Risk and Uncertainty,* 28, pp. 147–163.
Corter, J. E., & Zahner, D. C. (2007). Use of external visual representations in probability problem solving. *Statistics Education Research Journal*, *6*(1), 22-50.
Dave, C., Eckel, C. C., Johnson, C. A., & Rojas, C. (2010), “Eliciting Risk Preferences: When is Simple Better?”, *Journal of Risk and Uncertainty,* 41(3), pp. 219-243.
Deitrick, S., Edsall, R., 2006. “The influence of uncertainty visualization on decision making: An empirical evaluation”, in: Riedl, A., Kainz, W., Elmes, G.A. (Eds.), Progress in spatial data handling. Springer, Berlin, pp. 719–738.
Frederick, S. (2005), “Cognitive Reflection and Decision Making”, *The Journal of Economic Perspectives, Vo 19* No 4, pp. 25-42.
Gilbert, J.K. (2008), “Visualization: An emergent field of practice and enquiry in science education.” *In Visualization: Theory and practice in science education, pp. 3-24*. Springer, Dordrecht.
Gray, M., & Holyoak, K. (2018),”Individual Differences in Relational Reasoning*”, Proceedings of the 40th Conference of the Cognitive Science Society*
Habib, S., Friedman, D., Crockett, S., & James, D. (2016), “List Construction and Lottery Presentation Modulate Multiple Price List Responses”.
Hegarty, M. and Kozhevnikov, M. (1999), “Types of visual–spatial Representations and Mathematical Problem Solving”, *Journal of Educational Psychology, 91*(4), pp. 684.
Holt, C. A., & Laury, S. K. (2002), “Risk Aversion and Incentive Effects”, *American Economic Review,* 92(5), pp. 1644-1655.
Jacobson, S., & Petrie, R. (2009). Learning from mistakes: What do inconsistent choices over risk tell us?. *Journal of risk and uncertainty*, *38*(2), 143-158.
Khan, A., Breslav, S., & Hornbæk, K. (2018). Interactive instruction in bayesian inference. *Human–Computer Interaction*, *33*(3), 207-233.
Lipkus, I. M., Samsa, G., & Rimer, B. K. (2001), “General Performance on a Numeracy Scale among Highly Educated Samples”, *Medical Decision Making,* 21(1), pp.37.
MacCallum, R.C., Zhang, S., Preacher, K.J., & Rucker, D.D. (2002). On the practice of dichotomization of quantitative variables. Psychological Methods, 7, pp. 19–40.
Nelson, D.E., Hesse, B.W. and Croyle, R.T., 2009. “*Making data talk: Communicating public health data to the public, policy makers, and the press”.* New York, NY: Oxford University Press.
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), pp. 2.
Padilla, L., Creem-Regehr, S. H., Hegarty, M., & Stefanucci, J. K. (2018). Decision making with visualizations: a cognitive framework across disciplines. *Cognitive Research: Principles and Implications,* 3(1), 29.
Peters, E. (2008), “Numeracy and the Perception and Communication of Risk” *Annals of the New York Academy of Sciences,* 1128(2), pp. 1-7.
Peters, E., Västfjäll, D., Slovic, P., Mertz, C. K., Mazzocco, K., & Dickert, S. (2006). Numeracy and decision making. *Psychological science*, *17*(5), 407-413.
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), pp. 851-865.
Rohrmann, B. (2005). Risk attitude scales:concepts, question utilizations. *Project Report, 1-21.*
Starns, J. J., Cohen, A. L., Bosco, C., & Hirst, J. A (2018), A visualization technique for Bayesian reasoning*. Applied Cognitive Psychology,* 33(2), pp. 234-251.
Zahner, D., & Corter, J. E. (2010). The process of probability problem solving: Use of external visual representations. *Mathematical Thinking and Learning*, *12*(2), pp. 177-204.
Ziemkiewicz, C., Ottley, A., Crouser, R.J., Chauncey, K., Su, S.L., & Chang, R. (2012), “Understanding visualization by understanding individual users”, *IEEE computer graphics and applications, 32*(6), pp.88-94.
Zhang, J. (1997). The nature of external representations in problem solving. Cognitive science, 21(2), pp. 179-217.