RISK ATTITUDE AND INCONSISTENCIES - DOES THE CHOICE OF DISPLAY FORMAT AND RISK ELICITATION METHOD INFLUENCE THE OUTCOMES?

RISIKOEINSTELLUNG UND INKONSISTENZEN - BEEINFLUSST DIE WAHL DER DARSTELLUNGSFORM UND DER MESSMETHODE DIE ERGEBNISSE?

Golo-Friedrich Bauermeister

Department of Agricultural Economics and Rural Development,
Georg-August-Universität Göttingen

Oliver Mußhoff

Department of Agricultural Economics and Rural Development,
Georg-August-Universität Göttingen

Corresponding author: golo-friedrich.bauermeister@agr.uni-goettingen.de

2016

Paper prepared for presentation at the 56th annual conference of the GEWISOLA (German Association of Agricultural Economists)

„Agricultural and Food Economy: Regionally Connected and Globally Successful"
Bonn, Germany, September 28 – 30, 2016

RISK ATTITUDE AND INCONSISTENCIES - DOES THE CHOICE OF DISPLAY FORMAT AND RISK ELICITATION METHOD INFLUENCE THE OUTCOMES?

RISIKOEINSTELLUNG UND INKONSISTENZEN – BEEINFLUSST DIE WAHL DER DARSTELLUNGSFORM UND DER MESSMETHODE DIE ERGEBNISSE?

Abstract

In the past decade, many studies have measured individual risk attitude with different elicitation methods using within-subject design and have found significant disparities across the elicitation methods. According to the existing literature, there are also differences in the observed understanding of the elicitation methods measured by the inconsistency rate. However, there are no studies yet that compares the inconsistency rate across different elicitation methods in a within-subject design. Therefore, we compare both the inconsistency rate and the risk attitude of participants in two different lottery tasks in a within-subject design, namely the lottery task by HOLT and LAURY (2002) as well as the one by BRICK, VISSER and BURNS (2012). Moreover, we analyze in a between-subject design whether a visualization of a lottery task for a better understanding results in differences in the elicited risk attitude and thus leads to the desired reduction of the inconsistency rate. Results show that the elicited risk attitudes are more risk averse in the more complex Holt-and-Laury task in both display formats. Moreover, we find that the visualization results in more risk averse responses in both lottery tasks. According to the inconsistency rate, we find that the Brick-Visser-Burns task is better understood than the Holt-and-Laury task, especially in the textual display format. Furthermore, the visual display format of the Holt-and-Laury task results in a significantly better understanding compared to the textual display format.

Keywords

Brick-Visser-Burns task, display formats, Holt-and-Laury task, inconsistency rate, risk attitude

1 Introduction

Risk and uncertainty play an important role in a variety of financial and economic decisions (ABDELLAOUI, DRIOUCHI and L'HARIDON, 2011). Therefore, knowing the individual risk attitude is of central importance to predict and understand financial and economic behavior. For instance, in developing countries where the majority of people depend on agriculture, the individual risk attitude influences decisions in technology adoption (PURVIS et al., 1995), crop selection (PRICE and WETZSTEIN, 1999) or crop insurance markets (HILL and VICEISZA, 2012). Hence, it is important to reveal the individual risk attitude to predict or to explain people's economic behavior.

In the past, numerous research studies focused on measuring the individual risk attitude. Some studies included econometric approaches where the risk attitude was estimated using empirically observed data whereas others were based on experimental methodology, which has been more and more generally accepted. By now, there are many methods in the existing literature that experimentally determine the individual risk attitude. BINSWANGER (1980) was among the first who did not only conduct hypothetical surveys but also combined them with real monetary payouts. HOLT and LAURY (2002), TANAKA, CAMERER and NGUYEN (2010) and BRICK, VISSER and BURNS (2012) used a multiple price list (MPL) where the experimental participants were asked to decide in various rows between a safer and a riskier option. During the last decade, the Holt-and-Laury task (HL) using this MPL has become the "gold standard" in risk elicitation (ANDERSON and MELLOR, 2008). According to ANDERSEN et al. (2006), positive advantages of the MPL include both the fact that participants can notice relatively easily that a true answer is for their own advantage, and the fact that the MPL is easy to implement. Negative aspects, however, entail the fact that it is only possible to determine intervals of the individual risk attitude, and that the framework may influence the outcome. Further methods to determine the risk attitude are, for instance, the technique by ECKEL and GROSSMAN (2002; 2008) (EG), the self-estimation technique by DOHMEN et al. (2011) or the contextual statements by WEBER, BLAIS and BETZ (2002).

The problem of different elicitation methods is that they often lead to different results when determining risk attitude (ISAAC and JAMES, 2000; DAVE et al., 2010). Following ANDERSEN et al. (2006), different frameworks and risk aversion coefficients used in the elicitation methods can be an explanation for these comparison biases. Moreover, elicitation methods can lead to a wrong individual risk attitude through the participant's lack of understanding (GALARZA, 2009). In this case, the HL can be applied to determine an inconsistency rate, which provides information about how many participants did not correctly understand the survey methodology or gave inconsistent answers due to boredom or hurry. The values range from a few percentages (DAVE et al., 2010; HOLT and LAURY, 2002) to inconsistency rates of over 50 percent (CHARNESS and VICEISZA, 2015; GALARZA, 2009). BRICK, VISSER and BURNS (2012) also developed a method to measure the individual risk attitude that allows inferences about the inconsistency rate. The Brick-Visser-Burns task (BVB) is based on the idea by DOHMEN et al. (2010) of varying payoffs and fixing probabilities ("probability equivalence method") instead of the HL-idea, which varies probabilities and fixes payoffs ("certainty equivalence method"). In doing so, BRICK, VISSER and BURNS (2012) seek to facilitate the method and thus to reduce the inconsistency rate. Moreover, elicitation methods are illustrated in visual display formats in some studies (BOUGHERARA, GASSMANN and PIET, 2011; GALARZA, 2009) to improve the understanding and hence, reducing the inconsistency rate. A lower inconsistency rate results in more reliable data and therefore to more power of explaining economic behavior.

Consequently, beyond the existing literature in this field of research, the present study deals with three new aspects. First, to the best of our knowledge, this is the first study that compares the risk attitude determined by the HL and a BVB where the framework and the risk aversion coefficients are adjusted to the HL. Second, this is the first study dealing with the impact of different display formats and analyzing the determined risk attitude in due consideration of this aspect. It has not been investigated in the existing literature whether the allegedly simplification by a visual display format has an impact on the risk attitude determined by the HL and BVB. Third, this study analyzes the impact of different display formats on the inconsistency rate. It is true that there are previous experiments that used a visual display format to help participants to gain a better understanding of the experiment. However, the allegedly more complex textual format has not been directly compared to the allegedly simpler visual format in terms of the inconsistency rates.

The present study is structured as follows: the introduction in section 1 is followed by the generation of hypothesis in section 2. Subsequently, the experimental design is described in section 3. The results of the analysis are presented in section 4, and the study ends with conclusions and future prospects in section 5.

2 Relevant literature and generation of hypotheses

The existing literature contains many methods to determine the individual risk attitude. Some methods are compared with regard to the differences in their results. For example, DAVE et al. (2010) use Canadian students to analyze the differences between the common HL and the elicitation method of EG, which both have an incentive-compatible design. They come to the conclusion that the determined risk attitudes of the participants in the HL are significantly higher than those of the EG method. With German participants, MAART-NOELCK and

MUSSHOFF (2013) find a significant positive correlation between an incentive-compatible HL and a self-assessment as well as a business-related context-based statement. The problem of comparing different measuring methods lies in the different frameworks and risk aversion coefficients on which these methods depend, or when these features cannot be determined as it is the case for self-assessments or context-related statements. REYNAUD and COUTURE (2012) examine this problem with French farmers and compare the HL with the EG method. The latter is adjusted to the HL regarding the constant relative risk aversion (CRRA) and the framework. However, the EG method does not allow for a determination of inconsistent participants and the experiment is also not incentivized. The results reveal significantly lower CRRAs for the participants of the HL than of the EG method. As the HL and the BVB do not per se exclude inconsistent behavior, and since the framework of the BVB can be adjusted to the HL, we compare the "gold standard" method of HOLT and LAURY (2002) with the allegedly simpler method of BRICK, VISSER and BURNS (2012). The corresponding hypothesis is:

H1 "*Risk attitudes - HL vs. BVB*": There are not any significant differences between the individual risk aversion coefficients determined by a HL and a BVB.

In order to reduce complexity and improve the understanding, some measurement techniques are displayed in the form of graphs. For example, BOUGHERARA, GASSMANN and PIET (2011) use a wheel of fortune in order to illustrate probabilities and make the HL easier to understand. Also GALARZA (2009) visualizes the HL by using graphs. For each row, the author presents two bars that stand for the two different options. Depending on the probabilities, the corresponding proportions of the bars with the associated possible prizes are shaded. IHLI, CHIPUTWA and MUSSHOFF (2013) replace the textual forms in the HL and the BVB by pictures of bags filled with colored balls that stand for different prizes. Whether the visualization of the HL and the BVB lead to a significant difference in the determined risk attitude compared to the corresponding textual forms has not been analyzed yet. Thus, the following hypothesis was derived:

H2 "*Risk attitudes - textual form design vs. visual form design*": There are not any significant differences between the individual risk aversion coefficients determined by differently displayed HLs (BVBs).

Elicitation methods structured in the MPL-format, such as the HL and the BVB, bear the advantage that it is possible to determine the inconsistency rate. Participants who switch more than once between the secure and the risky option in the course of the lottery task show an inconsistent behavior (GALARZA, 2009). When applying the HL with a sample drawn from the rural population in Senegal, CHARNESS and VICEISZA (2015) find an inconsistent response behavior of 51 percent. Similarly, GALARZA (2009) detect an inconsistency rate of 52 percent for Peruvian cotton farmers. JACOBSON and PETRIE (2009) determine an inconsistency rate of 55 percent for a sample of the Rwandan population by varying the payout amounts. BRICK, VISSER and BURNS (2012), who work with fixed probabilities of 50 and 100 percent, find a multiple switching behavior of 41 percent for fishers from the west coast of South Africa. In a study in Mozambique, DE BRAUW and EOZENOU (2011) determine an inconsistent response behavior for 14 percent of the participants. This value is relatively low for developing countries. A possible explanation may be the verbal instructions that the participants received. In the industrialized countries, DAVE et al. (2010) determine an inconsistency rate of 8.5 percent for a sample of 881 adult Canadians. In a "low-payoff" experiment, HOLT and LAURY (2002) find an inconsistency rate of 13.2 percent for students in the USA. Using a multiplication of payout amounts with the factor 50 or 90, the inconsistent response behavior is reduced to 5.5 percent. Hence, a reason for a high inconsistency rate can be a low monetary incentive (LÉVY-GARBOUA et al., 2012). In contrast, REYNAUD and COUTURE (2012) and DAVE et al. (2010) foreground the very high complexity of the widely used HL. BOSCH-DOMÈNECH and SILVES-

TRE (2013) show that there are more inconsistent participants in the HL group than in the group where participants are confronted with fixed probabilities. For this reason, the BVB with constant probabilities and varying payout amounts seems to be easier to understand. So far, it has not been investigated if there is a significant difference in the level of the inconsistency rate of the widely used HL and the BVB. Therefore, the next hypothesis was formulated as follows:

H3 "*Inconsistency rates - HL vs. BVB*": There is not any significant difference between the inconsistency rate of the HL in textual form (visualized form) and of the BVB in textual form (visualized form).

Inconsistent participants are often excluded from analysis because of the the assumption that the participants did not completely understand the experiment and thus reflect a wrong individual risk attitude (GALARZA, 2009; CHARNESS, GNEEZY and IMAS, 2013). If a majority of the participants provides inconsistent responses, the validity of the determined risk attitude of the sample is vastly diminished. Hence, it is important to adapt the elicitation methods to the education of the participants and to present them as clearly as possible. Furthermore, it is useful to choose elicitation methods that allow detecting inconsistent participants. If this is not possible, there is no possibility to eliminate analyzed data, which will include the risk attitude of people who did not completely understand the task. Therefore, the elicitation methods should not per se exclude inconsistent behavior, and the presentation and complexity should be as easy to understand as possible. Large posters, examples and reading out the experimental instructions enhance the understanding (CARDENAS and CARPENTER, 2008). In an experiment with Canadian students, BRUNER (2011) find that a group who is first provided with verbal and subsequently with written instructions shows a more consistent response behavior than a group who is only provided with written instructions. The results of BRUNER (2011) reveal that verbal instructions diminish inconsistent behavior significantly. However, this assistance is time-consuming and not compatible with an online survey. Graphs or visual support as used by BOUGHERARA, GASSMANN and PIET (2011) or GALARZA (2009) are an alternative to improve the comprehension. To date, it has not been investigated if there are differences in the inconsistency rate of an elicitation method that is based on textual form and one that is based on visualization. Thus, the corresponding hypothesis is:

H4 "*Inconsistency rates - textual form design vs. visual form design*": There is not any significant difference in the inconsistency rate of different display formats of the HL (BVB).

3 Methodology

In the experiment, students face two different elicitation methods to measure their risk attitude; the lottery task by HOLT and LAURY (2002) as well as the one by BRICK, VISSER and BURNS (2012). In doing so, one group takes the lottery tasks in textual form, while the other group is provided with the visualized form of the lottery tasks. As monetary incentives promote the participants' extrinsic motivation to make a certain effort (BONNER and SPRINKLE, 2002), both lottery tasks have an incentive-compatible design.

3.1 Holt-and-Laury task

For the HL, participants are asked to decide one out of two options in ten different rows. All rows are presented to the participants at once, whereas the order of the rows does not vary. In the experiment, the possible payout amounts are set to € 60 and € 48 in option A and to € 115.50 and € 3 in option B (cf. table 1). In the visual display format, options are visualized as bags with ten colored balls in each (blue, red, green and yellow). The different colors represent the various possible payout amounts (€ 48, € 60, € 3 and € 115.50). Due to the lower difference between the possible payouts, option A is the more secure choice and option B is the riskier choice. The probabilities for higher payout amounts are 10 percent and consequently 90 percent for the lower amounts in the first row. From one to the next row, the probabilities change by 10 percent, so that in the tenth row, a probability of 100 percent occurs for the higher amount. Hence, the probabilities for the HL are varying and the possible payout amounts remain constant.

Table 1: Payout matrix of the HL

row	option A	option B	EV (A)[a]	EV (B)[a]	CRRA-span[a b]
1	with a 10% gain of € 60.00 with a 90% gain of € 48.00	with a 10% gain of € 115.50 with a 90% gain of € 3.00	€ 49.20	€ 14.25	r < - 1.71
2	with a 20% gain of € 60.00 with a 80% gain of € 48.00	with a 20% gain of € 115.50 with a 80% gain of € 3.00	€ 50.40	€ 25.50	- 1.71 < r < - 0.95
…	…	… …	…	…	…
9	with a 90% gain of € 60.00 with a 10% gain of € 48.00	with a 90% gain of € 115.50 with a 10% gain of € 3.00	€ 58.80	€ 104.25	0.97 < r < 1.37
10	with a 100% gain of € 60.00 with a 0% gain of € 48.00	with a 100% gain of € 115.50 with a 0% gain of € 3.00	€ 60.00	€ 115.50	1.37 < r

[a] Columns were not shown to the participants.
[b] A power utility function of the form $U(x) = x^{(1-r)}/(1-r)$ is assumed.

The expectation values of both options change from one row to another. Up to and including the fourth row, the expectation value for the more secure option A is higher than that for the more risky option B. As from the fifth row, the more risky option B has a higher expectation value. Consistent participants change at a certain time from the more secure to the more risky option depending on their individual risk attitude. By the time of this change, it is possible to determine the participants' individual risk attitude. A risk-seeking participant would change from option A to option B within the first four rows. The HL-value representing the number of safe choices is thus between 0 and 3 (HL-value < 4). A risk-neutral participant who is only interested in a higher expectation value would change to option B after the fourth row (HL-value = 4). Participants who decide after the fifth row for the riskier option B (HL-value > 4) are categorized as risk averse. Assuming a power utility function with a constant relative risk aversion: $U(x) = x^{(1-r)}/(1-r)$, where x represents the lottery payout, a span of the individual risk aversion coefficient (r) can be calculated for each row. A positive CRRA-value indicates risk averse behavior, while a negative value reveals risk-seeking behavior. In the last row, the participants receive € 60 for option A and € 115.50 for option B both with a probability of 100 percent. In this case, there is an absolute stochastic dominance of option B over option A. Participants who still choose the dominated option A in the last row – and thus have a HL-value of 10 – cannot act in the sense of a profit-maximizer and are not assigned a CRRA-value.

3.2 Brick-Visser-Burns task

The design and procedure of the lottery experiment by BRICK, VISSER and BURNS (2012) are similar to the HL. Both presentation forms (textual and visual) include all rows at once, whereas the order of the rows does not vary. The participants are asked to choose from two options in ten rows. In contrast to the HL, the possible payout amounts vary and the probabilities remain constant for the BVB (cf. table 2). For simplicity reasons, the probability is 100 percent in option A and 50 percent each in option B. In contrast to option B, there is no difference among the possible payout amounts in option A, and A is thus the secure alternative. In the visual form in option B, the red and blue balls with the values of € 115.50 and € 3 remain constant in their amount and value for each row. For option A, there are exclusively yellow balls in the bags, which change in their values. Unlike BRICK, VISSER and BURNS

(2012), we extend the number of rows from eight to ten and adjust the CRRA-values to the values of HOLT and LAURY (2002)[1].

Table 2: Payout matrix of the BVB

row	option A	option B	EV (A)[a]	EV (B)[a]	CRRA-span[a b]
1	with a 100% gain of € 115.50	with a 50% gain of € 115.50 with a 50% gain of € 3.00	€ 115.50	€ 59.25	-
2	with a 100% gain of € 89.50	with a 50% gain of € 115.50 with a 50% gain of € 3.00	€ 89.50	€ 59.25	r < -1.72
…	…	… …	…	…	…
9	with a 100% gain of € 19.50	with a 50% gain of € 115.50 with a 50% gain of € 3.00	€ 19.50	€ 59.25	0.68 < r < 1
10	with a 100% gain of € 10.50	with a 50% gain of € 115.50 with a 50% gain of € 3.00	€ 10.50	€ 59.25	1 < r < 1.37

[a] Columns were not shown to the participants.
[b] A power utility function of the form $U(x) = x^{(1-r)}/(1-r)$ is assumed.

For the option B, participants can win € 115.50 or € 3 with a probability of 50 percent in each row (cf. table 2). Therefore, the expectation value is constant at € 59.25. However, the expectation value of option A is reduced from € 115.50 in the first row to € 10.50 in the tenth row. In the first five rows, option A has a higher expectation value. From the sixth row onwards, this changes in favor of option B. Consequently, participants who have a BVB-value ("number of safe choices") of five, are risk neutral. A change in the first five rows (BVB-value < 5) indicates risk-seeking participants, while the change after the sixth row (BVB-value > 5) represents risk averse individuals. Assuming a risk utility function, it is possible to calculate a span of the individual risk aversion coefficient (*r*) for each row as it the case for the HL. In the first row, participants receive € 115.50 with a probability of 100 percent by choosing option A. In option B, they have a probability of 50 percent to win either € 115.50 or € 3. As the worst result of option A is at least as high as the best payout amount of option B, there is an absolute stochastic dominance of option A over option B in the first row. However, a person who decides for the dominant option B in the first row cannot act as a profit-maximizer and is not assigned a CRRA-value.

3.3 Data collection

The survey for the determination of the risk attitudes and the inconsistency rates was conducted online in 2014. Students from a faculty of agricultural sciences who will become part of the agricultural economics in the future were invited to participate in the experiment via e-mail lists. Previous studies show that there are matches in the behavior of real decision-makers and students regarding the information processing and the associated decision-making processes (ASHTON and KRAMER, 1980). The experiment consisted of three parts and took approximately 11 minutes. Each participant carried out once the HL as well as the BVB. The order of the elicitation methods was randomized. Therefore, some participants started with the HL in the first part of the experiment and continued with the BVB in the second part, while others carried out the BVB first and the HL in the second part. With regard to the display format, there existed two groups. One group was faced with both lottery tasks in textual form and the other group was provided with a visualized form of the lottery tasks. It was not possible that one participant had to deal with different display formats. The allocation of the participants in the respective group was randomized. In the third part of the experiment, socio-demographic and socio-economic data were collected. The first and second part of the experiment started with written instructions tailored to the respective method. Next to an exemplarily illustration of the respective elicitation method (textual or visualized form), an explana-

[1] We rounded to € 0.50 to reduce the complexity of the BVB; therefore, there are very small differences in the CRRA-spans.

tion of the determination of the cash prize was given. It was ex-ante communicated that 10 percent of the participants would win a cash prize. For the later determined prize, the lottery task the participant had to face (HL or BVB) was randomly determined and subsequently a randomized row was conducted. The determined payout amount was transferred to the winners' bank accounts after drawing. Depending on the lottery task, the row and the selected option, it was possible to win between € 3 and € 115.50. In the experiment, data conducted online were only collected from participants who completed all three parts. A consideration of participants who did not finish all three parts would not allow an intrapersonal comparison. Thus, data of 307 students can be analyzed, where 149 participants are provided with the textual form and 158 participants are faced with the visualized form. The mean age averages 23.5 years with a standard deviation of 3.3 years. There are 135 female and 172 male students who participated in the experiment

4 Results

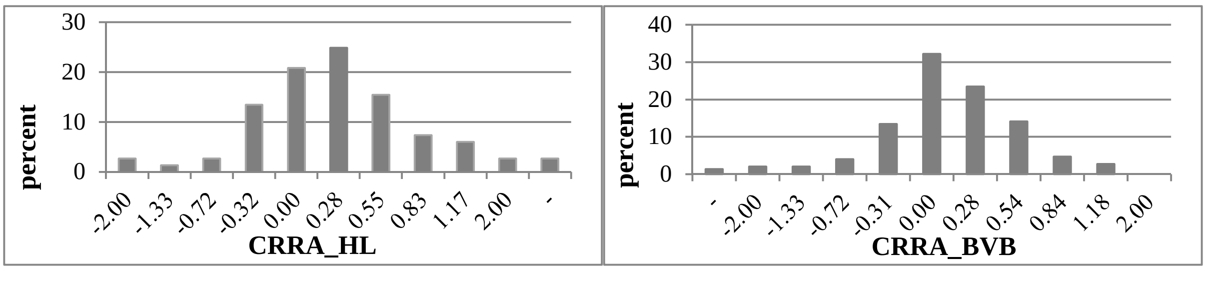
In the following, the hypotheses from section 2 are verified.

4.1 Verification of H1

"*Risk attitudes for textual form design - HL vs. BVB*"

The distributions of the CRRA-values from the 149 participants who were provided with the textual form are illustrated in figure 1. Similarly to REYNAUD and COUTURE (2012), we work with the means instead of CRRA-spans[2]. In the HL, 88 persons (59.1 percent) are classified as risk averse (HL-value > 4). The CRRA-value is positive for risk averse participants. Despite the absolute stochastic dominance of option B over option A in the last row of the HL, four participants always choose option A and therefore, no CRRA-value can be assigned. Moreover, 31 participants (20.8 percent) are classified as risk neutral (CRRA = 0) and 30 participants (20.1 percent) as risk seeking (CRRA < 0). In the BVB, 67 participants (45 percent) are classified as risk averse (CRRA > 0), 48 participants (32.2 percent) as risk neutral (CRRA = 0) and 34 participants (22.8 percent) are classified as risk seeking (CRRA < 0). Despite the absolute stochastic dominance of option A over option B in the first row, two participants always choose option B and thus, no CRRA-value can be assigned. To allow for an intrapersonal comparison of the means, the six participants with the missing CRRA-value are excluded from the analysis.

Figure 1: Distribution of CRRA-values for textual forms (n=149)



The CRRA-means average 0.20 in the textual form of the HL and 0.07 in the text format of the BVB, reflecting a small risk aversion in each case (cf. table 3). Furthermore, the participants are slightly more risk averse in the HL. To test the significance of the difference between the CRRA-means, CRRA-values need to be checked for normal distribution. According to the Kolmogorov–Smirnov test and the Shapiro–Wilk test, the CRRA-values of the HL and BVB do not follow a normal distribution with a probability of error of 5 percent. Conse-

[2] On the margins of the lottery tasks, where there are no CRRA-spans, we work with CRRA-values of -2 and 2 as done by REYNAUD and COUTURE (2012).

quently, we work with the non-parametric Wilcoxon signed-rank test to compare the CRRA-means. This intrapersonal test reveals that the CRRA-means of both lotteries are significantly different (p-value = 0.003). Thus, for the textual form design, the hypothesis H1 can be rejected.

Table 3: Comparison of the CRRA-Means [a)]

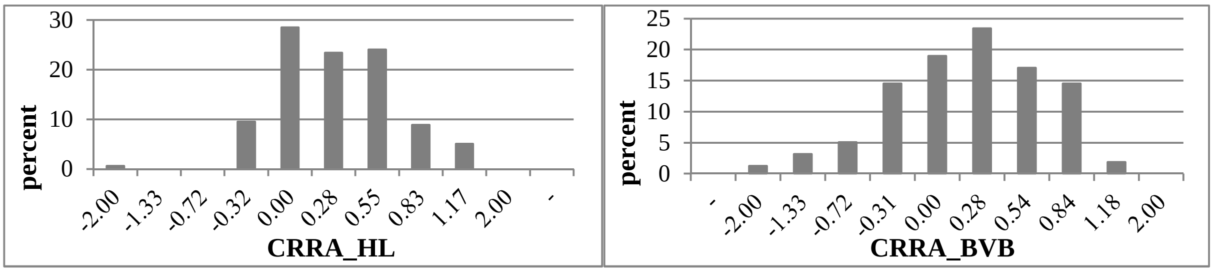
	CRRA-means HL	CRRA-means BVB	p-value Wilcoxon signed-rank-test
group face textual form	0.20 (0.66)	0.07 (0.53)	0.003
group face visual form	0.29 (0.42)	0.16 (0.57)	0.000
p-value Mann-Whitney U test	0.188	0.047	-

[a)] Standard deviation in brackets

"*Risk attitudes for visual form design - HL vs. BVB*"

The distributions of the CRRA-values from the 158 participants who face the visual form of the lottery tasks are illustrated in figure 2. In the HL, 97 participants (61.4 percent) are classified as risk averse (CRRA > 0), 45 participants (28.5 percent) as risk neutral (CRRA = 0) and 16 participants (10.1 percent) as risk seeking (CRRA < 0). In the BVB, 90 participants (57.0 percent) are classified as risk averse (CRRA > 0), 30 participants (19.0 percent) as risk neutral (CRRA = 0) and 38 participants (24.0 percent) as risk seeking (CRRA < 0). None of the participants who face the visual form chose a statistically dominated option, so a CRRA-value can be assigned to every participant.

Figure 2: Distribution of CRRA-values for visual forms (n=158)



The CRRA-means average 0.29 in the visualized HL and 0.16 in the visualized BVB reflecting a small risk aversion in each case (cf. table 3). Similarly to the group with the textual form, the replies of the participants were on average more risk averse in the HL than in the BVB. The Kolmogorov–Smirnov test and the Shapiro–Wilk test also show that the distribution of the CRRA-values is not normally distributed with a probability of error of 5 percent for the group faced with the visualization. Thus, we work with the non-parametric Wilcoxon signed-rank test to additionally analyze the CRRA-means. This intrapersonal test reveals that the CRRA-means of both lotteries are significantly different (p-value < 0.001) in the group provided with the visualization. Hence, hypothesis H1 can also be rejected for the visual form design.

4.2 Verification of H2

"*Risk attitudes - textual forms design vs. visual form design*"

The different CRRA-means and the respective standard deviations of both lottery tasks and display formats are shown in table 3. This table also provides the p-values calculated with a Mann–Whitney U test that may lead to a significant difference in the CRRA-means of one lottery task in different display formats. The means indicate that the group faced with the visual form is more risk averse in the HL and in the BVB than the group provided with the textual form. The p-value of the Mann–Whitney U test shows that the difference between the display formats is not significant regarding the CRRA-means of the HL (p-value = 0.188).

Accordingly, the display format used does not influence the participants' risk attitude measured by the HL. In contrast, there is a significant difference in the participants' risk attitude measured when applying different display formats of the BVB (p-value = 0.047). As a consequence, the hypothesis H2 can be rejected for the BVB, while it cannot be rejected for the HL.

4.3 Verification of H3

"*Inconsistency rates for textual form design - HL vs. BVB*"

To obtain a more detailed analysis of the inconsistency rate, participants are divided in four groups (cf. table 4). With regard to the HL, the first group is consistent in terms of their decision-making behavior and includes all participants who switched only once, from option A to option B. Participants of the second group who always chose option B are extremely risk seeking but still consistent in their decision-making behavior. Participants who always chose option A, and thus preferred the dominated option A in the last row, are classified as inconsistent in group three. The fourth group comprises of participants who switched at least twice between option A and option B or switched directly from option B to option A. These participants are also classified as inconsistent. With respect to the BVB, the first group also comprises of participants who first chose option A and switched to option B later at some point. The fourth group also includes participants who switched more than once or who switched directly from option B to option A. In contrast to the HL, the second group of the BVB consists of participants who are consistent by always choosing option A and, therefore, are extremely risk averse. Moreover, participants who always chose option B and thus chose the dominated option B in the first row are classified as inconsistent in group three.

Table 4: Classification of Consistent and Inconsistent Participants

		group	textual form (n=149)		visual form (n=158)	
			absolut	relative	absolut	relative
HL	consistent	1: One switch from option A to option B	110	73.83%	137	86.71%
		2: Always option B	4	2.68%	1	0.63%
	inconsistent	3: Always option A	4	2.68%	0	0.00%
		4: Multiple switch or directly from option B to option A	31	20.81%	20	12.66%
BVB	consistent	1: One switch from option A to option B	134	89.93%	139	87.97%
		2: Always option A	0	0.00%	0	0.00%
	inconsistent	3: Always option B	2	1.34%	0	0.00%
		4: Multiple switch or directly from option B to option A	13	8.72%	19	12.03%

According to this classification, 35 participants (23.49 percent) show an inconsistent behavior in the textual form design of the HL. In the textual form design of the BVB, however, only 15 participants (10.06 percent) can be classified as inconsistent. With regard to the inconsistency rates, the difference between the HL and BVB amounts to 13.42 percent. This large disparity indicates that the BVB may be easier to understand in the textual form design than the HL. To test whether this difference is significant, we apply a McNemar's test for related samples, which reveals a significant difference (p-value < 0.001) in the inconsistency rates between both lottery tasks (cf. table 5). Hence, the hypothesis H3 can be rejected for the textual form design.

Table 5: Comparison of the Inconsistency Rates

	inconsistency rate HL	inconsistency rate BVB	p-value McNemar's test
group face textual form	23.49%	10.06%	< 0.001
group face visual form	12.66%	12.03%	1.000
p-value Chi-squared test	0.013	0.585	-

"*Inconsistency rates for visual form design - HL vs. BVB*"

Next to the inconsistency rates of the textual form design, table 4 also shows the inconsistency rates of the visual form design. The classification is applied equally to the textual form design. Twenty participants (12.66 percent) indicate an inconsistent behavior in the HL, whereas none of the participants always chose option A (group 3). In the BVB, since none of the participants chose always option B, we classify only 19 participants (12.03 percent) of group 4 as inconsistent. With regard to the inconsistency rates, the difference between the HL and BVB amounts to 0.63 percent. Again, we apply the McNemar's test for related samples to test whether this small difference is significant or not. The p-value of 1.000 (cf. table 5) demonstrates that the inconsistency rates of both lotteries are not significantly different in the visual form design. Accordingly, hypothesis H3 cannot be rejected for the visual form design.

4.4 Verification of H4

"*Inconsistency rates - textual form design vs. visual form design*"

The results of H3 reveal that different kinds of lottery tasks can lead to divergent inconsistency rates. Table 5 indicates the inconsistency rates of the different form designs and lottery tasks. In regard to the inconsistency rate of the HL, there is a relatively large difference of 10.83 percent between the textual and visual form design. In comparison, in the BVB, a relatively small difference of 1.96 percent occurs between both form designs. To test whether these differences are significant, we apply the Chi-squared test. The analysis shows that the difference between both form designs in the HL is significant with a p-value of 0.013. Therefore, when looking at the inconsistency rates, the visual form design of the HL fared better than the textual form design. It seems that visualization improves the understanding of the participants for the HL. Regarding the BVB, the verification of both form designs does not reveal any significant difference (p-value = 0.585) for the inconsistency rates. In contrast to the HL, the visualization of the allegedly more comprehensible BVB does not show any significant impact of a better understanding of one of the two display formats tested. Hence, hypothesis H4 can be rejected for the HL but cannot be rejected for the BVB.

5 Conclusion

The individual risk attitude affects the choices of decision-makers. Hence, the knowledge of the individual risk attitude is of crucial importance to predict or explain economic behavior. However, it is quite problematic that the determined individual risk attitudes differ across different elicitation methods and inconsistent answers distort the true risk attitude. Therefore, we first applied the allegedly simpler lottery task by BRICK, VISSER and BURNS (2012) next to the "gold standard" method by HOLT and LAURY (2002). In a second step, we worked with two different display formats (textual and visual) that might influence the comprehension of the participants. We carried out a comparison in a within-subject design to examine whether there is a significant difference in the risk attitudes measured when using the BVB and the more complex HL or not. Moreover, we verified if the display formats of the HL and BVB influence the participants' risk attitude measured. Regarding inconsistency rates, we tested in a within-subject design if there are significant differences between both lotteries and we examined in a between-subject design whether using different display formats of one lottery task leads to significant differences or not.

Results can be summarized as follows: (i) The participants are classified as slightly risk averse in the HL and the BVB, while using both the textual display format as well as the visual display format. Thereby, a within-subject comparison between the lottery tasks results in significantly different risk attitudes measured in both display formats. For both display formats, the participants' responses are more risk averse in the more complex HL than in the BVB. The choice of the methods to determine the risk attitude has a significant influence on the elicit risk attitude. (ii) A between-subject comparison between the participants in both display formats indicates that the group provided with the visualization gives more risk averse responses in both lottery tasks. However, the tested differences are only significant for the BVB. (iii) Regarding the group faced with the text format, a within-subject test shows that the BVB is better understood than the widespread HL. The inconsistency rates of the HL and the BVB are significantly different. With respect to the group provided with the visual form, there are no significant differences between the inconsistency rates. (iv) The visualization of the HL leads to a significantly lower inconsistency rate compared to the text format. In contrast, there is no significant difference for the BVB. The visual display format of the more complex HL visibly leads to a better understanding of the elicitation method. The seemingly more understandable BVB cannot be simplified by using the visual form design. In summary, both elicitation methods determine a different risk attitude of the participants, even if the framework of the BVB has been adapted to the HL in this experiment and the possible payout amounts of the BVB were changed in a way that they corresponded to the CRRA-values of the HL. The high inconsistency rate of the HL applied with textual display format raises the question whether the HL should be referred to as the "gold standard". The visualization of the HL leads to a significantly better understanding, whereas the risk attitude measured does not change. To the best of our knowledge, this is the first study supporting the visualization of the HL with significant results. The visualization of the BVB does not significantly reduce the inconsistency rate although the risk attitude measured is significantly different. Hence, the visual form design of the BVB appears to be less suitable, at least in this particular context.

Elicitation methods to determine the participants' risk attitude have been simplified to a visual display format in numerous research studies. These determined CRRA-values have often been compared with CRRA-values of other studies. The fact that the visualization of a lottery task may result in a significantly different measured risk attitude of the participants challenges whether these comparisons are valid for different display formats. Furthermore, the use of visual display formats does not automatically lead to the desired simplification. For future research, it would be of interest to verify in which elicitation methods the visualization would lead to a reduction of the inconsistency rate and for which groups of participants a visual display format might be advantageous.

References

ABDELLAOUI, M., A. DRIOUCHI and O. L'HARIDON (2011): Risk Aversion Elicitation: Reconciling Tractability and Bias Minimization. In: Theory and Decision 71: 63-80.

ANDERSEN, S., G. HARRISON, M. LAU and E. RUTSTRÖM (2006): Elicitation using multiple price list formats. In: Experimental Economics 9: 383-405.

ANDERSON, L.R. and J.M. MELLOR (2009): Are risk preferences stable? Comparing an experimental measure with a validated survey-based measure. In: Journal of Risk and Uncertainty 39, 137–160.

ASHTON, R.H. and S.S. KRAMER (1980): Students as Surrogates in Behavioral Research: Some Evidence. In: Journal of Accounting Research 18: 1-15.

BINSWANGER, H.P. (1980): Attitudes toward risk: Experimental measurement in rural India. In: American Journal of Agricultural Economics 62(3): 395-407.

BONNER, S.E. and G.B. SPRINKLE (2002): The Effects of Monetary Incentives on Effort and Task Performance: Theories, Evidence, and a Framework for Research. In: Accounting, Organization and Society 27: 303-345.

BOSCH-DOMÈNECH, A. and J. SILVESTRE (2013): Measuring risk aversion with lists: a new bias. In: Theory and decision 75: 465-496.

BOUGHERARA, D. X. GASSMANN and L. PIET (2011): Eliciting Risk Preferences: A Field Experiment on a Sample of French Farmers. Paper presented at EAAE 2011 International Congress, Zurich, 30 August - September 2.

BRICK, K., M. VISSER and J. BURNS (2012): Risk Aversion: Experimental Evidence from South African Fishing Communities. In: American Journal of Agricultural Economics 94(1): 133-152.

BRUNER, D.M. (2011): Multiple switching behavior in multiple price lists. In: Applied Economics Letters 18: 417-420.

CARDENAS, J.C. and J. CARPENTER (2008): Behavioral Development Economics: Lessons from Field Labs in the Developing World. In: Journal of Development Studies 44: 311-338.

CHARNESS, G., U. GNEEZY and A. IMAS (2013): Experimental methods: Eliciting risk preferences. In: Journal of Economic Behavior & Organization 87: 43-51.

CHARNESS, G. and A. VICEISZA (2015): Three risk-elicitation methods in the field: Evidence from rural Senegal. In: Review of Behavioral Economics, forthcoming.

DAVE, C., C.C. ECKEL, C. JOHNSON and C. ROJAS (2010): Eliciting risk preferences: When is simple better? In: Journal of Risk and Uncertainty 41: 219-243.

DE BRAUW, A. and P. EOZENOU (2011): Measuring Risk Attitudes among Mozambican Farmers. In: IFPRI Harvest Plus Working Paper 6, Washington DC.

DOHMEN, T.J., A. FALK, D. HUFFMAN and U. SUNDE. 2010. "Are risk aversion and impatience related to cognitive ability?" American Economic Review 100: 1238-1260.

DOHMEN, T.J., A. FALK, D. HUFFMAN, U. SUNDE, J. SCHUPP and G. WAGNER (2011): Individual Risk Attitudes: Measurement, Determinants and Behavioral Consequences. In: Journal of the European Economic Association 9: 522-550.

ECKEL, C.C. and P.J. GROSSMAN (2002): Sex differences and statistical stereotyping in attitudes toward financial risk. In: Evolution and Human Behavior 23: 281-295.

ECKEL, C.C. and P.J. GROSSMAN (2008): Forecasting risk attitudes: An experimental study using actual and forecast gamble choices. In: Journal of Economic Behavior & Organization 68: 1-17.

GALARZA, F. (2009): Choices under risk in rural Peru. In: Munich Personal RePEc Archive Working Paper 17708, Munich.

HILL, R.V. and A. VICEISZA (2012): A field experiment on the impact of weather shocks and insurance on risky investment. In: Experimental Economics 15(2): 341-371.

HOLT, C.A. and S.K. LAURY (2002): Risk aversion and incentive effects. In: The American Economic Review 92: 1644-1655.

IHLI, H.H., B. CHIPUTWA and O. MUSSHOFF (2013): Does Changing Probabilities or Payoffs in Lottery-Choice Experiments Matter? Evidence from Rural Uganda. In: Global Food Discussion Paper Series No. 24.

ISAAC, R.M. and D. JAMES (2000): Just who are you calling risk averse? In: Journal of Risk and Uncertainty 20: 177-187.

JACOBSON, S. and R. PETRIE (2009): Learning from mistakes: What do inconsistent choices over risk tell us? In: Journal of Risk and Uncertainty 38: 143-158.

LÉVY-GARBOUA, L., H. MAAFI, D. MASCLET and A. TERRACOL (2012): Risk aversion and framing effects. In: Experimental Economics 15(1): 128-144.

MAART-NOELCK, S.C. and O. MUSSHOFF (2013): Measuring the risk attitude of decision-makers: are there differences between groups of methods and persons? In: Australian Journal of Agricultural and Resource Economics 58(3): 336-352.

PRICE, T.J. and M.E. WETZSTEIN (1999): Irreversible investment decisions in perennial crops with yield and price uncertainty. In: Journal of Agricultural and Resource Economics 24: 173-185.

PURVIS, A., W.G. BOGGESS, C.B. MOSS and J. HOLT (1995): Technology Adoption Decisions under Irreversibility and Uncertainty: 'An Ex Ante' Appproach. In: American Journal of Economics 77: 541-551.

REYNAUD, A. and S. COUTURE (2012): Stability of Risk Preference Measure: Results from a Field Experiment on French Farmers. In: Theory and Decision 73: 203-221.

TANAKA, T., C.F. CAMERER and Q. NGUYEN (2010): Risk and time preferences: Experimental and household survey data from Vietnam. In: American Economic Review 100: 557-571.

WEBER, E.U., A.R. BLAIS and N.E. BETZ (2002): A Domain-Specific Risk Attitude Scale: Measuring Risk Perceptions and Risk Behaviors. In: Journal of Behavioral Decision Making 15: 263-290.