Risk Aversion and Inconsistencies - Does the Choice of Risk Elicitation Method and Display Format Influence the Outcomes?

Golo-Friedrich Bauermeister and Oliver Musshoff

Georg-August University Goettingen, Department of Agricultural Economics and Rural Development, Platz der Goettinger Sieben 5, 37073 Goettingen, Germany

Author email: Golo-Friedrich.Bauermeister@agr.uni-goettingen.de

Selected Paper prepared for presentation at the 2016 Agricultural & Applied Economics Association Annual Meeting, Boston, Massachusetts, July 31-August 2

Risk aversion and inconsistencies - Does the choice of risk elicitation method and display format influence the outcomes?

Abstract

In the past decade, many studies measured individual risk attitude with different elicitation methods in a within-subject design and found significant disparity 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 that compare the inconsistency rate across different elicitation methods in a within-subject design. Therefore, we intrapersonally compare the inconsistency rate and the risk attitude of German agricultural students in two different lottery tasks: 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 the visualization of a lottery task for a better understanding results in differences in the elicited risk attitude and can ultimately lead to the desired reduction of the inconsistency rate. Results show that the elicited risk attitudes measured by the different lottery tasks are significantly different in both display formats since the participants' responses are more risk averse in the more complex Holt-and-Laury task. 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:

Between-subject design, Brick-Visser-Burns task, display formats, Holt-and-Laury task, inconsistency rate, risk attitude, within-subject design

JEL classifications

C91 – D80 – O10

Risk and uncertainty play an important role in a variety of financial and economic decisions (Abdellaoui, Driouchi and L'Haridon 2010). 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), crop insurance markets (Hill and Viceisza 2012) or conservation intervention (Winter-Nelson and Amegbeto 1998). Consequently, 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. On the one hand, these studies included econometric approaches, where the individual risk attitude was estimated using empirically observed data. On the other hand, many studies were based on experimental methodology that 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. According to Andersen et al. (2006), positive advantages of the MPL include both the fact that participants can see relatively easily that a true answer is for their own advantage, and the fact that the MPL is easy to implement. Negative aspects, however, consist of the fact that it is only possible to determine intervals of the individual risk attitude, and that the framework may induce participants to start choosing the safer option and switch to the riskier one in the middle of the list (Anderson et al. 2006). This problem can be solved by highlighting individual sequences (Jacobson and Petrie 2009) or by randomization (Kirby, Petry and Bickel 1999), which, however, increases the complexity of the method. 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), the contextual statements by Weber, Blais and Betz (2002) or the willingness-to-pay method by Kahneman, Knetsch and Thaler (1990).

The problem of the various aforementioned measurement methods is that they often lead to different results for the individual risk attitude measured (Isaac and James 2000; Andersen et al. 2006; Dave et al. 2010). During the last decade, the Holt-and-Laury task (HL) has often been applied to measure risk attitude and has therefore been characterized as the "gold standard" in risk elicitation (Anderson and Mellor 2008; Nielsen and Zeller 2013). The HL is a method that 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 lottery task to measure the individual risk attitude that allows inferences about the level of the inconsistency rate. The Brick-Visser-Burns task (BVB) is based on the idea by Dohmen et al. (2005; 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. To improve the understanding, methods to elicit risk attitudes are illustrated in visual display formats in some studies (Bougherara, Gassmann and Piet (2011); Johnson, Eckel and Engle-Warnick (2007). A lower inconsistency rate results in more reliable data. Particularly in developing countries, where inconsistency rates are usually high, reducing complexity and improving understanding play an important role.

Consequently, the present study focuses particularly on four different aspects. First, the risk attitude measuring methods HL and BVB are compared in a within-subject design by using a text format and a visual format. Secondly, it examines in a between-subject design if the determined risk attitude differs within one elicitation method for different display formats (textual and visual format). Third, it investigates how high the inconsistency rates are for both lotteries in the different display formats and if there are significant differences between the HL and the BVB regarding the inconsistency rate for equal display formats. Fourth, it analyzes if there are significant differences in the inconsistency rates of the two display formats of one method.

Previous studies show that there are indeed matches in the behavior of real decision-makers and students regarding the information processing and the associated decision-making processes (Ashton and Kramer 1980). As students normally have comparatively low opportunity costs, they are relatively easy to recruit and to remunerate in an incentive-compatible way (Falk and Fehr 2003). For this reason, the present study is based on a convenient group of students. In order to gather data for this study, students were asked to participate in both incentive-compatible lottery tasks. The order of the two lotteries as well as the display format, which remained the same in both lotteries for each individual, was randomized.

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 critical risk aversion coefficient are adjusted to the HL. When comparing different measurement techniques, as previously done by Andersen et al. (2006) or Dave et al. (2010), different critical risk aversion coefficients and different frameworks are used, which may lead to comparison biases (Andersen et al. 2006). Reynaud and Couture (2012), who conducted a non-incentivized experiment, perform an adjustment of the critical risk aversion coefficients. However, the used EG-method was not designed as a MPL and, therefore, has a different framework than the HL that did not allow the determination of inconsistency rates. 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 figure 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.

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 widespread HL and the survey technique 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) conduct a correlation analysis between an incentive-compatible HL and a self-assessment as well as a business-related context-based statement. In this case, the risk attitude has a significantly positive correlation. In an experiment in Uganda, Ihli, Chiputwa and Musshoff (2013) compare a visualized HL with a visualized BVB and demonstrate that there is a statistically significant difference between the observed risk attitudes of the incentive-compatible methods. The problem of the comparison of different measuring methods is the fact that the methods usually have different frameworks and risk aversion coefficients, or that 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. In doing so, the EG method does not allow for a determination of inconsistent participants and, furthermore, the experiment is 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 row-making behavior, and since the framework of the BVB can be adjusted, we compare the "golden 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, popular measurement techniques are often 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 Johnson, Eckel and Engle-Warnick (2007) visualize the HL by using graphs. For each row, the authors present two circles that stand for the two options. Depending on the probabilities, the corresponding proportions of the circles 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 forms vs. visualization": There are not any significant differences between the individual risk aversion coefficients determined by differently displayed HL methods (BVB methods).

Methods structured in the MPL-format, such as the HL and the BVB, have 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 experiment 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), by varying the payout amounts (instead of the probabilities, as done by Holt and Laury (2002)) and by showing the lottery pairs in sequences, determine an inconsistency rate of 55 percent for a sample of the Rwandan population. Brick, Visser and Burns (2012), who work with fixed probabilities of 50 and 100 percent and who also do not change the probabilities but only the payout amounts for the option, find out that 41 percent of the participating fishers from the west coast of South Africa change the option at least twice. In a survey in Mozambique, De Brauw and Eozenou (2011) detect 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 receive. The low inconsistency rate in Mozambique is comparable with the rate of industrialized countries. For a sample of 881 adult Canadians, Dave et al. (2010) determine an inconsistency rate of 8.5 percent. 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 Silvestre (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 advanced method by Brick, Visser and Burns (2012) of 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 (Galarza 2009; Charness, Gneezy and Imas 2013). The reason for this is the assumption that the participants did not completely understand the experiment and thus reflect a wrong risk attitude by showing an inconsistent response behavior (Galarza 2009; Charness, Gneezy and Imas 2013). If a majority of the participants provides inconsistent responses, the validity of the examined risk attitude of the sample is vastly diminished. It is therefore important to adapt the methods to the cognitive skills of the participants and to present them as clearly as possible. Furthermore, it is useful to choose 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 methods should take into account inconsistent individuals, 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 response 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), Reynaud and Couture (2012) 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 a method that is based on textual form and one that is based on visualization. Thus, the corresponding hypothesis is:

H4 "inconsistency rates - textual forms vs. visualization": There is not any significant difference in the inconsistency rate of different presentation forms of the HL method (BVB method).

Methodology

In the experiment, students face two different 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 lotteries have an incentive-compatible design.

Holt-and-Laury task

For the HL, participants are asked to decide for one out of two options in ten different rows. All rows are presented to the participants at once. 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* and *figure 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) (cf. *figure 1*). Due to the lower difference between the possible payouts, option A is the more secure 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.

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 changes 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 risk 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 every row. A consistent person who changes, for instance, between the sixth and seventh row (HL-value = 6), is therefore assigned to a CRRA-coefficient of between 0.41 and 0.68. 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.

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 a glance. The participants are asked to choose from two options in various rows (cf. *table 2* and *figure 2*). In contrast to the HL, the possible payout amounts vary and the probabilities remain constant for the BVB. For simplicity reasons, the probability is 100 percent in option A and 50 percent in option B. In contrast to option B, there is no difference between the possible payout amounts in option A, and B is thus the secure alternative. In the visual form (cf. *figure 2*) 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].

For the option B, participants can win € 115.50 or € 3 with a probability of 50 percent in each row. 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. Participants who always decide for option A in the BVB, have a BVB-value of 10 and a CRRA-coefficient that is higher than 1.37. An inconsistent behavior can occur if participants change more than once between the options or decide for option B in the first row. In the latter, 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. When 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. 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.

Data collection

The survey for the determination of the risk attitude and the inconsistency rates was conducted online in 2014. Students from the faculty of agricultural sciences were invited to participate in the experiment via e-mail lists and the information board of the university. The experiment consisted of three parts and took 20 minutes. Each participant carried out once the HL as well as the BVB. The order of the elicitation methods was randomized. Some participants therefore 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 presentation forms, there existed two groups. One group was faced with both lotteries in textual form and the other group was provided with a visualized form of the lotteries. It was not possible that one participant had to deal with different presentation forms. 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 explanation 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. In doing so, 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

Results

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

Verification of H1

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

The distributions of the CRRA-values from the 149 participants who were provided with the textual form are illustrated in *figure 3*. 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 of these participants always choose option B and thus, no CRRA-value can be assigned. To allow for an intrapersonal comparison of the medians, the six participants with the missing CRRA-value are excluded.

The CRRA-means average 0.20 in the text format of the HL and 0.07 in the textual form 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. Consequently, we work with the non-parametric Wilcoxon signed-rank test to compare the CRRA-means. This intrapersonal test reveals that although the CRRA-means of both lotteries are significantly different (p = 0.003), the participants' responses are more risk averse in the HL. Thus, for the textual form, the hypothesis H1 can be rejected.

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

The distributions of the CRRA-values from the 158 participants who face the visual form of the lotteries are illustrated in *figure 4.* 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

The CRRA-means average 0.29 in the visual form of the HL and 0.16 in the visualization of the BVB reflecting a small risk aversion in each case (*cf. table 4*). 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 although the CRRA-means of both lotteries are significantly different ($p < 0.001$) in the group provided with the visualized lotteries, the participants made on average more risk averse responses in the HL. Hence, hypothesis H1 can also be rejected for the visual form.

Verification of H2

"Risk attitudes in the HL and the BVB - textual forms vs. visualizations"

The different CRRA-means and the respective standard deviations of both lotteries 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-values of one lottery task in different display formats. The values indicate that the group faced with the visual form made more risk averse replies 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 ($p = 0.188$). 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 ($p = 0.047$). As a consequence, the hypothesis H2 can be rejected for the BVB, whereas it can be confirmed for the HL.

Verification of H3

"Inconsistency rates for text format - 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 additionally 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.

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 applied a McNemar's test for related samples, which revealed a significant difference ($p < 0.001$) in the inconsistency rates between both lotteries (cf. *table 5*). Hence, the hypothesis H3 can be rejected for the textual form design.

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

Apart from the inconsistency rates of the textual form design, *table 4* also shows the inconsistency rates of the visual form. The classification is applied similarly to the textual form design. Twenty participants (12.66 percent) indicate an inconsistent behavior, whereas none of the participants always chose option A (group 3). In the BVB, since none of the participants chose option B, we classify 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.

Verification of H4

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

The results of H3 reveal that different kinds of lottery methods can lead to divergent inconsistency rates. *Table 5* indicates the inconsistency rates of the different form designs and lotteries. 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. Apparently, 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 ($p = 0.585$) 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-method and confirmed for the BVB-method.

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. The problem is that different elicitation methods lead to varying participants' risk attitude measured 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 "golden 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 an intrapersonal comparison 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 intrapersonally tested if there are significant differences between both lotteries and we interpersonally examined whether using different display formats of one lottery task leads to significant differences or not.

Results can be summarized as follows: (i) The participants were 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, an intrapersonal comparison between the lottery tasks resulted in significantly different risk attitudes measured in both display formats. For both display formats, the participants' responses were 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) An interpersonal comparison between the participants in both display formats indicates that the group that is provided with the visualization makes more risk averse responses in both lotteries. However, the tested differences are only significant for the BVB. (iii) Regarding the group faced with the text format, an intrapersonal test showed that the BVB was 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 was no significant difference for the BVB. The visual display format of the more complex HL visibly leads to a better understanding of the method. The apparently more understandable BVB cannot be simplified by using the visual format. In summary, both methods elicit 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 “golden 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 article 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 display format of the BVB appears to be less suitable, at least in this particular context.

Methods to elicit the participants’ risk attitude had been simplified by a visual display format in numerous research studies. These CRRA-values measured had often been compared with CRRA-values of other studies. The fact that the visualization of a method may result in a significantly different measured risk attitude of the participants leads to the question 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 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.

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

[2]On the margins of the lotteries, where there are no CRRA-spans, we work with CRRA-values of -2 and 2 as done by Reynaud and Couture (2012).

References

Abdellaoui, M., A. Driouchi, and O. L'Haridon. 2011. "Risk Aversion Elicitation: Reconciling Tractability and Bias Minimization." *Theory and Decision* 71: 63-80.

Andersen, S., G. Harrison, M. Lau, and E. Rutström. 2006. "Elicitation using multiple price list formats." *Experimental Economics* 9: 383-405.

Ashton, R.H., and S.S. Kramer. 1980. "Students as Surrogates in Behavioral Research: Some Evidence." *Journal of Accounting Research* 18: 1-15.

Binswanger, H.P. 1980. "Attitudes toward risk: Experimental measurement in rural India." *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." *Accounting, Organization and Society* 27: 303-345.

Bosch-Domènech, A., and J. Silvestre. 2013. "Measuring risk aversion with lists: a new bias." *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." *American Journal of Agricultural Economics* 94(1): 133-152.

Bruner, D.M. 2011. "Multiple switching behavior in multiple price lists." *Applied Economics Letters* 18: 417-420.

Cardenas, J.C., and J. Carpenter. 2008. "Behavioral Development Economics: Lessons from Field Labs in the Developing World." *Journal of Development Studies* 44: 311-338.

Charness, G., U. Gneezy, and A. Imas. 2013. "Experimental methods: Eliciting risk preferences." *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." *Review of Behavioral Economics*, forthcoming.

Dave, C., C.C. Eckel, C. Johnson, and C. Rojas. 2010. "Eliciting risk preferences: When is simple better?" *Journal of Risk and Uncertainty* 41: 219-243.

De Brauw, A., and P. Eozenou. 2011. "Measuring Risk Attitudes among Mozambican Farmers." IFPRI Harvest Plus Working Paper 6, Washington DC.

Dohmen, T.J., A. Falk, D. Huffman, J. Schupp, U. Sunde, and G.G. Wagner. 2005. "Individual Risk Attitudes: New Evidence from a Large, Representative, Experimentally-Validated Survey." IZA discussion paper No. 1730.

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." *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." *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.” *Journal of Economic Behavior & Organization* 68: 1-17.

Falk, A., and E. Fehr. 2003. “Why labour market experiments?” *Labour Economics* 10: 399-406.

Galarza, F. 2009. “Choices under risk in rural Peru.” 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.” *Experimental Economics* 15(2): 341-371.

Holt, C.A., and S.K. Laury. 2002. “Risk aversion and incentive effects” *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.” Global Food Discussion Paper Series No. 24.

Isaac, R.M., and D. James. 2000. “Just who are you calling risk averse?” *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?” *Journal of Risk and Uncertainty* 38: 143-158.

Johnson, C., C.C. Eckel, and J. Engle-Warnick. 2007. “An adaptive instrument to elicit risk preferences.” Working paper, University of Montreal.

Kahneman, D., J.L. Knetsch, and R.H. Thaler. 1990. "Experimental tests of the endowment effect and the Coase theorem." *Journal of Political Economy* 98: 1325-1348.

Kirby, K.N., N. Petry, and W.K. Bickel. 1999. "Heroin Addicts Have Higher Discount Rates for Delayed Rewards than Non-drug-using Controls." *Journal of Experimental Psychology: General* 128: 78-87.

Lévy-Garboua, L., H. Maafi, D. Masclet, and A. Terracol. 2012. "Risk aversion and framing effects." *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?" *Australian Journal of Agricultural and Resource Economics* 58(3): 336-352.

Nielsen, T., and M. Zeller. 2013. "The impact of shocks on risk preference changes between seasons for smallholder farmers in Vietnam." Paper presented at 53rd Annual Conference of the German Society of Economic and Social Sciences in Agriculture. 25-27 September.

Price, T.J., and M.E. Wetzstein. 1999. "Irreversible investment decisions in perennial crops with yield and price uncertainty." *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." *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." *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." *American Economic Review* 100: 557-571.

Winter-Nelson, A., and K. Amegbeto. 1998. "Option values to conservation and agricultural price policy: application to terrace construction in Kenya." *American Journal of Agricultural Economics* 80(2): 409-418.

Weber, E.U., A.R. Blais, and N.E. Betz. 2002. "A Domain-Specific Risk Attitude Scale: Measuring Risk Perceptions and Risk Behaviors" *Journal of Behavioral Decision Making* 15: 263-290.

Table 1. Payout Matrix of the HL [a)]

row	option A	option B	EV (A)	EV (B)	CRRA-Span [b)]
1	with a 10% gain of € 60 with a 90% gain of € 48	with a 10% gain of € 115.50 with a 90% gain of € 3	€ 49.20	€ 14.25	$r \leq -1.71$
2	with a 20% gain of € 60 with a 80% gain of € 48	with a 20% gain of € 115.50 with a 80% gain of € 3	€ 50.40	€ 25.50	$-1.71 \leq r \leq -0.95$
3	with a 30% gain of € 60 with a 70% gain of € 48	with a 30% gain of € 115.50 with a 70% gain of € 3	€ 51.60	€ 36.75	$-0.95 \leq r \leq -0.49$
4	with a 40% gain of € 60 with a 60% gain of € 48	with a 40% gain of € 115.50 with a 60% gain of € 3	€ 52.80	€ 48	$-0.49 \leq r \leq -0.14$
5	with a 50% gain of € 60 with a 50% gain of € 48	with a 50% gain of € 115.50 with a 50% gain of € 3	€ 54	€ 59.25	$-0.14 \leq r \leq 0.15$
6	with a 60% gain of € 60 with a 40% gain of € 48	with a 60% gain of € 115.50 with a 40% gain of € 3	€ 55.20	€ 70.50	$0.15 \leq r \leq 0.41$
7	with a 70% gain of € 60 with a 30% gain of € 48	with a 70% gain of € 115.50 with a 30% gain of € 3	€ 56.40	€ 81.75	$0.41 \leq r \leq 0.68$
8	with a 80% gain of € 60 with a 20% gain of € 48	with a 80% gain of € 115.50 with a 20% gain of € 3	€ 57.60	€ 93	$0.68 \leq r \leq 0.97$
9	with a 90% gain of € 60 with a 10% gain of € 48	with a 90% gain of € 115.50 with a 10% gain of € 3	€ 58.80	€ 104.25	$0.97 \leq r \leq 1.37$
10	with a 100% gain of € 60 with a 0% gain of € 48	with a 100% gain of € 115.50 with a 0% gain of € 3	€ 60	€ 115.50	$1.37 \leq r$

[a)] Participants did not see the last three columns.

[b)] Assuming: $U(x) = x^{(1-r)}/(1-r)$

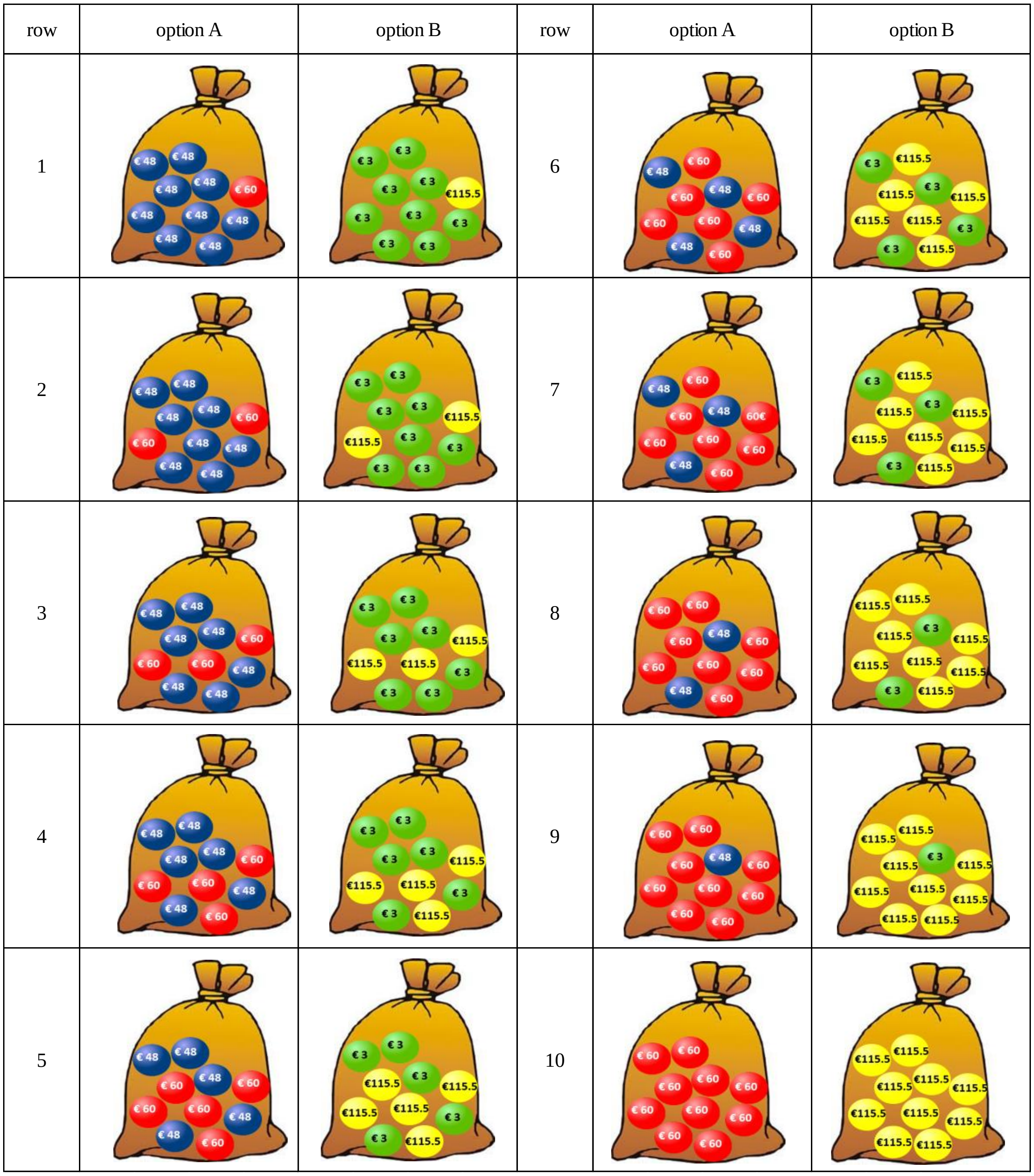

Figure 1. Visualized HL

Table 2. Payout Matrix of the BVB [a)]

row	option A	option B	EV (A)	EV (B)	CRRA-span [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	$r \leq -1.72$
3	with a 100% gain of € 81.00	with a 50% gain of € 115.50 with a 50% gain of € 3.00	€ 81.00	€ 59.25	$-1.72 \leq r \leq -0.95$
4	with a 100% gain of € 72.50	with a 50% gain of € 115.50 with a 50% gain of € 3.00	€ 72.50	€ 59.25	$-0.95 \leq r \leq -0.48$
5	with a 100% gain of € 63.50	with a 50% gain of € 115.50 with a 50% gain of € 3.00	€ 63.50	€ 59.25	$-0.48 \leq r \leq -0.13$
6	with a 100% gain of € 54.00	with a 50% gain of € 115.50 with a 50% gain of € 3.00	€ 54.00	€ 59.25	$-0.13 \leq r \leq 0.15$
7	with a 100% gain of € 43.00	with a 50% gain of € 115.50 with a 50% gain of € 3.00	€ 43.00	€ 59.25	$0.15 \leq r \leq 0.41$
8	with a 100% gain of € 31.00	with a 50% gain of € 115.50 with a 50% gain of € 3.00	€ 31.00	€ 59.25	$0.41 \leq r \leq 0.68$
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	$0.68 \leq r \leq 1$
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	$1 \leq r \leq 1.37$

[a)] Participants did not see the last three columns.

[b)] Assuming: $U(x) = x^{(1-r)}/(1-r)$

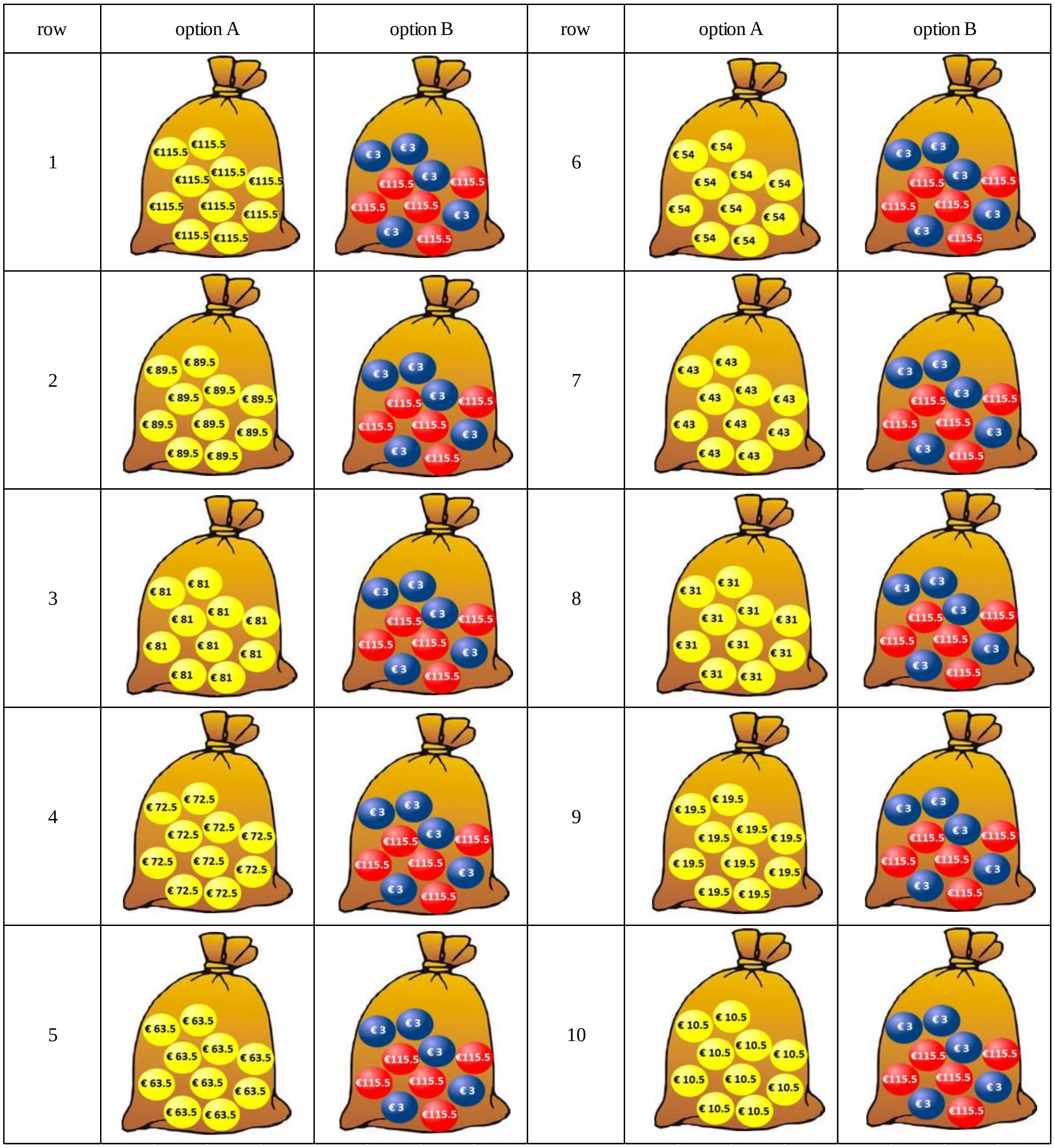

Figure 2. Visualized BVB

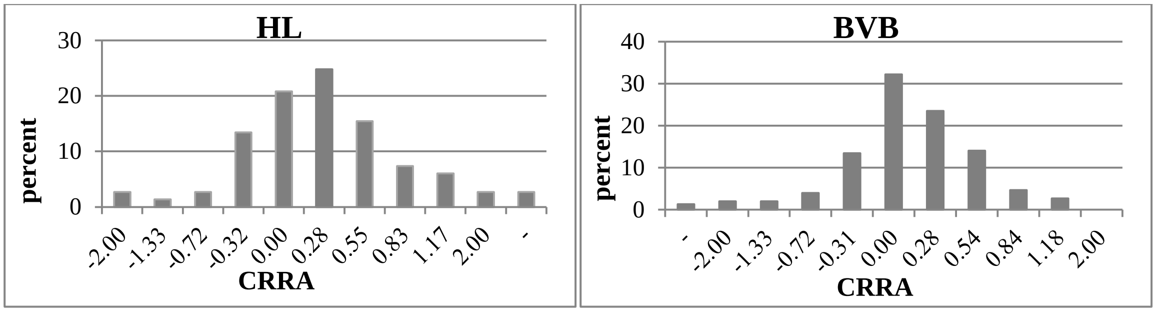

Figure 3. Distribution of CRRA-values for textual form (N=149)[a)]

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

	CRRA-value HL	CRRA-value 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

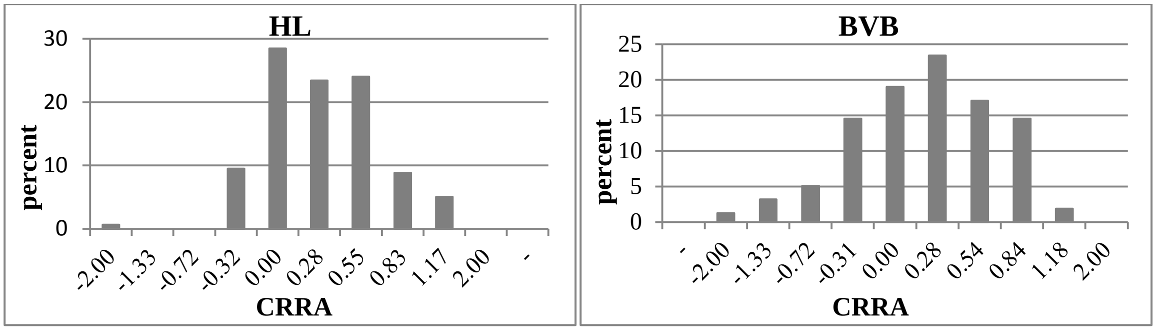

Figure 4. Distribution of CRRA-values for visual form (N=158)

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: More than one 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: More than one switch or directly from option B to option A	13	8.72%	19	12.03%

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	-