COLOGNE GRADUATE SCHOOL
in Management, Economics and Social Sciences

Vol 2 | No 3

CGS WORKING PAPER

Measuring Individual Risk Attitudes in the Lab: Task or Ask? An Empirical Comparison

Jan-Erik Lönnqvist
Markku Verkasalo
Gari Walkowitz
Philipp C. Wichardt

Cologne Graduate School
in Management, Economics
and Social Sciences
Albertus-Magnus-Platz
50923 Köln
www.cgs.uni-koeln.de

University of Cologne

Measuring Individual Risk Attitudes in the Lab: Task or Ask? An Empirical Comparison*

Jan-Erik Lönnqvist, Markku Verkasalo
Institute of Behavioural Sciences, University of Helsinki

Gari Walkowitz
Department of Management, University of Cologne
Laboratory for Experimental Economics, University of Bonn

Philipp C. Wichardt[a]
Department of Economics, University of Bonn

This version: March 13, 2011

Abstract This paper compares two prominent empirical measures of individual risk attitudes — the Holt and Laury (2002) lottery-choice task and the multi-item questionnaire advocated by Dohmen, Falk, Huffman, Schupp, Sunde and Wagner (forthcoming) — with respect to (a) their within-subject stability over time (one year) and (b) their correlation with actual risk-taking behaviour in the lab — here the amount sent in a trust game (Berg, Dickaut, McCabe, 1995). As it turns out, the measures themselves are uncorrelated (both times) and, most importantly, only the questionnaire measure exhibits test-re-test stability ($\rho = .78$), while virtually no such stability is found in the lottery-choice task. In addition, only the questionnaire measure shows the expected correlations with a Big Five personality measure and is correlated with actual risk-taking behaviour. The results suggest that the questionnaire is the more adequate measure of individual risk attitudes for the analysis of behaviour in economic (lab) experiments. Moreover, with respect to trust, the high re-test stability of trust transfers ($\rho = .70$) further supports the conjecture that trusting behaviour indeed has a component which itself is a stable individual characteristic (Glaeser, Laibson, Scheinkman and Soutter, 2000).

Keywords: Risk Attitudes, Trust, Personality, Lab Experiments
JEL codes: D81, C91, Z10

*We are grateful to Vincent Crawford, Thomas Dohmen, Armin Falk, Ernst Fehr, Glenn Harrison, Heike Hennig-Schmidt, Sven Hoeppner, Bernd Irlenbusch, Daniel Schunk, Avner Shaked, Jörn Tenhofen and Daniel Wiesen for helpful comments and discussions. Moreover, we would like to thank Julia Berndt, Andreas Born, Antti Hulsi and Rainer Rilke for helpful research assistance. Financial support of the Academy of Finland (127641) and the Deutsche Forschungsgemeinschaft (PN74548) is gratefully acknowledged.

[a]Corresponding author: Economic Theory 3, University of Bonn, Adenauerallee 24-26, D-53113 Bonn, Germany; e-mail: philipp.wichardt@uni-bonn.de.

1 Introduction

Risk-aversion has long since been a standard ingredient of economic theory (seminal works being Arrow, 1965, and Pratt, 1964). Moreover, with the growing body of experimental studies in the social sciences in general and in economics in particular (see Falk and Heckman, 2009, for a discussion), individual risk attitudes have also been linked empirically to a good many behavioural patterns. For example, Goeree, Holt and Palfrey show how the distribution of bids in a private-values auction experiment, including the prevalence of overbidding, can be captured in a quantal response equilibrium model with risk averse bidders (Goeree et al., 2002), and how deviations from Nash-equilibrium in a generalised matching pennies game can be explained if the subjects' risk aversion is taken into account (Goeree et al., 2003). Dohmen and Falk (forthcoming) in turn find that individuals self-select into different types of payment schemes according to their risk attitudes with more risk averse individuals choosing less performance dependent payment schemes (see also Guiso and Paiella, 2005). And these are but some of the many examples that have been discussed in the literature (see Harrison and Rutström, 2008, for further examples).

Despite the empirical relevance of individual risk attitudes, however, there is still an ongoing debate as to their adequate elicitation (cf. Harrison and Rutström, 2008). Among other aspects — including specific procedural issues as well as the general question whether risk aversion is at all a plausible phenomenon for the small stakes which are commonly offered in the lab (cf. Rabin, 2000) — the relevance of proper incentives has given rise to discussions: Should risk preferences be inferred from incentivised behavioural measures such as lottery choice tasks (e.g. Holt and Laury, 2002) as commonly done in economics? Or should they rather be assessed using non-incentivised questionnaires based on so called Likert statements in which subjects specify their level of agreement to a certain statement as predominantly used in psychology (see Lauriola and Levin, 2001, for a historical review; see also Dohmen, Falk, Huffman, Schupp, Sunde and Wagner, forthcoming)? Of course, lottery choices in general will be easier to translate into formal indices and, hence, will be preferable when questions about structural parameters of utility functions are at issue (see, for example, Andersen, Harrison, Lau and Rutström, 2008a). Yet, also questionnaire measures, which have witnessed a growing popularity in recent years (e.g. Dohmen and Falk, forthcoming; Dohmen, Falk, Huffman and Sunde, 2008 and 2010), have their virtues as predictors of behaviour, especially when used in connection with laboratory

experiments — not least because they are both cheaper to gather and arguably easier to respond to without further instructions.

In the present study, we take up the question about the adequate elicitation of risk attitudes in connection with economic lab experiments and investigate the relative performance of a lottery-choice measure validated by Holt and Laury (2002) and a questionnaire measure advocated by Dohmen et al. (forthcoming) with regard to the correlation between them, their construct validity (being correlated with an external predictor of risk-taking behaviour, namely personality), and their test-re-test stability over time (one year). Moreover, we test the ability of the two types of measures to predict actual behaviour in a domain that is typically associated with risk.[1] In doing so, we operationalise risk-taking behaviour in the lab via a standard trust / investment game (Berg, Dickhaut and McCabe, 1995) taking the amount sent by first movers as a measure for their willingness to take risks. Although the question whether the investment decision in trust games eventually measures risk or trust (or both) is contentious — with some studies emphasising the role of risk (e.g. Snijders and Keren, 1998, Ben-Ner and Putterman, 2001) and others rather questioning it (e.g. Eckel and Wilson, 2004, or Bohnet and Zeckhauser, 2004) or arguing directly in favour of trust (e.g. Houser, Schunk and Winter, 2010) — it appeared to us as an intuitive item to analyse. It fact, the popular view that giving money to strangers is risky (also expressed, for example, in Kosfeld, Heinrichs, Zak, Fischbacher and Fehr, 2005) intuitively suggests that doing so should somehow be correlated with risk attitudes.[2] In conjunction with the mixed evidence on this correlation, it thus seemed appealing to try and scrutinise whether the type of risk measure applied may be crucial in this context.

In the experiment we conducted, subjects first played a standard one shot trust game – the main experimental task. Once that was completed, subjects were given the Holt and Laury (2002) lottery choice task and, eventually, the questions about individual risk-attitudes advanced by Dohmen et al. (forthcoming) as part of the post experimental questionnaire.[3] Moreover, in order to further analyse the connection

[1]Note that biases in the measurement – as might result from one of two measures is incentivised while the other is not – are not problematic for our purposes as we "only" intend to ascertain which of the two (well established) measures is most suited as a tool to relate observed behaviour in economic (lab) experiments to individual risk attitudes but do not draw any inferences about the shape of individual utility functions. For an interesting discussion of hypothetical biases in the elicitation of risk attitudes and a review of the respective literature, see Harrison (2006).

[2]For an instructive general discussion of the different facets of trust see, for example, Ben-Ner and Halldorsson (2010).

[3]Risk-measures were gathered after the trust game (the main behavioural task) and without

between risk attitudes and other fundamental personal characteristics, we also gathered data on the subjects' Big Five personality traits (Neuroticism, Extraversion, Openness to Experience, Agreeableness, Conscientiousness), which subsume a huge variety of personality attributes and provide a concise summary of stable individual differences in personality (Digman, 1990).[4] These traits had previously been shown to be correlated with risk-taking behaviour (Nicholson, Soane, Fenton-O'Creevy and Willman, 2005; or Lauriola and Levin, 2001), so that it seemed plausible to expect them to be correlated with the measures of individual risk attitudes under scrutiny, i.e. with behaviour in the lottery-choice task and the answers to the questionnaire. The data on personality where gathered prior to and independent of the actual (first) experiment, though, in order to obtain an additional external reference point against which to compare our later results. Finally, in order to further check on the stability of our results, we reran the actual experiment with (some of) the same subjects after about a year's time and compared test and re-test results.

The results show the following general patterns: (1) The two measures of risk-attitudes themselves are uncorrelated (both studies). (2) Both the questionnaire measure (all single items and a factor that was statistically extracted from the questionnaire measure about risk attitudes) and individual risk-taking (trusting) behaviour exhibit a high and significant test-re-test stability, while virtually no such stability is present in the lottery-choice task. (3) The factor derived from the questionnaire as well as all single items are correlated with the Big Five personality measure, while behaviour in the lottery-choice task is mostly not. (4) The investment decision in the trust game is *correlated* with the risk factor as well as with some of the single items, but is virtually *uncorrelated* with the subjects' choices in the Holt and Laury lottery-choice task (this holds for both test and re-test); and we do not find any evidence for hedging or reference point effects that could explain the lack of correlation with the subjects' lottery choices (see Section 3 for details).

From these observations, we conclude that the type of multi-item questionnaire advocated by Dohmen et al. (forthcoming) is a useful and reliable tool to measure

being explicitly announced before in order to avoid any external indications that trusting behaviour is supposed to be risky. It is noteworthy in this respect that Houser et al. (2010), who ran a similar experiment combining a trust game with the Holt and Laury lottery-choice task in changing order, do not find any order effects. Inasmuch as possible, we also do not find any evidence in this direction in our data (see Section 3).

[4]The Big Five personality measure is also used, for example, by Bartling, Fehr, Marechal and Schunk (2005) who provide evidence that Agreeableness is related to a person's competitiveness.

individual risk attitudes in connection with economic lab experiments, whereas the behavioural measure, i.e. the Holt and Laury (2002) lottery-choice task, is rather problematic in this respect. Moreover, regarding the discussion about the determinants of trust (as measured in the Berg-Dickhaut-McCabe trust game), we interpret the data as suggesting that individual risk attitudes, once elicited in the appropriate way, indeed play a significant role in such behaviour (in particular attitudes towards "risk in financial matters," see Section 3 for details). Last but not least, we see the high overtime stability of investment behaviour in the trust game as a strong indication that trusting behaviour indeed has a component which is a stable individual characteristic as conjectured by Glaeser, Laibson, Scheinkman and Soutter (2000, p. 827).[5]

The remainder of this paper is structured as follows: In Section 2, we present the experimental design and procedures. The empirical results of our study are gathered and briefly discussed in Section 3. Section 4 concludes.

2 Experimental Design and Procedures

In the sequel, we describe the design of the lab-experiment (2.1) and the procedures of the overall study (2.2).

2.1 Design of the Experiment

The lab-experiment consisted of three different parts: The Holt and Laury lottery-choice task, a post experimental questionnaire including the questionnaire measure of risk attitudes, and the trust game played at the beginning of the experimental sessions. These are described below.

The Lottery-Choice Task

As an incentivised behavioural measure of individual risk attitudes, we used the well established lottery choice task proposed by Holt and Laury (2002).[6] This measure presents participants with ten choices between paired lotteries with payoffs ranging from .10 to 3.85 Euros (see Table 1 in Holt and Laury, 2002, P. 1645, who pay in US-Dollars instead of Euros). The first four pairs of lotteries give a higher expected payoff for the safer choice, whereas the latter six pairs give a higher expected payoff

[5] In fact, this observation also squares well with the results by Cesarini et al. (2008) which indicate a correlation between trusting behaviour and the subjects genetic code; see also Fehr (2009) for a recent summary on the literature on the determinants of trust.

[6] See also Harrison, Johnson, McInnes and Rutström (2005a), and Holt and Laury (2005).

for the more risky choice. Thus, the risk neutral option is to make four safe choices and then six risky choices. At the end of the experiment, one of the ten pairs of lotteries is randomly selected, the preferred lottery of that pair is conducted and the subject is payed accordingly. Following Holt and Laury (2002), we took the number of risky decisions the participant made as our measure of individual risk attitudes.[7]

Asking About Risk Attitudes - the questionnaire measure

As a non-incentivised questionnaire measure of individual risk-attitudes, we adhered to the questionnaire items presented by Dohmen et al. (forthcoming) and asked subjects to express their willingness to take risks in (1) general (2) driving (3) financial matters (4) sport and leisure activities (5) career decisions (6) health behaviours and (7) trusting strangers, using an eleven-point scale with zero indicating complete unwillingness to take risks and ten indicating complete willingness to take risks.[8] This type of questionnaire measure has been shown by Dohmen et al. to predict risky behaviors in various life domains. As our measure of individual risk attitudes, we used a factor which was statistically extracted from the answers to all seven questions (using a standard factor analysis), henceforth referred to as the *general risk factor* - GRF for short.[9] However, in the sequel, we occasionally report also correlations of single items.

The Trust Game – Risk-associated behaviour

The main experimental task for the subjects was the type of trust game first proposed by Berg, Dickhaut and McCabe (1995). In this game, there are two players, an investor and a responder. Initially, all players were given 10 Talers (1 Taler = .60 Euro). Responders just pocketed the money. By contrast, investors were given the opportunity to transfer any non-negative integer part of the endowment to a randomly matched and anonymous responder. All transfers were tripled by the experimenter. Upon receiving the transfer, responders had to decide how much to return to the investor; responders' back-transfers, which were elicited with the strategy method

[7]We also conducted the subsequent analyses using the point at which subjects crossed over from the safe choice to the risky choice as our measure of risk-taking (215 subjects, i.e. 93%, had such a crossover point and did not switch back from the risky choices to the safe choices). However, as the results were virtually identical, we report only the results for the former type of coding.

[8]The last item is not part of the questions initially developed for the German Socio-Economic-Panel (SOEP) and analysed by Dohmen et al. (forthcoming) but was added for purposes of the present study. In view of the later discussion, it is noteworthy that its inclusion into the factor model is not crucial for the results.

[9]Parallel analysis revealed that the questionnaire data can be modelled by just one principal component that accounted for 39% of the variance. In the present setting, results for the factor, which is a kind of (normalised) sophisticated average, are almost identical to those for the standard average, though.

(Selten, 1967), i.e. without any information about actual transfers, are not analysed in the present paper. The amount transferred by investors was taken as our measure of risk-taking behaviour. The trust game concluded with the elicitation of senders' and responder's first order beliefs which are partly reported in the sequel.

2.2 Procedures

The study was conducted in three phases. In the first phase, participants' personality profiles were gathered using a questionnaire posted on a University of Helsinki website;[10] the questionnaire also asked about some further items which are unrelated to the present study as well as some statistical data which allowed us to track subjects over time. The questionnaire was open from November 15^{th} to December 6^{th} 2007. Participants were contacted through ORSEE (Greiner, 2004) based on a mailing list of the Laboratory for Experimental Economics (University of Bonn) to which they had signed up in order to take part in research conducted at the lab. The mailing list consisted of around 3000 names. As an incentive to participate, participants were offered feedback on their personality profiles and a potential payoff from a short non-related decision task placed after the questionnaires; 945 subjects participated in this phase.

Out of the 945 subjects for which personality profiles had been gathered, a total of 232 responded to our subsequent invitation (which made no reference to phase 1) and participated in the experimental sessions conducted in phase 2 of our study at the Laboratory for Experimental Economics in Bonn between December 2007 and February 2008 (109 male, mean age 23.5 years, *SD* of age = 3.2).[11] Upon arriving at the lab, subjects were seated in computer cubicles and, thereby, randomly assigned to either of two roles for the trust game so that we ended up with $N = 116$ investors and equally many responders. The experiment, which was implemented using z-tree (Fischbacher, 2007), started once the subjects had read the instructions and (correctly) answered a set of control-questions. After decisions for the one shot trust game had been made, all subjects had to complete the Holt and Laury lottery-choice task and, eventually, the questionnaire measure about risk attitudes as part of the

[10] Personality was measured with a 60-item questionnaire (Konstabel, Lönnqvist, Walkowitz, Konstabel and Verkasalo, forthcoming) that captures the five personality factors conceptualised in the Five-Factor Model (Costa and McCrae, 1992).

[11] To be precise, 252 subjects participated in the experiments. However, due to technical problems, the first session happened to end up without the questionnaire measure of individual risk attitudes. The session was excluded for the later analysis (this also affects the re-test to which 2 additional subjects from the respective session turned up but were excluded from the analysis).

post experimental questionnaire (i.e. both measures had not been announced before as they were not part of the primary experiment). Feedback about the actual outcome of the trust game and the lottery was given individually at the end of the experiment.

In phase 3, we reinvited subjects from the 232 participants of phase 2 for a new experiment which took place on March 16 2009; 44 out of the 232 subjects from phase 2 participated in phase 3 (22 male, mean age 24.0 years, SD of age = 4.1).[11] The actual experiment was identical to the one described in phase 2 except for the fact that the assignment of subjects to their roles in the trust game was not random but determined by phase-2-assignments. At the very last, subjects were asked whether they had taken part in a similar experiment before (without suggesting that they had).[12]

3 Results

In the sequel, we report the results of the our experiments from phase 2 (the test) and phase 3 (the re-test) of our study focusing in particular on the correlation between the different measures of individual risk attitudes, their correlation with an external predictor of risk-taking behaviour and with investment decisions in the trust game. The results are collapsed across genders as we do not find gender to be correlated with either the lottery-choice behaviour, the general risk factor[13] or trusting behaviour.

3.1 The Test Study

As a first result, we find that the lottery-choice measure and the GRF themselves are uncorrelated (Spearman's $\rho = -.04$, $p > .10$)[14] suggesting that they measure different things.

In order to find out more about the reliability of the respective risk-attitude measures, we analysed the correlations with the Big Five personality factors; see Table 1. Based on a self-report measure of risk-attitudes in six decision domains Nicholson, Soane, Fenton-O'Creevy and Willman (2005) find that risk-taking is associated

[12]The additional items were added at the very end of the post-experimental questionnaire in order to avoid any interference with the earlier aspects of the experiment.

[13]Gender was correlated with one single item of the underlying questionnaire, though, namely the willingness to take risks in financial matters. For this the average for females was 2.642 as opposed to 3.553 for males (a Mann Whitney Test shows that those groups are different, $p < .01$)

[14]All reported correlations are Spearman rank correlation coefficients of two-sided tests; if not separately specified, Pearson correlation coefficients are almost identical.

positively with Extraversion and Openness, and negatively with Neuroticism, Agreeableness and Conscientiousness (see also Lauriola and Levin, 2001).[15] Accordingly, we expected similar relations to hold for the risk-attitude measures tested in the present study. Yet, as shown in Table 1, all five personality factors significantly predict the GRF but do not predict behaviour in the lottery-choice task in multiple regression analyses which model the respective risk-attitude as a composite of the Big Five. Thus, these findings cast some doubt on the reliability of the lottery-choice task as a measure of individual risk attitudes.

	(1)	(2)
Measure	LM	GRF
Neuroticism	-.0249	-.0672***
	(.0247)	(.0223)
Extraversion	.0333	.0500***
	(.0208)	(.0186)
Openness	-.0104	.0801***
	(.0228)	(.0205)
Agreeableness	.0323	-.0515**
	(.0248)	(.0223)
Conscientiousness	.0118	-.0638***
	(.0194)	(.0174)
Constant		-.143
		(.113)
Observations	232	232
(Pseudo) R^2	.009	.180

Table 1: Ordered Probit (1) and ordinary least squares (2) regression analyses predicting risk-attitude measures (Lottery Measure [LM] and Questionnaire Measure [GRF]) as a function of the Big Five personality factors (all subjects; $N = 232$). Table entries are the predictors' un-standardised B-coefficients from multiple regression analyses; numbers in brackets indicate standard errors. *** $p < .01$, ** $p < .05$, * $p < .10$.

As a second step of our anaylsis, we compared the relationship between the different measures of individual risk attitudes and observed trusting behaviour. As pointed out earlier, the question about the correlation between risk taking and trusting behaviour is contentious. However, as we show below, the GRF provides indeed a powerful predictor not only for current but also for future (re-test) trusting behaviour

[15]The six decision domains considered by Nicholson et al. are: recreation, health, career, finance, safety and social.

thereby giving further support to a position arguing for a connection between risk and trust.

To begin with, we focus solely on the results of the test study, though. For this, the average transfer in the trust game was 4.664 Taler (*SD* 3.095) with single transfers ranging from the full endowment of the truster (10 Taler) to none of it (0 Taler). Furthermore, only the questionnaire measure shows a significant correlation with investment behaviour in the trust game (ρ=.151, p=.107 for LM, and ρ=.231, p=.013 for GRF).

In order to further assess the predictive power of the two risk-attitude measures, we also ran several multiple ordinary least squares regression analyses trying to predict trusting behaviour by the lottery measure and the questionnaire measure (see Table 2).[16] Controlling for gender, age and senders' first-order beliefs regarding responders' trustworthiness (in % of transferred amounts), and step-wise entering the lottery- and the questionnaire measure, respectively, actual trusting behaviour in fact turns out to be associated with both measures. However, the contribution of the lottery-choice measure is significant only at a 10% level and its inclusion into the regression adds only little to the explanatory power of the model. In particular, compared to model (2), which controls only for gender, age and first-order (FO) beliefs, the changes in the explained variance and tests of statistical significance are considerably smaller when adding the lottery-measure in model (3) than when adding the GRF in model (4) (R^2 = .292 and p = .072 for LM, R^2 = .312 and p = .012 for GRF). Taking also into account the results for model (5) in which both LM and GRF are added (R^2 = .325 and p = .079 for LM and p = .013 for GRF), this strongly suggests that the questionnaire measure has considerably more predictive power than the lottery measure. And, in fact, the results of the re-test study presented below further support this interpretation of the results.

A possible explanation for the fact that the data from the lottery-choice task do not correlate with the general risk factor, personality or only weakly with actual risk-taking behaviour, of course, may be found in the fact that the lottery was also incentivised and was run after the trust-game. Thus, subjects may have used the lottery to hedge their behaviour in the trust game in a way that confounds the results of the lottery choice task (instead of treating the lottery-choice as an isolated item). Yet, we find no evidence for such behaviour in the data: As reported above, lottery

[16] Due to missing values (gender and age) we had to omit one subject.

Transfers in *test*	(1)	(2)	(3)	(4)	(5)
LM			.242*		.231*
			(.133)		(.130)
GRF				.669**	.653**
				(.260)	(.258)
FO-Belief		.0838***	.0829***	.0811***	.0803***
		(.0129)	(.0128)	(.0127)	(.0126)
Female	-1.024*	-.818*	-.789	-.691	-.667
	(.574)	(.492)	(.487)	(.483)	(.478)
Age	-.0483	-.0809	-.0718	-.0756	-.0671
	(.0794)	(.0681)	(.0676)	(.0665)	(.0660)
Constant	6.305***	3.741**	2.504	3.632**	2.457
	(1.922)	(1.691)	(1.807)	(1.651)	(1.765)
N	115	115	115	115	115
Adjusted R^2	.0125	.277	.292	.312	.325

Table 2: Ordinary least squares regression analyses predicting trusting behavior (amount transferred in trust game) as a function of LM, GRF, senders' first order beliefs, gender and age; senders: $N = 115$). Table entries are the predictors' unstandardised B-coefficients from multiple regression analyses; numbers in parentheses indicate standard errors. *** $p < .01$, ** $p < .05$, * $p < .10$.

choices are uncorrelated with investment behaviour. Moreover, there is also no significant correlation between the subjects' beliefs about their expected outcome from the trust game and their lottery-choice behaviour. And, as we will report in more detail further below, there is also no significant correlation between the change of behaviour in the trust game and that in the lottery choice task over time. As also Houser et al. (2010) find no order effects in a study on the determinants of trust, which combined the Berg et al. trust game with the Holt and Laury lottery-choice task in varying order, we are confident that order effects did not affect the outcomes in the present study.

3.2 The Re-Test

In order to investigate the test-re-test stability of both measures within a time period of one year, the re-test study was conducted under almost identical conditions (cf. Section 2.2); the results are reported below. In order to ensure that participants could not remember their prior responses, we asked them at the very end of the re-test study whether they had taken part in a similar study before. One participant

could remember taking part in the first study and was therefore excluded from the below analyses leaving us with 22 trusters and 21 responders. Moreover, in view of the results it is noteworthy that the 44 participants who volunteered for phase 3 of the study did not, on either risk measure administered at phase 2, differ from those who did not (mean scores of 4.33 and 4.47 for risky lottery decisions, and .01 and -.02 for factor scores on the risk questionnaire, for phase 3 drop-outs and volunteers, respectively; Mann-Whitney Test, all $p > .10$) so that attrition should not have biased the reported results in favor of either measure of individual risk attitudes.[17]

Experimental Results - Re-Test Only

All in all, the experimental results of the re-test study are very similar to those of the first study. The mean transfer was 4.409 Taler (*SD* 3.647) with transfers again ranging from the full endowment (10 Taler) to nothing (0 Taler). Moreover, as before, only the general risk factor is strongly correlated with trusting behaviour in the trust game (ρ=.494, p=.019) while behaviour in the lottery-choice task is not (ρ=.238, p=.286).

As above, we conducted several ordinary least squares regression analyses in order to examine the relation among the two measures of individual risk attitudes and observed trusting behaviour (see Table 3). This time, again holding gender, age and senders' first-order beliefs constant, we only find the questionnaire measure to have some predictive power regarding actual trusting behavior in the re-test. Moreover, adding the lottery measure of risk attitudes does not significantly affect the predictive power of the model; R^2 changes from .235 in model (2) to .242 in model (3). By contrast, adding the general risk factor significantly contributes to the predictive power of the model; R^2 changes from .235 in Model (2) to .496 in Model (4). Moreover, the complete model (5) again confirms this finding: only the general risk factor has significant predictive power regarding subjects' trusting behaviour. Thus, again the data strongly suggest that trusting behaviour (i.e. transfers in the trust game) is correlated with individual risk attitudes provided if these are measured by the type of questionnaire advocated by Dohmen et al. (forthcoming).

Test-Re-Test Comparison

The main aspect of running the re-test study, however, was to scrutinise the over-time stability of the different measures elicited in the course of our experiment, i.e. the risk-attitude measures and the amount transferred in the trust game.

[17] We also investigated demographic variables and trust decisions, but no differences between the groups were close to conventional levels of statistical significance; $p > .10$ for all items.

Transfers in *re-test*	(1)	(2)	(3)	(4)	(5)
LM			.324		-.449
			(.304)		(.343)
GRF				1.718***	2.422***
				(.549)	(.760)
FO-Belief		.109***	.109***	.0915***	.0840**
		(.0363)	(.0362)	(.0300)	(.0299)
Female	.255	-1.924	-2.050	-.398	.401
	(1.731)	(1.616)	(1.614)	(1.400)	(1.500)
Age	.0576	.0295	.0313	-.0407	-.0719
	(.185)	(.154)	(.154)	(.127)	(.127)
Constant	3.087	1.356	-.0566	2.899	5.485
	(4.376)	(3.692)	(3.908)	(3.038)	(3.571)
N	21	21	21	21	21
Adjusted R^2	-.101	.235	.242	.496	.517

Table 3: Ordinary least squares regression analyses predicting trusting behavior in the re-test (amount transferred in trust game) as a function of LM, GRF, senders' first order beliefs, gender and age; senders: $N = 21$). Table entries are the predictors' un-standardised B-coefficients from multiple regression analyses; numbers in brackets indicate standard errors. *** $p < .01$, ** $p < .05$, * $p < .10$.

Regarding the stability of individual risk attitudes over time, we first analysed the correlations between the test and the re-test study responses. The general risk factor (as well as all single items of the underlying questionnaire) shows a very high over-time stability ($\rho = .780$, $p < .001$), while the lottery-choice measure shows only a trend towards stability ($\rho = .258$, $p = .095$; the corresponding Pearson correlation coefficient is not significant: $r = .205, p = .188$); see Table 4 for details. Being surprised by the weak stability of choices in the lottery-choice task, we also examined the test-re-test stability of the lottery-choice measure by categorising participants into risk-averse (cross-over point after more than four safe decisions), risk-neutral (cross-over point after exactly four safe decisions), and risk-taking (cross-over point before more than four safe decisions). This could be done for 33 participants (those who had cross-over points in both studies). Yet, also the category shows no stability over time (χ^2 (4) = 5.38, $p > .10$). Thus, the test-re-test comparison casts some further doubt on the reliability of the lottery-choice task as a measure of individual risk-attitudes in connection with other laboratory experiments.

Risk attitudes	Spearman's ρ
LM	.258*
GRF	.780***
General	.768***
Driving	.573***
Financial matters	.550***
Sport and leisure	.670***
Career	.623***
Health behaviour	.632***
Trusting strangers	.641***

Table 4: Test-re-test stability of risk measures (all subjects; $N = 43$). *** $p < .01$, ** $p < .05$, * $p < .10$.

For comparison purposes and to examine what type of test-re-test correlation that could reasonably be expected for a behavioural decision-making measure, we also computed the test-re-test correlation of the amount transferred in the trust game. This correlation was almost as high as that of the general risk factor ($\rho = .692$, $p < .001$) suggesting that trusting behaviour indeed has a component which is a stable individual characteristic as hypothesised by Glaeser et al. (2000, p. 827).

As a final step of our analysis regarding the reliability and over time stability of the two different measures of individual risk attitudes, we investigated the extend to which risk attitudes as measured in the *test* study can predict trusting behaviour in the *re-test* study (see Table 5); note that trusting behaviour in the re-test is not confounded as it is not preceded by any other decision or questionnaire item. Consistent with our previous procedures, we used ordinary least squares regression analyses to try and predict investment decisions in the re-test by risk attitudes measures as gathered in the test study conducted one year earlier; again we controlled for gender, age and sender's first-order beliefs in the re-test study. As it turns out, the questionnaire measure indeed has significant predictive power regarding actual trusting behavior in the re-test; adding it increases R^2 from .235 in model (2) to .446 in model (4). Not surprisingly, though, the lottery measure shows no such effect. Moreover, also extending model (4) by adding the lottery measure in model (5) shows no change in R^2.[18]

[18]Single correlations of risk attitude measures from the test study and trust transfers in the re-test yield no significant correlation coefficients.

Transfers in *re-test*	(1)	(2)	(3)	(4)	(5)
LM (*test*)			.0856		-.00283
			(.395)		(.339)
GRF (*test*)				1.374**	1.375**
				(.503)	(.522)
FO-Belief (*re-test*)		.109***	.106**	.123***	.123***
		(.0363)	(.0393)	(.0313)	(.0341)
Female	.255	-1.924	-1.809	-1.910	-1.913
	(1.731)	(1.616)	(1.746)	(1.375)	(1.491)
Age	.0576	.0295	.0224	.0204	.0206
	(.185)	(.154)	(.162)	(.131)	(.138)
Constant	3.087	1.356	1.216	1.110	1.114
	(4.376)	(3.692)	(3.854)	(3.144)	(3.292)
N	21	21	21	21	21
Adjusted R^2	-.101	.235	.190	.446	.409

Table 5: Ordinary least squares regression analyses predicting trusting behavior in the re-test as a function of LM, GRF from the *test* study, and senders' first order beliefs (from the re-test study) as well as gender and age ($N = 21$). Table entries are the predictors' un-standardised B-coefficients from multiple regression analyses; numbers in brackets indicate standard errors. *** $p < .01$, ** $p < .05$, * $p < .10$.

Testing for Confounding Order Effects.

Finally, we used the results of the re-test study for a further test of the possibility that the results of the lottery-choice task are confounded due to order effects, i.e. a conditioning of lottery choices on investment behaviour in the trust game that might have blurred the underlying correlation between risk and trust. In particular, we tested whether the direction of over-time changes in investment behaviour (trust game) are correlated with direction of changes in lottery-choice behaviour. If subjects had conditioned their behaviour in the lottery choice task on their earlier investment behaviour, it seemed reasonable to expect the direction of the deviations to be correlated. However, we do not find any evidence for such a correlation: 11 trusters change their investments over time (3 invest more, 8 less). Of those who invest more in the re-test-study, one makes a safer choice in the lottery task and two make a more risky one, while the 8 subjects who invest less in the re-test-study split equally into safer and more risky choices (4 each). Accordingly, we find that the direction of change in the trust game is not correlated with the direction of change in the lottery-choice

task (Fisher's exact test $p = .576$).[19] Thus, also the combination of test and re-test data gives no indication of order effects which might have confounded the results of the lottery-choice task.

4 Concluding Remarks

In this paper, we have compared the empirical power of two different measures of individual risk attitudes as instruments to analyse behaviour in economic lab experiments (here the Berg et al., 1995, trust-game): the Holt and Laury (2002) lottery choice task and a questionnaire similar to the one advocated by Dohmen et al. (forthcoming). The data presented above show that the general risk factor derived from the questionnaire about individual risk attitudes has good construct validity (being correlated with an external predictor of risk-taking behaviour, namely personality, almost exactly as expected), reasonable predictive power (regarding transfer behaviour in the trust game), and a very good test-re-test stability over time (one year). By contrast, the Holt and Laury (2002) lottery-choice measure shows no construct validity (again when related to personality), almost no predictive power, and most importantly no robust test-re-test stability.[20] Furthermore, we also find no evidence supporting the convergent validity between general risk factor and the Holt and Laury (2002) lottery-choice task, i.e. both are uncorrelated in our study. Taken together, the results thus suggest that behaviour in the lottery-choice task is not a very reliable measure of stable individual differences in risk-attitudes when measured in connection with other (primary) behavioural tasks in the lab.

Moreover, the lack of reliability of the lottery-choice measure can neither be clearly attributed to the method of measurement (i.e. relying on a single behavioural item), as our other one-shot behavioural measure – transferred amount as truster – shows a very high test-re-test stability. Nor can it be attributed to the construct being

[19] The same remains to hold if also trustees are drawn into the picture: 10 of them change their back-transfer behaviour over time. Combining those who invest more (less) with those who back-transfer more (less), we again find that the direction of change in one task is not correlated with that in the other (Fisher's exact test $p = .410$).

[20] Note that we are not the first to study over-time stability of individual risk attitudes. In a different context, using a different subject pool (a representative sample of the Danish population) and a different set-up (for different scalings of the lottery with no further primary experimental task), for example, Harrison, Johnson, McInnes and Rutström (2005b) find that CRRA coefficients when inferred from choices in the Holt and Laury lottery-choice measure exhibit some stability over time; see also Andersen, Harrison, Lau and Rutström (2008b). We are not aware of any other study analysing and comparing the test-retest stability of (different) measures of individual risk attitudes in connection with other primary tasks in economic lab experiments, though.

measured (i.e. risk attitudes), as the multi-item questionnaire measure shows a very high test-re-test stability. Thus, the problem appears specific to the behavioural measurement of individual risk attitudes by means of lottery choice tasks such as the one proposed by Holt and Laury (2002).[21] From an applied point of view, our results therefore recommend using the type of questionnaire measure advocated by Dohmen et al. (forthcoming) rather than — or at least complementary to — common lottery-choice measures when studying the connection between individual risk-attitudes and behaviour in laboratory experiments.

Apart from the aspects about the adequate measurement of individual risk attitudes mentioned above, our results also contribute to the current discussion on the relation between risk and trust. Challenging the claim that there is an intrinsic relation between these constructs (suggested e.g. by Snjiders and Keren, 1998, or Ben-Ner and Putterman, 2001), Eckel and Wilson (2004) reported that their risk instruments (including the Holt and Laury measure) could not predict the decision to trust in a binary trust game (for similar results, see Bohnet and Zeckhauser, 2004; Houser et al., 2010). By contrast, our results support the intuitive contention that trusting behaviour, as revealed in the Berg et al. trust / investment game, does entail an element of risk. In how far these results translate to more general instances of trust, of course, remains an open question.[22] The results of the present study, however, strongly suggest that future research into the nature of trust and its connection to risk should include a questionnaire measure of individual risk-attitudes of the type advocated by Dohmen et al. (forthcoming) instead of relying on lottery-choice tasks alone.

Finally, we want to reemphasise that the high test-re-test stability of the observed investment behaviour in the trust game strongly suggests that trusting behaviour indeed has a component which is a stable individual characteristic as hypothesised by Glaeser et al. (2000, p. 827).[23] Thus, although the exact determinants of trust remain an open question, it seems that trust is at least a comparably stable phenomenon — a fact that, despite all remaining difficulties, should facility future enquiries into the nature of trust.

[21] The more general claim going beyond the Holt and Laury task, of course, calls for further research to be substantiated.

[22] See Fehr (2009) for an interesting recent discussion about the determinants of trust.

[23] As mentioned earlier, the respective hypothesis is also supported through a study by Cesarini et al. (2008) who identify a correlation between trusting behaviour and the subjects' genetic code.

References

Andersen, S., Harrison, G.W., Lau, M.I., Rutström, E.E. (2008a). Eliciting Risk and Time Preferences. *Econometrica*, 76, pp. 583-618.

Andersen, S., Harrison, G.W., Lau, M.I., Rutström, E.E. (2008b). Lost in State Space: Are Preferences Stable? *International Economic Review*, 46, pp. 1091-1112.

Arrow, K. (1965). *Aspects of the Theory of Risk Bearing*. Yrjö Jahnssonin Saeaetioe, Helsinki.

Bartling, B., Fehr, E., Marechal, A., Schunk, D. (2009). Egalitarianism and Competitiveness. *American Economic Review*, 99, pp. 93-98.

Ben-Ner, A., Halldorsson, F. (2010). Trusting and trustworthiness: What are They, How to Measure Them, and What Affects Them. *Journal of Economic Psychology*, 81, pp. 64-79.

Ben-Ner, A., Putterman, L. (2001). Trusting and Trustworthiness. *Boston University Law Review*, 81, pp. 523-551.

Berg, J., Dickhaut, J., McCabe, K. (1995). Trust, Reciprocity, and Social History. *Games and Economic Behavior*, 10, pp. 122-142.

Bohnet, I., Zeckhauser, R. (2004). Trust, Risk and Betrayal, *Journal of Economic Behavior & Organization*, 55, pp. 467-484.

Cesarini, D., Dawes, C., Fowler, J., Johannesson, M., Lichtenstein, P., Wallace, B. (2008). Heritability of Cooperative Behavior in the Trust Game. *Proceedings of the National Academy of Sciences*, 105, pp. 3721-3726.

Costa, P.T., McCrae, R.R. (1992). *NEO PI-R Professional Manual*. Odessa, FL: Psychological Assessment Resources.

Digman, J.M. (1990). Personality Structure: Emergence of the Five-Factor Model. *Annual Review of Psychology*, 41, pp. 417-440.

Dohmen, T., Falk, A. (forthcoming). Performance Pay and Multi-Dimensional Sorting – Productivity, Preferences and Gender. *American Economic Review*.

Dohmen, T., Falk, A., Huffman, D., Sunde, U. (2008). The Intergenerational Transmission of Risk and Trust Attitudes. *CEPR Discussion Paper*, No. 6844 (2008).

Dohmen, T., Falk, A., Huffman, D., Sunde, U. (2010). Are Risk Aversion and Impatience Related to Cognitive Ability? *American Economic Review*, Vol. 100, pp. 1238-1260.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J, Wagner, G. G. (forthcoming). Individual Risk Attitudes: Measurement, Determinants and Behavioral Consequences. *Journal of the European Economic Association*.

Eckel C.C., Wilson R.K. (2004). Is Trust a Risky Decision? *Journal of Economic Behavior & Organization*, 55, pp. 447-465.

Falk, A., Heckman, J. (2009). Lab Experiments are a Major Source of Knowledge in the Social Sciences. *Science*, 326, pp. 535-538.

Fehr, E. (2009). On the Economics and Biology of Trust. *Journal of the European Economic Association*, 7, pp. 235-266.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10, pp. 171-178.

Glaser, E.L., Laibson, D.I., Scheinkman, J.A., Souter, C.L. (2000). Measuring Trust. *Quarterly Journal of Economics*, 115, pp. 811-846.

Goeree, J.K., Holt, C.A., Palfrey, T.R. (2002). Quantal Response Equilibrium and Overbidding in Private-Value Auctions. *Journal of Economic Theory*, 104, pp. 247-272.

Goeree, J.K., Holt, C.A., Palfrey, T.R. (2003). Risk Averse Behavior in Generalized Matching Pennies Games. *Games and Economic Behavior*, 45, pp. 97-113.

Greiner, B. (2004). An Online Recruitment System for Economic Experiments. In: Kurt Kremer, Volker Macho (Eds.): *Forschung und wissenschaftliches Rechnen 2003*. GWDG Bericht 63, Gttingen: Ges. fr Wiss. Datenverarbeitung, 79-93, 2004.

Guiso, L., Paiella, M. (2005). The Role of Risk Aversion in Predicting Individual Behavior. *Bank of Italy Economic Working Paper*, No. 546.

Harrison, G.W. (2006). Hypothetical Bias Over Uncertain Outcomes. Pp. 41-69 in J. List (ed.), *Using Experimental Methods in Environmental and Resource Economics*, Edward Elgar, Northampton, MA.

Harrison, G.W., Johnson, E., McInnes, M.M., Rutström, E.E. (2005a). Risk Aversion and Incentive Effects: Comment. *American Economic Review*, 95, pp. 897-901.

Harrison, G.W., Johnson, E., McInnes, M.M., Rutström, E.E. (2005b). Temporal Stability of Estimates of Risk Aversion. *Applied Financial Economics Letters*, 1, pp. 31-35.

Harrison, G.W., Rutström, E.E. (2008). Risk Aversion in the Laboratory. In J.C. Cox and G.W. Harrison (eds.), *Risk Aversion in Experiments (Research in Experimental Economics 12)*, Bingley, UK: Emerald, pp. 41-196.

Holt, C.A., Laury, S.K. (2002). Risk Aversion and Incentive Effects. *American Economic Review*, 95, pp. 902-904.

Holt, C.A., Laury, S.K. (2005). Risk Aversion and Incentive Effects: New Data without Order Effects. *American Economic Review*, 92, pp. 1644-1655.

Houser, D., Schunk, D., Winter, J. (2010). Distinguishing Trust from Risk - An Anatomy of the Investment Game. *Journal of Economic Behavior and Organization*, 74, pp. 72-81.

Konstabel, K., Lönnqvist, J.-E., Walkowitz, G., Konstabel, D., Verkasalo, M. (forthcoming). The "Short 5": Measuring Personality with Comprehensive Single Items. European Journal of Personality.

Kosfeld, M., Heinrichs, M., Zak, P., Fischbacher, U., Fehr, E. (2005). Oxytocin Increases Trust in Humans. *Nature*, 435, pp. 673-676.

Lauriola, M., Levin, I.P. (2001). Personality Traits and Risky Decision-Making in a Controlled Experimental Task: An Exploratory study. *Personality and Individual Differences*, 31, pp. 215-236.

Nicholson, N., Soane, E., Fenton-O'Creevy, M., Willman, P. (2005). Personality and Domain-Specific Risk-Taking. *Journal of Risk Research*, 8, pp. 157-176.

Pratt, J. (1964). Risk Aversion in the Small and in the Large. *Econometrica*, 32, pp. 122-136.

Rabin, M. (2000). Risk Aversion and Expected Utility Theory. *Econometrica*, 68, pp. 1281-1292.

Selten, R. (1967). Die Strategiemethode zur Erforschung des eingeschränkt rationalen Verhaltens im Rahmen eines Oligopolexperimentes. Pp. 136-168 in H. Sauermann (Ed.), *Beiträge zur experimentellen Wirtschaftsforschung*, J.C.B. Mohr (Paul Siebeck), Tübingen.

Snijders, C., Keren, G. (1998). Determinants of Trust. In: Budescu, D., Erev, I., Zwick, R. (Eds.), Games and Human Behavior: Essays in Honor of Amnon Rappoport. Lawrence Erlbaum, Mahwah, NJ., pp. 355-385.