Maastricht University

Measuring time and risk preferences: Reliability, stability, domain specificity

Citation for published version (APA):

Wölbert, E. M., & Riedl, A. M. (2013). *Measuring time and risk preferences: Reliability, stability, domain specificity*. Maastricht University, Graduate School of Business and Economics. GSBE Research Memoranda No. 041 https://doi.org/10.26481/umagsb.2013041

Document status and date:
Published: 01/01/2013

DOI:
10.26481/umagsb.2013041

Document Version:
Publisher's PDF, also known as Version of record

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy
If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Download date: 24 Apr. 2024

Maastricht University

Eva Wölbert, Arno Riedl

Measuring Time and Risk Preferences: Reliability, Stability, Domain Specificity

RM/13/041

GSBE

Maastricht University School of Business and Economics
Graduate School of Business and Economics

P.O Box 616
NL- 6200 MD Maastricht
The Netherlands

Measuring Time and Risk Preferences: Reliability, Stability, Domain Specificity*

Eva Wölbert Arno Riedl

July 2013

Abstract

To accurately predict behavior economists need reliable measures of individual time preferences and attitudes toward risk and typically need to assume stability of these characteristics over time and across decision domains. We test the reliability of two choice tasks for eliciting discount rates, risk aversion, and probability weighting and assess the stability of these characteristics over time and across situations. We find high reliability and that individual characteristics are remarkably stable over time. The estimated parameters correlate well with self-reported decisions in financial domains, but are largely uncorrelated with decisions in other important life domains involving intertemporal trade-offs and risk.

Keywords: time discounting, risk aversion, probability weighting, reliability, stability, domain specificity, experiment

JEL Classification: C18, C91, D03, D80, D90

---

*Eva Wölbert: Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience,) & Department of Economics (AE1), School of Business and Economics, Maastricht University, P.O.Box 616, 6200 MD Maastricht, the Netherlands, evawoelbert@outlook.com; Arno Riedl (corresponding author): CESifo, IZA, Netspar, Department of Economics (AE1), School of Business and Economics, Maastricht University, P.O.Box 616, 6200 MD Maastricht, the Netherlands, a.riedl@maastrichtuniversity.nl. We are grateful for valuable comments of participants in seminars and conferences at Barcelona, Bonn, Evanston, Helsinki, and Maastricht.

1 Introduction

Many important economic and societal decisions involve outcomes that are delayed or uncertain. Consequently, individual time preferences and attitudes toward risk are of fundamental importance for economic models. Beyond that they also play an important role in research on human decision making in other social sciences, like psychology and neuroscience (see,e.g., Kirby et al. 1999, Trepel et al. 2005, Kable and Glimcher 2007, Fox and Tannenbaum 2011).

In order to make accurate predictions, economic research and applications need reliable instruments for measuring individual preferences. Moreover, typically it needs to be assumed that individual preferences are stable over time and across various decision contexts. Despite their obvious importance to the validity of economic research and the effectiveness of policy implementations, these assumptions have not been sufficiently addressed in the literature (Borghans et al. 2008). In this paper we provide an assessment of the reliability of commonly used experimental time and risk preference elicitation tasks. In addition, we explore the temporal stability of elicited discount rates, risk aversion, and probability weighting over several weeks and ask whether they generalize to important domains of life outside the laboratory.

In economics, using experiments and surveys, time and risk preferences are usually quantified based on real or hypothetical behavior in decision making tasks. Subjects are asked to make several decisions involving uncertain and/or delayed (monetary) rewards. It is typically assumed that these tasks are reliable measures, i.e. that their results are reproducible and measurement error is small. However, the actual reliability of these tasks is largely unknown. Knowledge about reliability is important, however. If reliability is low, elicited preferences in one task cannot be expected to be temporally stable or correlate well with behavior in other settings, simply due to high measurement error. Hence, only with reliable instruments accurate assessments about temporal and across-domain stability of preferences can be made.

When using elicited time preferences and attitudes toward risk to explain field behavior and derive potential policy implications, one needs to assume that a person's decisions in such tasks reflect underlying dispositions that are stable, both over time and across

different decision situations. That is, a person's preferences at one point in time have to be assumed to be predictive of her preferences at another point in time, and preferences in one life domain have to be predictive of preferences in a different domain. However, behavior in choice tasks, even if suitably incentivized, need not necessarily reflect longer-lasting, all-encompassing attitudes of the individual towards every type of risk or delay. Instead, it might be highly dependent on the subject's state or the decision context (Fox and Tannenbaum 2011). Whether and how time and risk preferences are stable over time and across different decision domains is an important, ultimately empirical question that economists have only recently begun to address.

Investigating the temporal stability of time preferences, Kirby (2009) and Meier and Sprenger (2010b) both found that discount rates elicited in a monetary intertemporal choice task were relatively stable over a period of one year. Although many different elicitation methods for time preferences have been explored (see Frederick et al. 2002, for review), there is little evidence on the stability of individual time preferences across different choice contexts. The few studies that have investigated how elicited discount rates relate to field behavior find mixed results. Whereas Ashraf et al. (2006) and Meier and Sprenger (2010a) reported that present bias in an intertemporal choice task is associated with credit card debt and preference for a commitment savings product, Chabris et al. (2008) as well as Borghans and Golsteyn (2006) found that experimentally elicited discount rates correlate only very weakly with health-related behavior such as exercising and smoking. Interestingly, Reuben et al. (2009) observed that discount rates over monetary rewards predicted procrastination on an unpleasant task.

Evidence on the stability of risk attitudes over time and across domains is equally scarce, despite the many empirical studies on risk taking (see Fehr-Duda and Epper 2012, for a recent review). Andersen et al. (2008b) found that individual risk aversion measured with a typical multiple price list (Holt and Laury 2002) was only moderately stable over time and also Zeisberger et al. (2012) observe considerable instability of risk aversion and probability weighting over a period of one month. First field studies addressing the stability of risk aversion across different contexts suggest that many individuals do not exhibit comparable degrees of risk aversion in different life domains, such as health, disability or car insurance (Barseghyan and Prince 2011, Einav et al. 2012). Pennings and

Smidts (2000) found that responses in an experimental lottery choice task were correlated with managerial decisions involving risk, providing some initial support for the construct validity of such tasks.

In this paper we address the questions of reliability of measurement of time preferences and attitudes toward risk and their stability over time and decision domains in one encompassing study. We use intertemporal choice tasks and lottery choice tasks that have been frequently used in economic research (e.g. Andersen et al. 2008b, Chabris et al. 2008, Bruhin et al. 2010, Dohmen et al. 2010, Ifcher and Zarghamee 2011). In each task, subjects decide repeatedly between different amounts of money that are associated with varying delays in intertemporal choice, or with varying winning probabilities in lottery choice. Subjects' choices are used to derive parameter estimates describing the individual's discount rate, risk aversion and degree of nonlinear probability weighting. We assess reliability of preference elicitation techniques by correlating the individual results from mutually exclusive choice subsets. Establishing good reliability provides the foundation for estimating the stability of preferences over time and addressing questions of ecological validity. In order to estimate test-retest stability, we repeated the experiment on a subset of subjects after several weeks and correlate parameter estimates derived at the two different test dates. Lastly, to contribute towards establishing validity of these measurement techniques across decision domains, we report correlations of estimates of time preferences and risk attitudes with conceptually related self-report scales that include questions on self-control, future orientation, savings, and risk attitudes.

Our results can be summarized as follows. First, we find that the used choice tasks produce behavioral measures that are reliable and provide reproducible estimates. Second, all estimated measures show high intertemporal stability over a period of 5–10 weeks. Hence, we can conclude that the elicited time preferences, risk aversion, and probability weighting are stable at least over the tested period of time. Third, we find that the elicited preferences are predictive for (self-reported) choices in life domains related to financial decisions, like investments and savings. However, they are largely unrelated to decisions in other important life domains involving delayed and risky consequences, like health and leisure. Overall, our results indicate that the explored elicitation tasks can be useful for

predicting behavior regarding intertemporal and risky financial decisions but are likely uninformative for behavior outside the narrow financial realm.

The rest of the paper is organized as follows. The next section describes our design and procedures. Section 3 reports the results on reliability, stability, and domain specificity first for the measure of time preferences followed by the measures of risk aversion and probability weighting. Section 4 offers a discussion of our results and concludes. The appendices offer additional descriptive statistics and tests (Appendix A), details on questionnaires (Appendix B), and contain the experiment instructions (Appendix C).

2 Design and Procedures

We conducted experiments where we measured time preferences and risk attitudes of participants. All subjects completed a series of choices between smaller-sooner and larger-later amounts of money as well as a series of choices between guaranteed and probabilistic payments, all of which were fully incentivized. After the decision tasks, subjects completed a set of questionnaires measuring general and specific dimensions of impulsiveness and self-assessed risk attitudes in various domains of life. In order to assess the intertemporal stability of estimated time and risk preferences, a subset of subjects was asked to return to the lab for the same experiment 5–10 weeks after the first data acquisition. All tasks and questionnaires are described more specifically below. At the end of this section details on the experimental procedure are given.

Measuring Time Preference

A set of intertemporal decision problems over monetary rewards was used to assess time preference. Subjects faced 27 choices between a smaller-sooner and a larger-later amount of money. For example, subjects were asked "Would you prefer to receive € 25 today or € 30 in 80 days?". Smaller-sooner amounts ranged from € 11 to € 54, and the larger-later amounts ranged from € 25 to € 60. Delays ranged from 7 days to 200 days. All choice problems were presented individually and in no obvious logical order, following a sequence suggested by Kirby et al. (1999).

The set of intertemporal choices that we used consisted of 3 subsets of 9 choices each. Set 1 contained items from the intertemporal choice questions introduced by Kirby et al. (1999). These questions ask the subject to choose between an amount to be paid out immediately after the experiment session and a larger amount available after a delay of d days. In order to assess whether subjects exhibit time inconsistent behavior we created two more sets of 9 items each. Specifically, we modified the items of set 1 by adding a fixed delay of one day (set 2) or 14 days (set 3) to both the sooner and the later amount. Consequently, the smaller-sooner amounts in set 2 were paid out after one day and the larger-later amounts after $d + 1$ days. In set 3 all smaller-sooner amounts were paid out after 14 days and all larger-later amounts after $d + 14$ days.

We used these choice sets to estimate an exponential discount parameter ρ for each subject and each choice set according to

$$V = Ae^{-\rho D}, \tag{1}$$

where V denotes the current subjective value of the delayed amount, A denotes the monetary amount, and D the delay in days.[1]

Table 1 shows the values and delays used for all 27 intertemporal choices subjects had to make as well as the order of presentation. In the table the intertemporal choice questions are ordered into the three sets described above. Within each set the questions are sorted according to the discount rate that would make an individual indifferent between the two options.[2] As can be seen from the table, within each set, choices range from requiring extreme patience when preferring the later option (e.g. choosing to wait 186 days to receive € 35 instead of € 34 immediately) to extreme impatience when preferring the sooner option (e.g. going for immediate € 19 instead of waiting 7 days to receive € 30). Subjects with time preferences that lie in between these extremes will prefer the

[1]Several specifications for individual discount rates are feasible. Importantly, our results do not depend on the specification of the discount rate. The ρ values are not estimated by fitting an exponential function to the data, but are determined by subjects' switching points between smaller-sooner and large-later amounts. Moreover, as will be shown later, our subjects did not exhibit a present bias, which makes an exponential specification appropriate.

[2]We used a slightly different version of set 3 for the first 96 subjects in the first experiment. The set displayed in the table turned out to cover the range of subjects' choices better.

Table 1: Intertemporal Choice Tasks

Set	Order of presentation	Smaller-sooner amount	Larger-later amount	Days to smaller amount	Days to later amount	Indifference ρ
	13	€34	€35	0	186	0.00016
	20	€28	€30	0	179	0.00039
	26	€22	€25	0	136	0.00094
	22	€25	€30	0	80	0.00228
1	3	€19	€25	0	53	0.00518
	18	€24	€35	0	29	0.01301
	5	€14	€25	0	19	0.03052
	7	€15	€35	0	13	0.06518
	11	€11	€30	0	7	0.14333
	1	€34	€35	1	187	0.00016
	6	€28	€30	1	180	0.00039
	24	€22	€25	1	137	0.00094
	16	€25	€30	1	81	0.00228
2	10	€19	€25	1	54	0.00518
	21	€24	€35	1	30	0.01301
	14	€14	€25	1	20	0.03052
	8	€15	€35	1	14	0.06518
	27	€11	€30	1	8	0.14333
	9	€34	€35	14	200	0.00016
	17	€28	€30	14	193	0.00039
	12	€22	€25	14	150	0.00094
	15	€25	€30	14	94	0.00228
3	2	€19	€25	14	67	0.00518
	25	€24	€35	14	43	0.01301
	23	€14	€25	14	33	0.03052
	19	€15	€35	14	27	0.06518
	4	€11	€30	14	21	0.14333

Note. For presentation in this table, questions are grouped into the three sets. Within each set, they are ordered according to the exponential discount rate that would make a decision maker indifferent between the two options. Column 2 indicates the position in the order of presentation to the subjects.

smaller-sooner option up to a given point—their switch point—and prefer the larger-later option in all remaining tasks. For them the discount rate lies in the interval between the two indifference parameters of the tasks where the switch occurs. We use the geometric mean of this interval as an estimate of the discount rate in that case (cf. Kirby et al. 1999). When a subject displayed more than one switch point within a set, the discount rate was assigned such that it minimized the number of choices that would have been a mistake according to this value.[3] This procedure was applied separately for each of the three intertemporal choice sets, yielding three estimates for the discount rate per subject, which we denote as ρ_1, ρ_2 and ρ_3.

Measuring Risk Preference

When assessing attitudes toward risk we estimate the curvature of each subject's value function as well as the degree of nonlinear probability weighting (Kahneman and Tversky 1979). For that we elicit certainty equivalents for 20 lotteries. Each lottery was presented to the subjects on a decision screen that showed the specific lottery as well as a list of 20 sure amounts. Lotteries were described in written form and probabilities were also visualized as a pie chart in order to facilitate comprehension. The sure amounts ranged from the higher to the lower outcome of the specific lottery and were equally spaced. For each of these sure amounts, subjects had to indicate whether they preferred to receive the lottery or this sure amount. In each decision screen subjects were allowed to switch only once between preferring the sure amount and preferring the lottery, i.e. monotonicity was enforced in this task.[4]

For each lottery the certainty equivalent was determined as the arithmetic mean of the smallest sure amount the subject preferred over the lottery and the next smaller amount on the list. All lotteries consisted of two non-negative outcomes in the range of € 0 to

[3]It turned out that within one intertemporal choice set most subjects answer consistently, i.e. they display a single switch point (85.8% of all subjects were consistent in all three sets and 93.9% of all sets were answered consistently). Only two subjects displayed multiple switch points in all three choice sets. We excluded them from the analysis of the intertemporal choice task.

[4]See Epper et al. (2009, 2011) and Bruhin et al. (2010) for successful implementations of this procedure. The experiment instructions shows a screenshot of a decision screen of one lottery (see Appendix C, p. 38).

Table 2: Lotteries in the Lottery Choice Task

Order	p_{x1}	$x1$	$x2$	Order	p_{x1}	$x1$	$x2$
9	0.1	€12	€6	7	0.25	€30	€12
13	0.5	€12	€6	19	0.5	€30	€12
6	0.9	€12	€6	17	0.75	€30	€12
11	0.05	€24	€6	2	0.95	€30	€12
18	0.25	€24	€6	20	0.05	€90	€30
3	0.5	€24	€6	1	0.5	€6	€0
12	0.75	€24	€6	8	0.5	€12	€0
14	0.95	€24	€6	5	0.05	€24	€0
4	0.05	€30	€12	10	0.95	€30	€0
16	0.1	€90	€0	15	0.25	€24	€0

Note. Outcomes $x1$ and $x2$ are stated in Euro, p_{x1} denotes the probability for outcome $x1$. Order denotes the position in the order of presentation to the subject.

€ 90, probabilities attached to these outcomes ranged from 0.05 to 0.95. The full set of lotteries as well as the order of presentation are shown in Table 2.[5]

Elicited certainty equivalents were used to jointly estimate a value function and a probability weighting function for each subject. When choosing the functional forms we were guided by requirements of parsimony and past evidence indicating good fit of data. Regarding the value function it has been shown that a simple power function

$$v(x) = x^{\alpha}, x \geq 0, \alpha > 0, \tag{2}$$

where x denotes the monetary value of the outcome, is the best compromise between these requirements (Stott 2006, Wakker 2008, Bruhin et al. 2010). It is also conveniently interpreted as $0 < \alpha < 1$ indicates risk aversion, $\alpha = 1$ indicates risk neutrality, and $\alpha > 1$ indicates risk seeking.

[5]Since we were mainly interested in individual variability in this task, we presented the lotteries in the same order to all participants. Since no lottery was resolved before the end of the experiment, no history effects are expected. Any potential effects of the order of presentation should affect all subjects equally.

To capture probability weighting we use a parametrization originally suggested by Prelec (1998) (see also Epper et al. 2009),

$$\omega(p) = \exp(-(-\ln p)^{\gamma}), 0 < \gamma < 1, \tag{3}$$

where p denotes the objective probability. This specification allows for subproportionality and is inverse S-shaped, properties in line with previous empirical evidence (Gonzalez and Wu 1999, Stott 2006, Bruhin et al. 2010, Epper et al. 2011). The degree of deviation from linearity is conveniently captured by the single parameter γ, where smaller values refer to larger deviations in the form of an inverted S-shape. That is, small probabilities are overweighted and large probabilities are underweighted.

For each subject, α and γ were estimated simultaneously by minimizing the squared distance between the estimated subjective value of the lottery and the measured certainty equivalent. To correct for heteroscedasticity, the outcomes of each lottery were normalized to uniform length.

Self-Reported Impulsiveness and General Risk Attitudes

In order to investigate how behavior in the monetary choice tasks in the laboratory relates to the conceptually close psychological measure of impulsiveness and to (perceived) risk preferences in various life domains subjects completed a series of questionnaires.

Impulsiveness is a psychological construct that is thought to be closely related to intertemporal choice. The inability to delay gratification is considered the core problem of impulsive behaviors, for example when the immediate pleasure of eating sweet food is preferred over the delayed health benefits associated with refraining from it (Ainslie 1975, Logue 1988). Likewise, in economics it is assumed that an intertemporal choice between an immediately available and a delayed reward requires self-control and that impatience is exacerbated by impulsiveness (Laibson 1997, Frederick et al. 2002, McClure et al. 2004). The evidence on the association of impulsiveness with intertemporal choice tests is mixed, however (Kirby et al. 1999, Reynolds et al. 2006, de Wit et al. 2007, Reimers et al. 2009).

For the purpose of this study, the Barratt Impulsiveness Scale (BIS-11) was employed (Patton et al. 1995). The BIS-11 is a well-validated and standardized self-report question-

naire for measuring impulsiveness.[6] There is ample evidence that high scores on the BIS-11 are associated with behaviors that reflect a lack of self-control and future-orientation, such as smoking, alcohol and illegal substance abuse, attention deficits, as well as a higher sensitivity to reward and punishment (for a review, see Stanford et al. 2009). Next to a sum score that reflects general impulsiveness, the BIS-11 allows computing a subscore for non-planning impulsiveness. This subscore is supposed to capture interindividual differences in self-control and future-orientation that are often assumed to contribute to individual differences in intertemporal choice behavior.

To assess self-reported willingness to take risks in different domains of life subjects completed a risk taking questionnaire from the German Socio-Economic Panel (SOEP). Specifically, subjects answered questions regarding their general risk attitudes ("Generally speaking, are you a person who is ready to take risks or are you trying to avoid risks?") as well as questions about risk taking in six specific domains: car driving, financial matters, leisure and sports, health, professional career, and confidence in strangers. Table 3 shows the precise wording of the questions. We used these questions because it has been shown that answers to them are significantly related to actual behavior in the field regarding investment in stocks, being self-employed, participating in sports, and smoking (Dohmen et al. 2011). Therefore, (non-)correlations between the laboratory measures and answers to these questions can be viewed es indicative for the ecological (in)validity of those measures.

General Procedure

The data acquisition took place at the Behavioral and Experimental Economics laboratory (BEElab) of the School of Business and Economics at Maastricht University. Subjects were recruited with ORSEE (Greiner 2004) from the BEElab subject pool which consists of students of all fields of study at Maastricht University, the majority being students of business or economics. In total, 144 students participated in the experiment, 50.7% of these were female, 75% were enrolled in the School of Business and Economics. The time interval between the first time a subject was tested, and the retest ranged from 5 to 10 weeks. We invited about two-thirds of the original subjects for the retest. Subjects did not

[6] Table B.1 in Appendix B shows the questionnaire.

Table 3: Questions for Self-Reported "Attitudes Towards Risk"

Generally speaking,
are you a person who is ready to take risks or are you trying to avoid risks?
One can behave differently in different circumstances.
In the following circumstances, how would you assess your readiness to take risks?
Driving a car?
Making a financial investment?
In leisure and when doing sports?
Regarding your professional career?
Regarding your health?
Regarding confidence in strangers?

Note. Subjects rated each item on a scale from 0 (unwilling to take risks) to 10 (fully prepared to take risks) (see Dohmen et al. 2011).

know that they were invited for the same experiment. In total 53 participants returned for the retest.[7]

All decisions and questionnaires were presented on a computer screen using the software package z-Tree (Fischbacher 2007). At the beginning of the experiment, subjects were informed that there would be several decision making tasks and that they will receive detailed instructions before each task. They were allowed to make their decisions at their own speed.

Subjects received a show-up fee of € 3 in cash. In addition, one of all decisions made during the experiment was selected at random, the subject's choice in this decision situation was implemented and the associated earnings were paid out. This procedure allowed us to provide salient incentives for several tasks while minimizing the problem that decisions might be influenced through house money effects or risk hedging.[8] In case

[7]In the results section, the retested group is compared to those subjects that did not return for the retest in order to assess whether the retest data is affected by selective dropout of subjects. There are no significant differences in the tested variables between retested and non-retested subjects.

[8]Hey and Lee (2005) have shown that subjects' behavior is in line with the assumption that each decision is considered as an independent decision under this incentive scheme.

a lottery was selected for payout, risk was resolved by a random draw of the computer. Subjects received detailed information on the payment procedure at the beginning of the experiment and were again reminded of the determination of earnings before each of the different decision making tasks.

All earnings — except delayed ones — were paid out in private immediately after the experiment in cash. Delayed payments were delivered by bank transfer in order to minimize transaction costs.[9] On average, the experiment lasted 1.5 hours. Total earnings ranged from € 3 to € 129.50, with an average of € 27.55.

3 Results

In the following, we first describe the data on time preferences, and second the data on risk preferences. The central questions addressed are: First, how reliably does each task measure individual preferences? Second, how stable are individual preferences over the two separate test dates? In addition, for time preferences we ask how well the elicited preferences correlate with psychological constructs of impulsiveness? For risk preferences we also explore the association between the laboratory measure and validated self-report questions in different decisions domains. In the beginning of each section, descriptive statistics are presented in order to compare results to the existing literature and to check for general effects of repeated testing and selective drop-out. If not otherwise indicated reported correlations are Spearman rank-order correlations and p-values refer to two-sided tests.

Time preferences

At both test dates subjects made decisions in three sets of intertemporal choice tasks. For each of these sets we calculated exponential discount rates, denoted ρ_1, ρ_2, and ρ_3,

[9]Money in a Dutch bank account is accessible at virtually no transaction costs, since payment by debit card is widely accepted and a large number of conveniently situated ATMs is available 24h and free of charge. We therefore are confident that the difference in transaction costs between immediate and delayed payments is negligible.

respectively, for each subject. Table 4 shows descriptive statistics of these discount rates for both test dates. The results are in keeping with discount rates reported in studies using a similar procedure or procedures that involve comparable reward sizes and delays (see Thaler 1981, Chabris et al. 2008, Kirby 2009, and Frederick et al. 2002 for a review).

Table 4: Descriptive Statistics of Time Preferences

	N		Median	Mean	95% CI of Mean
Date 1		ρ_1	0.00821	0.00685	[0.00538, 0.00872]
whole	142	ρ_2	0.00821	0.00622	[0.00499, 0.00776]
sample		ρ_3	0.00737	0.00694	[0.00558, 0.00862]
Date 1		ρ_1	0.00821	0.00646	[0.00430, 0.00970]
retest	53	ρ_2	0.00821	0.00540	[0.00377, 0.00773]
subsample		ρ_3	0.00737	0.00602	[0.00415, 0.00875]
		ρ_1	0.00821	0.00589	[0.00382, 0.00907]
Date 2	53	ρ_2	0.00821	0.00620	[0.00422, 0.00911]
		ρ_3	0.00821	0.00618	[0.00401, 0.00953]

Note. As discount rates are positively skewed, reported mean discount rates are based on log-transformed values (Kirby 2009) and reported means correspond to the geometric means of the discount rates. For the same reason, confidence intervals are reported instead of standard errors; lower and upper bounds were determined on the log-transformed values and then transformed back into the original space.

In total, 53 subjects returned for the retest. To test whether our results from the retested sample might be biased by selective dropout, we compared the distribution of date 1 discount rates of those that returned for the retest ($N = 53$) and those that did not ($N = 89$). Kolmogorov-Smirnov and Mann-Whitney tests indicate that discount rates were not significantly different between these two groups ($p \geq 0.354$).[10] We take this as

[10]Kolmogorov-Smirnov and Mann-Whitney test statistics of all pairwise comparisons can be found in Table A.1 in Appendix A.

indication that — with respect to the variable of interest — there is no selection effect on the data obtained at test date 2.

Reliability. In order to test the reliability of the measured discount rates, we correlated the individual estimates obtained from the three different sets within one test date. All correlations are high (Spearman's Rho ≥ 0.80, see Table 5), which means that subjects who discount steeply when there is an immediate reward available do so very consistently also when there is no immediate reward available. This indicates that measurement error in the intertemporal choice task is small.

Table 5: Discount Rates - Reliability and Stability

	Within-Session Correlations						Test-Retest
	Date 1 ($N = 142$)			Date 2 ($N = 53$)			Correlation Date 1
	ρ_1	ρ_2	ρ_3	ρ_1	ρ_2	ρ_3	with Date 2 ($N = 53$)
Avg. ρ	0.95^{***}	0.93^{***}	0.88^{***}	0.97^{***}	0.98^{***}	0.92^{***}	0.68^{***}
ρ_1		0.87^{***}	0.80^{***}		0.94^{***}	0.85^{***}	0.61^{***}
ρ_2			0.80^{***}			0.90^{***}	0.68^{***}
ρ_3							0.67^{***}

Note. *** $p < .001$; intercorrelations (Spearman's Rho) of discount rate estimates from three non-overlapping choice sets within the same test session, both for test date 1 and test date 2 (Within-Session Correlations). Test-retest correlations (Spearman's Rho) of discount rates estimated at two separate test dates using the same choice set (Test-Retest).

Note, that this result contrasts with the influential view that immediate rewards are valued separately from delayed rewards (Laibson 1997, McClure et al. 2004), which gives rise to present bias. To further explore this, we tested whether discount rates obtained from our choice set 1 (with immediately available options) are higher than those obtained from the other two sets (without immediately available options). We find that the three estimates for ρ do not differ significantly, neither at test date 1 (Friedman Test, $\chi^2 = 1.369$, $p = 0.490$, $N = 142$), nor at test date 2 (Friedman Test, $\chi^2 = 1.714$, $p = 0.432$, $N = 53$). That is, our subjects are not systematically less patient when the smaller-sooner reward is available immediately, the next day, or in 14 days. Hence, we do not find evidence for a

present bias, which casts some doubt on the generality and robustness of this result (see also Dohmen et al. 2012).[11]

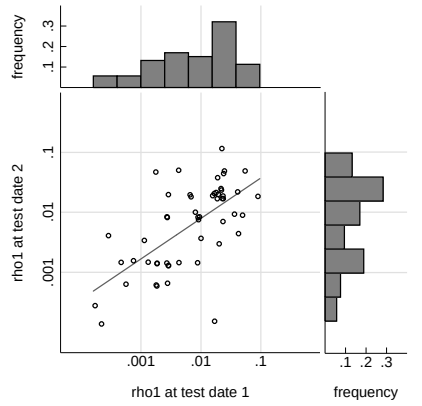
Stability. In the aggregate, discount rates are stable across the two test dates. That is, for the subsample of all subjects that were tested twice, the discount rates do not differ between the two test dates (Wilcoxon Signed-Rank tests, all $p \geq 0.152$). In order to assess stability on the individual level, we correlated the measured discount rates across the two test dates. Figure 1 plots the individual discount rates measured at date 2 against those measured on date 1. We find that discount rates measured at the two dates are highly correlated (Spearman's Rho ≥ 0.61, see also Table 5), indicating that subjects who discount steeply at one test date do so also at the second test date.

Domain Specificity. In both the economic and the psychological literature it is often assumed that intertemporal choice invokes a self-control problem, which results in impatient choice behavior (Laibson 1997, Frederick et al. 2002, McClure et al. 2004, Fudenberg and Levine 2006, Berns et al. 2007). In fact, monetary intertemporal choice tasks, like the one used in the present study, are often employed to measure impulsiveness (e.g. Kirby and Finch 2010, Figner et al. 2010, Crockett et al. 2010). There is, however, no clean evidence on the actual association between such monetary measures and any other psychological construct of impulsiveness.[12]

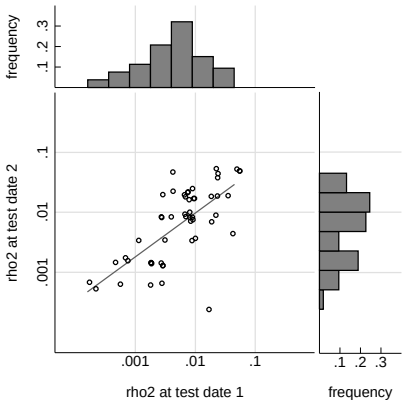
We assessed how strongly behavior in the intertemporal choice task is related to impulsiveness by correlating the estimated discount rates with the total score of the BIS-11 as well its subscale 'Nonplanning Impulsiveness', which is conceptually most closely related to intertemporal choice (see Section 2). In addition, we correlated the discount rates with selected individual items that ask for self-control, saving, spending, and future orientation,

[11] Andreoni and Sprenger (2012a) also do not find evidence for hyperbolic discounting. They, however, use a different presentation format (calendar dates) than we and most studies that report to find hyperbolic discounting do (time distance). It has been shown that calendar date presentation can indeed mitigate hyperbolic discounting (Read et al. 2005).

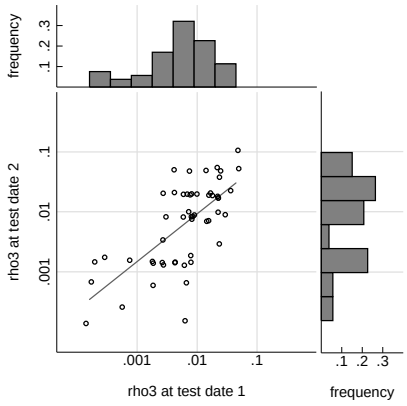
[12] For some earlier attempts to explore this relationship, see Reynolds et al. (2006), de Wit et al. (2007), Kirby and Finch (2010), Verdejo-García et al. (2010).



(a) Set 1



(b) Set 2



(c) Set 3

Note: For each intertemporal choice set, the discount rate at the second test date is plotted as a function of the discount rate at the first test date, together with univariate histograms of each estimated discount rate. Cloud is jittered.

Figure 1: Test-Retest Data for Discount Rates.

core factors that are believed to contribute to intertemporal choice behavior (Frederick et al. 2002).[13]

We find a small significantly positive correlation between the estimates for the discount rate and items relating to saving and spending behavior (see Table 6, items 10 and 25). That is, people who report to rarely save and to spend more than they earn are less patient

[13]Table A.2 in Appendix A reports the test-retest correlations for this questionnaire (overall as well as single items). These are generally high.

Table 6: Correlations of Discount Rates with Self-Reported Impulsiveness

	Barratt Impulsiveness Scale ($N = 142$)					
	Total Score	Subscale Nonplanning	Item 8 Self-Control	Item 10 Saving	Item 25 Spending	Item 27 Present
Average ρ	-0.03	0.00	-0.01	0.19^{*}	0.18^{*}	0.01
ρ_1	-0.03	-0.02	0.01	0.18^{*}	0.13	0.00
ρ_2	-0.05	-0.02	0.05	0.18^{*}	0.19^{*}	-0.03
ρ_3	0.03	0.07	0.02	0.17^{*}	0.20^{*}	0.03

Note. * $p < .05$; correlations (Spearman's Rho) at test date 1.

in the intertemporal choice task. However, responses to an item directly addressing the individual's future-orientation ("I am more interested in the present than in the future") are uncorrelated with the elicited discount rates (Table 6, item 27). Discount rates are also not correlated with responses to the item "I am self-controlled", the subscale "Nonplanning Impulsiveness" or the total sum score of the BIS-11. To check the robustness of these results, we analyzed data of an independent sample that was measured at a different point in time. In this data set we find a similar pattern. The discount rate again correlated most strongly with self-report items on saving and spending.[14]

To summarize, we find that discount rates elicited in a monetary intertemporal choice task are highly correlated for three different intertemporal choice sets and we observe good test-retest correlations for discount rates over an interval of 5–10 weeks. This shows that the used choice task has satisfactory reliability and that time preferences over monetary rewards are remarkably stable over time.

However, we also find that the elicited discount rates do not correlate with self-reported impulsiveness and we do also not observe time inconsistent behavior in the intertemporal choice task. Our results thus add to accumulating evidence suggesting that intertemporal decisions over monetary rewards are not, or only very weakly related to general impulsiveness and lack of self-control.

[14]These results are reported in Table A.5 in Appendix A.

As Borghans et al. (2008) suggest, this could be due to the fact that a typical intertemporal choice task in an economic experiment does not mimic self-control problems outside the lab very well. Alternatively, what is referred to as 'time preference' and 'self-control' might differ across situations within one individual. For example, people might be very concerned about their future health, but not so much about their future financial situation (see Frederick et al. 2002). This idea is supported by our observation that discount rates over money correlate most strongly with self-report items on spending and saving behavior, and not at all with aggregate measures of self-reported impulsiveness. Hence, discounting monetary payoffs in a lab experiment seems to relate specifically to those everyday decisions that involve money. This interpretation is also in line with the observation that discount rates for different types of outcomes, such as health, money and chocolate, are at most moderately correlated (Chapman 1996, Reuben et al. 2010, Tsukayama 2010).

Together, our findings indicate that behavior in an intertemporal choice task is stable over time and captures individual discount rates reliably. How patiently someone acts, however, seems to be domain specific.

Risk Attitudes

Descriptive statistics for the parameter estimates of the curvature of the utility function and the probability weighting function are shown in Table 7. At both test dates we observe moderate levels of risk aversion (median $\alpha = 0.746$ at date 1 and 0.717 at date 2), which are in keeping with earlier results (Andersen et al. 2008a, Bruhin et al. 2010). With regard to the probability weighting function we observe a median γ of 0.413 (date 1) and 0.339, which correspond to pronounced deviations from linear probability weighting ($\gamma = 1$). These estimates are also in line with earlier results (Epper et al. 2009, 2011, Bruhin et al. 2010). When testing for differences at date 1 in estimated risk aversion and probability weighting between subjects that were retested ($N = 53$) and those that did not return for the retest ($N = 91$), Mann-Whitney and Kolmogorov-Smirnov tests did not return significant results at the 5 percent level.[15] We take this as evidence that — regarding the

[15]For α, $p \geq 0.071$, and for γ, $p \geq 0.263$; detailed test results are reported in Appendix A Table A.1.

Table 7: Descriptive Statistics of Risk Attitudes

	N		Median	Mean	95% CI of Mean
Date 1	144	α	0.746	0.784	[0.726, 0.843]
whole sample		γ	0.413	0.442	[0.398, 0.485]
Date 1	53	α	0.693	0.715	[0.649, 0.781]
retest subsample		γ	0.391	0.410	[0.340, 0.480]
Date 2	53	α	0.717	0.659	[0.570, 0.748]
		γ	0.339	0.374	[0.306, 0.441]

variables of interest — any potential selection effect in the retested subsample is negligible.

Reliability. We assessed reliability of these measures of risk attitudes using a split-half procedure. For that purpose, the set of all 20 lotteries was split into two non-overlapping sets using an odd-even split of the lottery numbers. The estimation procedure was then repeated for all odd and all even numbered lotteries separately and the resulting parameter estimates were correlated. If the estimated parameters were distorted by erratic choice behavior and, hence, not reliably measured, then the estimates from the two sets should not be highly correlated. We find that the resulting split-half correlations are in fact very high (Spearman's Rho ≥ 0.78), indicating that the estimated parameters are highly reproducible (see Table 8 for all correlations). Hence, our data suggest that the used lottery choice task is a reliable method for measuring risk aversion and probability weighting.

Stability. To assess whether risk attitudes are stable over time on the aggregate level, we compared the estimates for α and γ across the two test dates. There are no significant differences in risk aversion or probability weighting between the two test dates (Wilcoxon Signed-Rank Tests, all $p \geq 0.174$). In order to evaluate individual test-retest stability, parameter estimates obtained at test date 1 were correlated with those obtained at test date 2. Figure 2 illustrates the test-retest data for α and γ using scatterplots with univariate histograms. As is evident from the scatterplots, test-retest correlations are high (Spearman's Rho ≥ 0.73, see Table 8 for details). Thus, subjects who display strong

Table 8: Risk Attitudes - Reliability and Stability

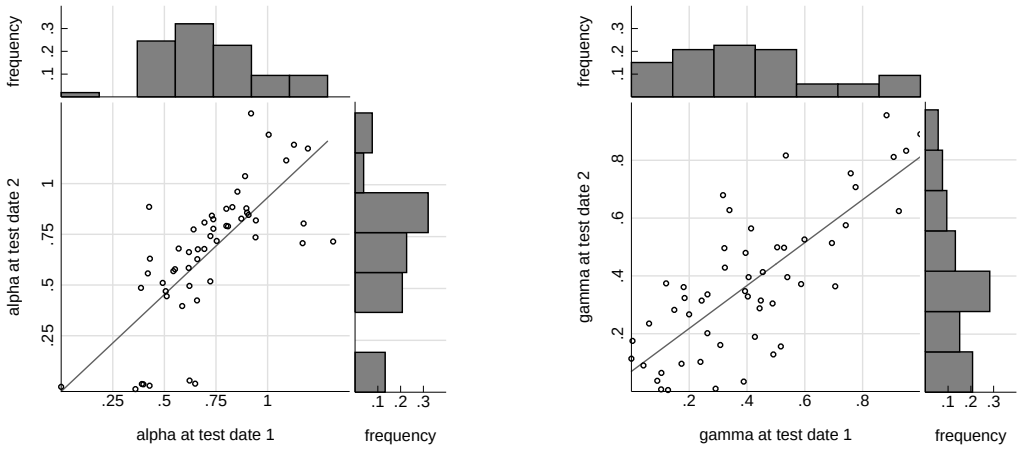
	Within-Session Correlations		Test-Retest
	Split-Half Reliability		Correlation Date 1
	Test Date 1 ($N = 144$)	Test Date 2 ($N = 53$)	with Date 2 ($N = 53$)
α	0.82^{***}	0.82^{***}	0.77^{***}
γ	0.78^{***}	0.83^{***}	0.73^{***}

Note. *** $p < .001$; intercorrelations (Spearman's Rho) of parameter estimates from two non-overlapping subsets of lotteries within the same test session, both for test date 1 and test date 2 (Within-Session Correlations). Test-retest correlations of parameters estimated at two separate test dates using the same choice set (Test-Retest).

risk aversion and high probability weighting at one test date do so also at the other test date.[16] We also looked at test-retest correlations for the individual certainty equivalents of all lotteries. They range from Spearman's Rho = 0.36 to Spearman's Rho = 0.68 and are thereby somewhat lower than the correlations for the aggregate measures. Hence, aggregating choice data from several decisions seems valuable as it increases the reproducibility of the results. This observation could also explain why we observe higher stability than Andersen et al. (2008b), who used a single multiple price list. In our view, this also illustrates that it is important to establish reliability of a measurement technique before assessing stability over time. If reliability at one test date is unknown, low stability over time can be either due to high measurement error or due to true fluctuations over time.

Domain Specificity. Investigating how the tested measures relate to other measures of risk attitudes is important for assessing their ecological validity. Moreover, it can also be informative for a better understanding of the nature of risk attitudes in different domains of life. We correlated the parameter estimates from the lottery choice task with the answers to selected risk questions of the German Socio-Economic Panel (SOEP) discussed

[16]Interestingly, our test-retest correlations are comparable to those reported for the Balloon Analogue Risk Task (White et al. 2008), a behavioral measure of risk attitudes commonly employed in the psychological literature.



(a) Curvature Value Function (b) Probability Weighting

Note: For both parameters, the estimate derived at the second test date is plotted as a function of the estimate from the first test date, together with univariate histograms of each estimated parameter. Cloud is jittered.

Figure 2: Test-Retest Data for Risk Attitudes.

in Section 2 (Table 3).[17] These questions were chosen because they have been shown to correlate well with revealed field behavior in the specified domains.

Table 9 reports the correlations of date 1 estimates of individuals' α and γ with their answers to the various risk questions. It shows that correlations with attitudes towards risk in general and the sum of scores of all risk questions are highly significant but moderate in size (Spearman's Rho between 0.26 and 0.34). The same holds for the risk question regarding the specific domains financial investments (Spearman's Rho 0.27 and 0.30) and confidence in strangers (Spearman's Rho 0.23 and 0.19). In these domains, subjects who report to take more risks show also less risk aversion and less deviations from linear probability weighting in the lottery choice task. However, risk attitudes elicited with the lottery choice task are entirely uncorrelated to self-reported attitudes towards risk in other domains such as car driving, leisure and sports activities, and health.[18] This strongly

[17]Table A.3 in Appendix A reports the test-retest correlations for this questionnaire in our sample. These are generally high.

[18]These results seem robust as we replicated these findings in an entirely independent sample (see

Table 9: Correlations of Risk Attitudes with Self-Reported Risk Taking

	Attitudes Towards Risk ($N = 144$)							
	In General	Driving a Car	Financial Investments	Leisure & Sports	Career	Health	Strangers	Sum Score
α	0.34^{***}	0.07	0.27^{**}	0.10	0.18^{*}	0.11	0.23^{**}	0.31^{***}
γ	0.31^{***}	0.09	0.30^{***}	0.04	0.09	0.15	0.19^{*}	0.26^{**}

Note. $^{***\ (**)\ \{*\}}$ $p < .001$ $(p < .01)$ $\{p < .05\}$; correlations (Spearman's Rho) at test date 1.

suggests that individual risk attitudes are domain specific.

To summarize, we find that parameter estimates for risk aversion and probability weighting parameters obtained from lottery tasks show high correlation across non-overlapping (split-half) sets of lotteries. Further, test-retest correlations over an interval of 5–10 weeks are high for risk aversion as well as probability weighting. This shows that the used lottery choice task measures important aspects of risk attitudes with satisfactory reliability and that behavior in this task is remarkably stable over time even on the individual level. The evidence regarding ecological validity and generalizability of the investigated measures is mixed, however. On the one hand, estimates of risk aversion and probability weighting both correlate statistically significantly with overall (self-reported) risk taking as well as with risk taking in the financial domain, which may be considered as similar to the examined lottery tasks. On the other hand, these correlations are economically only moderately significant. Moreover, for domains that are very dissimilar from money lottery tasks, risk taking is not at all correlated with the estimated parameters. Hence, it seems that attitudes toward risk are highly domain specific and different domains may demand different measures for reliably assessing risk taking behavior.[19]

Table A.4).

[19]For recent evidence suggesting a similar interpretation of risk taking behavior see Barseghyan and Prince (2011), Dohmen et al. (2011), and Einav et al. (2012).

4 Discussion and Concluding Remarks

We provide empirical evidence on the reliability, stability, and domain specificity for behavioral measures that are frequently used to elicit time preferences and risk attitudes in economic experiments. First, we find that parameters values for time discounting, risk aversion, and probability weighting, estimated from mutually exclusive subsets of the choice tasks are highly correlated, indicating that these behavioral measures are reliable and provide reproducible estimates. Second, all three estimated measures show high test-retest correlations over a period of 5–10 weeks. Together with the established reliability of the explored measures this allows us to conclude that the elicited attitudes are stable at least over the tested period. Third, we find that discount rates elicited with a monetary decision task correlate with self-reports on excess spending and low savings but not with those for other domains, like general planning abilities and self-control. Similarly, we observe that risk attitudes elicited with a lottery choice task correlate most with self-reported attitudes towards risks in financial investments but are largely unrelated with risk attitudes in other important domains of life, like health and leisure.

Together our results indicate that the explored measures may be well-suited for predicting field behavior regarding financial decisions involving time and risk trade-offs but ill suited when it comes to behavior outside the narrow financial realm.

In our study, we measure time and risk preferences independently. However, time preferences may be intertwined with risk preferences (Frederick et al. 2002, Andersen et al. 2008a, Andreoni and Sprenger 2012b, Epper and Fehr-Duda 2013), in which case discount rates should correlate with risk aversion. We observe only a small and marginally significant correlation between the discount rate and the risk aversion parameter. Subjects that are less risk averse tend to be more patient.[20] The portion of explained variance is small, however. This suggests that risk inherent to future payoffs governs intertemporal decision making only to a small extent. Alternatively, as we find, people react differently to various forms of risk, and this may also apply to the risk inherent in an intertemporal choice as opposed to a lottery choice task.

[20]For details, see Table A.6 in Appendix A.

We cannot control for credit constraints of our subjects. Therefore, the observed correlation of excess spending and low savings with discounting could arise from current cash constraints of these subjects. However, Meier and Sprenger (2010a) found that present-biased preferences in an intertemporal choice task were associated with increased credit card debt even after controlling for credit constraints. This indicates that cash constraints are not the only factor driving the association between intertemporal choice in the laboratory and spending behavior.

We provide evidence for the relation of behaviorally elicited preference parameters with conceptually related self-report scales. When evaluating the association between self-report instruments and behavioral measures of the same construct it is important to consider the differences in methodology (Campbell and Fiske 1959, Robins et al. 2007). Self-report scales might only imperfectly reflect subjects' true dispositions because self-reports can be distorted by motives like self-enhancement. Further, self-report scales like the BIS-11 focus on traits, which prompts subjects to reflect on their habits across different situations. Therefore, we have carefully chosen questionnaires which have been validated to correlate with behavior in domains of interest. Nevertheless, the identified correlations may establish only a lower bound for the true relational strength, as the behavioral measures may be influenced by situational factors, such as task-specificity or the current financial situation of the subject.

We find that measures of individual time preferences and risk attitudes are stable over time, but not across different decision domains. This strongly suggests that these individual characteristics are domain specific. Hence, in order to make good predictions, the differences as well as similarities of preferences and behavior in different domains of life need to be more accurately understood. Therefore, an important avenue for future research would be to more systematically explore the domain specificity of time preferences and risk attitudes and to establish preference elicitation tasks that possess high predictive validity across as well as within different decision domains.

References

Ainslie, G. (1975). Specious reward: a behavioral theory of impulsiveness and impulse control. *Psychological Bulletin*, 82(4):463–496.

Andersen, S., Harrison, G. W., Lau, M. I., and Rutström, E. E. (2008a). Eliciting risk and time preferences. *Econometrica*, 76(3):583–618.

Andersen, S., Harrison, G. W., Lau, M. I., and Rutström, E. E. (2008b). Lost in state space: Are preferences stable? *International Economic Review*, 49(3):1091–1112.

Andreoni, J. and Sprenger, C. (2012a). Estimating time preferences from convex budgets. *American Economic Review*, 102(7):3333–3356.

Andreoni, J. and Sprenger, C. (2012b). Risk preferences are not time preferences. *American Economic Review*, 102(7):3357–3376.

Ashraf, N., Karlan, D., and Yin, W. (2006). Tying Odysseus to the mast: Evidence from a commitment savings product in the philippines. *The Quarterly Journal of Economics*, 121(2):635–672.

Barseghyan, L. and Prince, J. (2011). Are risk preferences stable across contexts? Evidence from insurance data. *American Economic Review*, 101(2):591–631.

Berns, G. S., Laibson, D., and Loewenstein, G. (2007). Intertemporal choice - toward an integrative framework. *Trends in Cognitive Sciences*, 11(11):482–488.

Borghans, L., Duckworth, A. L. A., Heckman, J. J., and ter Wel, B. (2008). The economics and psychology of personality traits. *Journal of Human Resources*, 43(4):972 – 1059.

Borghans, L. and Golsteyn, B. H. (2006). Time discounting and the body mass index: Evidence from the Netherlands. *Economics and Human Biology*, 4(1):39 – 61.

Bruhin, A., Fehr-Duda, H., and Epper, T. (2010). Risk and rationality: Uncovering heterogeneity in probability distortion. *Econometrica*, 78(4):1375–1412.

Campbell, D. T. and Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56(2):81–105.

Chabris, C. F., Laibson, D., Morris, C. L., Schuldt, J. P., and Taubinsky, D. (2008). Individual laboratory-measured discount rates predict field behavior. *Journal of Risk and Uncertainty*, 37(2-3):237–269.

Chapman, G. B. (1996). Temporal discounting and utility for health and money. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22(3):771 – 791.

Crockett, M., Clark, L., Lieberman, M., Tabibnia, G., and Robbins, T. (2010). Impulsive choice and altruistic punishment are correlated and increase in tandem with serotonin depletion. *Emotion*, 10(6):855–862.

de Wit, H., Flory, J. D., Acheson, A., McCloskey, M., and Manuck, S. B. (2007). IQ and nonplanning impulsivity are independently associated with delay discounting in middle-aged adults. *Personality and Individual Differences*, 42(1):111–121.

Dohmen, T., Falk, A., Huffman, D., and Sunde, U. (2010). Are risk aversion and impatience related to cognitive ability? *American Economic Review*, 100(3):1238–1260.

Dohmen, T., Falk, A., Huffman, D., and Sunde, U. (2012). Interpreting time horizon effects in inter-temporal choice. *IZA Discussion Paper 6385*.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., and Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants and behavioral consequences. *Journal of the European Economic Association*, 9(3):522–550.

Einav, L., Finkelstein, A., Pascu, I., and Cullen, M. R. (2012). How general are risk preferences? Choices under uncertainty in different domains. *American Economic Review*, 102(6):2606–2638.

Epper, T., Bruhin, A., and Fehr-Duda, H. (2009). Uncertainty breeds decreasing impatience: The role of risk preferences in time discounting. *IEW Discussion Paper*.

Epper, T. and Fehr-Duda, H. (2013). The missing link: Unifying risk taking and time discounting. *Working Paper No. 96*.

Epper, T., Fehr-Duda, H., and Bruhin, A. (2011). Viewing the future through a warped lens: Why uncertainty generates hyperbolic discounting. *Journal of Risk and Uncertainty*, 43(3):169–203.

Fehr-Duda, H. and Epper, T. (2012). Probability and risk: Foundations and economic implications of probability-dependent risk preferences. *Annual Review of Economics*, 4:567–593.

Figner, B., Knoch, D., Johnson, E. J., Krosch, A. R., Lisanby, S. H., Fehr, E., and Weber, E. U. (2010). Lateral prefrontal cortex and self-control in intertemporal choice. *Nature Neuroscience*, 13(5):538–539.

Fischbacher, U. (2007). z-Tree: Zurich Toolbox of Ready-made Economics Experiments. *Experimental Economics*, 10(2):171–178.

Fox, C. R. and Tannenbaum, D. (2011). The elusive search for stable risk preferences. *Frontiers in Psychology*, 2:298.

Frederick, S., Loewenstein, G., and O'Donoghue, T. (2002). Time discounting and time preference: A critical review. *Journal of Economic Literature*, 40(2):351–401.

Fudenberg, D. and Levine, D. K. (2006). A dual-self model of impulse control. *American Economic Review*, 96(5):1449–1476.

Gonzalez, R. and Wu, G. (1999). On the shape of the probability weighting function. *Cognitive Psychology*, 38:129–166.

Greiner, B. (2004). An online recruitment system for economic experiments. In Kremer, K. and Macho, V., editors, *Forschung und wissenschaftliches Rechnen. Beiträge zum Heinz-Billing-Preis 2003. GWDG Bericht 63*, pages 79–93. Gesellschaft für wissenschaftliche Datenverarbeitung mbH, Göttingen.

Hey, J. D. and Lee, J. (2005). Do subjects separate (or are they sophisticated)? *Experimental Economics*, 8(3):233–265.

Holt, C. A. and Laury, S. K. (2002). Risk aversion and incentive effects. *American Economic Review*, 92(5):1644–1655.

Ifcher, J. and Zarghamee, H. (2011). Happiness and time preference: The effect of positive affect in a random-assignment experiment. *American Economic Review*, 101(7):3109–29.

Kable, J. W. and Glimcher, P. W. (2007). The neural correlates of subjective value during intertemporal choice. *Nature Neuroscience*, 10(12):1625–1633.

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–291.

Kirby, K. N. (2009). One year temporal stability of discount rates. *Psychonomic Bulletin & Review*, 16(3):457–462.

Kirby, K. N. and Finch, J. C. (2010). The hierarchical structure of self-reported impulsivity. *Personality and Individual Differences*, 48(6):704–713.

Kirby, K. N., Petry, N. M., and Bickel, W. K. (1999). Heroin addicts have higher discount rates for delayed rewards than non-drug-using controls. *Journal of Experimental Psychology: General*, 128(1):78–87.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics*, 112(2):443–477.

Logue, A. W. (1988). Research on self-control: an integrating framework. *Behavioral and Brain Sciences*, 11(4):665–709.

McClure, S. M., Laibson, D. I., Loewenstein, G., and Cohen, J. D. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science*, 306(5695):503–507.

Meier, S. and Sprenger, C. (2010a). Present-biased preferences and credit card borrowing. *American Economic Journal: Applied Economics*, 2(1):193–210.

Meier, S. and Sprenger, C. (2010b). Temporal stability of time preferences. *Working Paper*.

Patton, J. H., Stanford, M. S., and Barratt, E. S. (1995). Factor structure of the Barratt Impulsiveness Scale. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, 51(6):768–774.

Pennings, J. and Smidts, A. (2000). Assessing the construct validity of risk attitude. *Management Science*, 46(10):1337–1348.

Prelec, D. (1998). The probability weighting function. *Econometrica*, 66(3):497–527.

Read, D., Frederick, S., Orsel, B., and Rahman, J. (2005). Four score and seven years from now: The date/delay effect in temporal discounting. *Management Science*, 51(9):1326–1335.

Reimers, S., Maylor, E. a., Stewart, N., and Chater, N. (2009). Associations between a one-shot delay discounting measure and age, income, education and real-world impulsive behavior. *Personality and Individual Differences*, 47(8):973–978.

Reuben, E., Sapienza, P., and Zingales, L. (2009). Procrastination and Impatience. *NBER Working Paper*.

Reuben, E., Sapienza, P., and Zingales, L. (2010). Time discounting for primary and monetary rewards. *Economics Letters*, 106(2):125–127.

Reynolds, B., Ortengren, A., Richards, J., and Dewit, H. (2006). Dimensions of impulsive behavior: personality and behavioral measures. *Personality and Individual Differences*, 40(2):305–315.

Robins, R. W., Fraley, R. C., and Krueger, R. F., editors (2007). *Handbook of Research Methods in Personality Psychology*. Guilford, New York.

Stanford, M. S., Mathias, C. W., Dougherty, D. M., Lake, S. L., Anderson, N. E., and Patton, J. H. (2009). Fifty years of the Barratt Impulsiveness Scale: An update and review. *Personality and Individual Differences*, 47(5):385–395.

Stott, H. P. (2006). Cumulative prospect theory's functional menagerie. *Journal of Risk and Uncertainty*, 32(2):101–130.

Thaler, R. (1981). Some empirical evidence on dynamic inconsistency. *Economics Letters*, 8(3):201–207.

Trepel, C., Fox, C. R., and Poldrack, R. A. (2005). Prospect theory on the brain? Toward a cognitive neuroscience of decision under risk. *Cognitive Brain Research*, 23(1):34–50.

Tsukayama, E. (2010). Domain-specific temporal discounting and temptation. *Judgment and Decision Making*, 5(2):72–82.

Verdejo-García, A., Lozano, O., Moya, M., Alcázar, M. A., and Pérez-García, M. (2010). Psychometric properties of a Spanish version of the UPPS-P impulsive behavior scale: reliability, validity and association with trait and cognitive impulsivity. *Journal of Personality Assessment*, 92(1):70–7.

Wakker, P. P. (2008). Explaining the characteristics of the power (CRRA) utility family. *Health Economics*, 17(12):1329–1344.

White, T. L., Lejuez, C. W., and de Wit, H. (2008). Test-retest characteristics of the Balloon Analogue Risk Task (BART). *Experimental and Clinical Psychopharmacology*, 16(6):565–570.

Zeisberger, S., Vrecko, D., and Langer, T. (2012). Measuring the time stability of prospect theory preferences. *Theory and Decision*, 72(3):359–386.

A Additional Descriptive Statistics and Tests

Table A.1: Comparison of the Retested and Not-Retested Subsamples at Test Date 1

	Median		K-S	M-W
	not retested ($N = 89$)	retested ($N = 53$)	p	p
ρ_1	0.00821	0.00821	0.345	0.919
ρ_2	0.00821	0.00821	0.951	0.354
ρ_3	0.00737	0.00737	0.782	0.587
	not retested ($N = 91$)	retested ($N = 53$)	p	p
α	0.790	0.693	0.178	0.071
γ	0.429	0.391	0.506	0.263

Note. Comparison of the discount rates and risk attitudes for retested and non-retested subsamples, using Kolmogorov-Smirnov Test (K-S) and Mann-Whitney Test (M-W). No test indicates significant differences between retested and not retested subsamples.

Table A.2: Test Retest Correlations of Barratt Impulsiveness Scale

	Barratt Impulsiveness Scale ($N = 53$)					
	Total Score	Subscale Nonplanning	Item 8 Self-Control	Item 10 Saving	Item 25 Spending	Item 27 Present
Test-Retest Correlation	0.869^{***}	0.855^{***}	0.317^{*}	0.735^{***}	0.627^{***}	0.558^{***}

Note. *** $p < .001$, ** $p < .01$, * $p < .05$; correlation statistics are Spearman's Rho.

Table A.3: Test-Retest Correlations of General Risk Questions

	Attitudes Towards Risk ($N = 53$)							
	In General	Driving a Car	Financial Investments	Leisure and Sports	Career	Health	Strangers	Sum Score
Test-Retest Correlation	0.813^{***}	0.818^{***}	0.677^{***}	0.540^{***}	0.688^{***}	0.792^{***}	0.708^{***}	0.763^{***}

Note. *** $p < .001$, ** $p < .01$, * $p < .05$; correlation statistics are Spearman's Rho.

Table A.4: Correlations of Risk Attitudes with Self-Report Items in Independent Sample

	Attitudes Towards Risk ($N = 128$)							
	In General	Driving a Car	Financial Investments	Leisure and Sports	Career	Health	Strangers	Sum Score
α	0.28^{**}	0.14	0.30^{***}	0.09	0.22^{*}	0.07	0.27^{**}	0.31^{***}
γ	0.22^{*}	0.11	0.29^{**}	0.08	0.16	−0.03	0.16	0.22^{*}

Note. *** $p < .001$, ** $p < .01$, * $p < .05$; correlation statistics are Spearman's Rho; data were obtained from an entirely independent sample and measured in exactly the same way as reported in the main text.

Table A.5: Correlations of Discount Rates with Self-Report Items in Independent Sample

	Barratt Impulsiveness Scale ($N = 128$)					
	Total Score	Subscale Nonplanning	Item 8 Self-Control	Item 10 Saving	Item 25 Spending	Item 27 Present
Averageρ	0.13	0.19^{*}	0.26^{**}	0.39^{***}	0.30^{***}	0.02
ρ_1	0.14	0.18^{*}	0.24^{**}	0.37^{***}	0.29^{***}	0.07
ρ_2	0.04	0.12	0.23^{**}	0.32^{***}	0.22^{*}	0.01
ρ_3	0.12	0.16	0.20^{*}	0.33^{***}	0.27^{**}	−0.03

Note. *** $p < .001$, ** $p < .01$, * $p < .05$; correlation statistics are Spearman's Rho; data were obtained from an entirely independent sample and measured in exactly the same way as reported in the main text.

Table A.6: Correlation of Time Preferences with Risk Preferences

	α	γ
Average ρ	-0.14^{a}	-0.09
ρ_1	-0.16^{a}	-0.11
ρ_2	-0.15^{a}	-0.11
ρ_3	−0.05	-0.01
	N=142	

Note. ^{a} $p < .1$

B Self-Report Scales

Table B.1: Items of the Barratt Impulsiveness Scale (Patton et al. 1995)

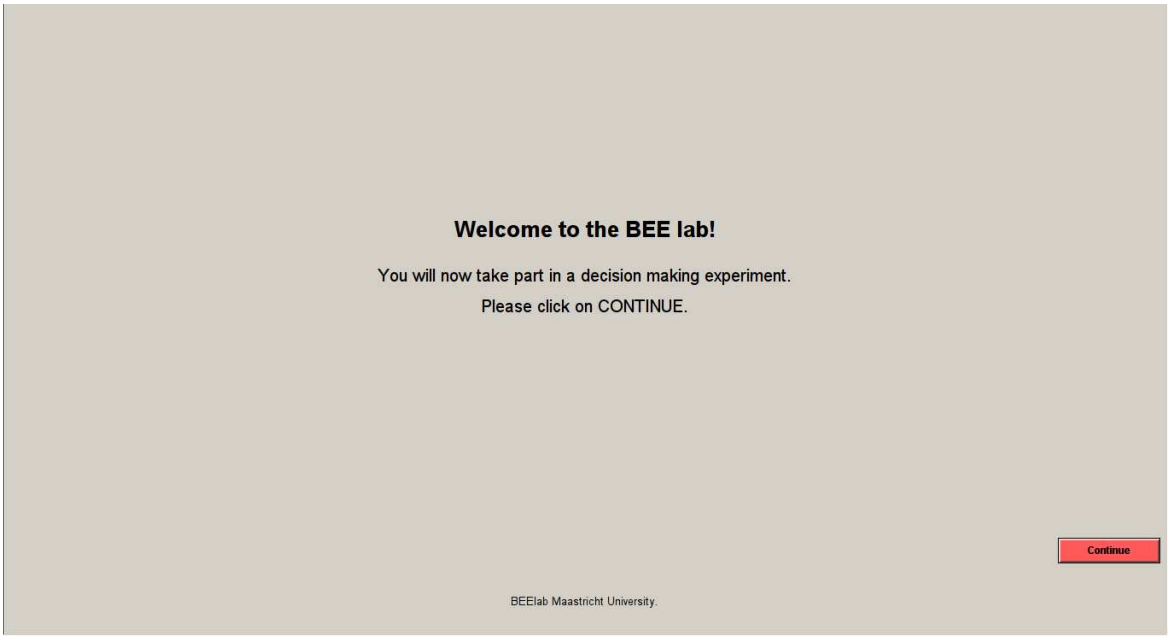
Order	Item[a]	Subscale
1	I plan tasks carefully.[b]	Nonplanning
2	I do things without thinking.	Motor
3	I make up my mind quickly.	Motor
4	I am happy go-lucky.	Motor
5	I don't pay attention.	Attention
6	I have "racing" thoughts.	Attention
7	I plan trips well ahead of time.[b]	Nonplanning
8	I am self-controlled.[b]	Nonplanning
9	I concentrate easily.[b]	Attention
10	I save regularly.[b]	Nonplanning
11	I squirm at plays or lectures.	Attention
12	I am a careful thinker.[b]	Nonplanning
13	I plan for job security.[b]	Nonplanning
14	I say things without thinking.	Nonplanning
15	I like to think about complex problems.[b]	Nonplanning
16	I change jobs.	Motor
17	I act "on impulse".	Motor
18	I easily get bored when solving thought problems.	Nonplanning
19	I act on the spur of the moment.	Motor
20	I am a steady thinker.[b]	Attention
21	I change residences.	Motor
22	I buy things on impulse.	Motor
23	I can only think about one problem at a time.	Motor
24	I change hobbies	Attention
25	I spend or charge more than I earn.	Motor
26	I often have extraneous thoughts when thinking.	Attention
27	I am more interested in the present than in the future.	Nonplanning
28	I am restless at the theatre or lectures.	Attention
29	I like puzzles.[b]	Nonplanning
30	I am future oriented.[b]	Motor

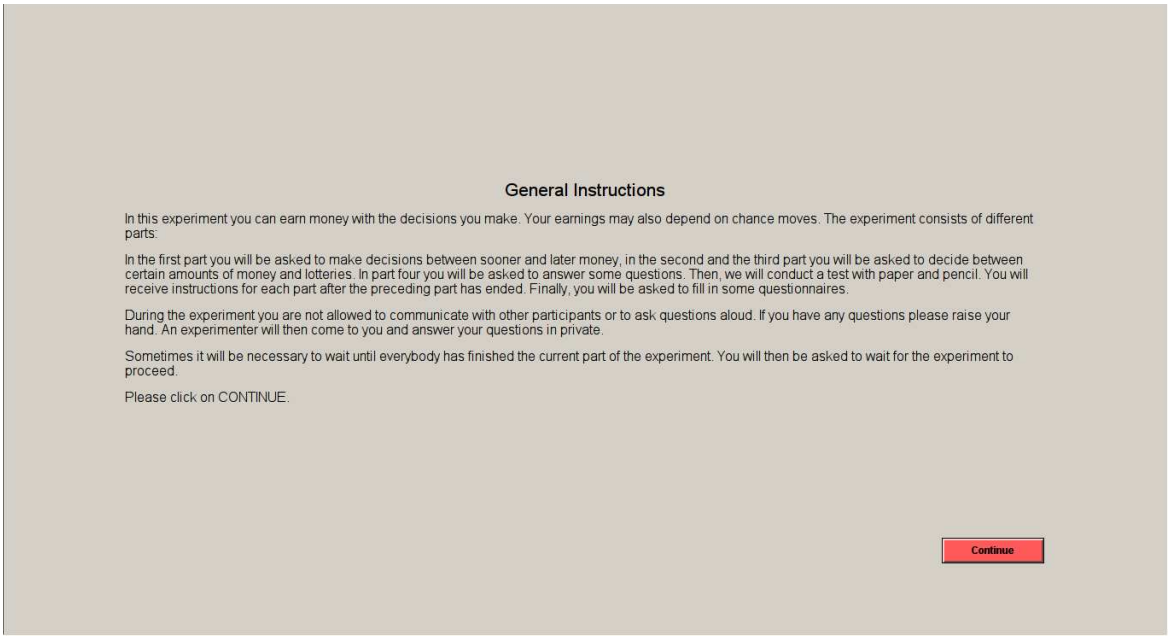
[a] Subjects rated each item on the following 4-point rating scale: 1 (rarely/never) 2 (occasionally) 3 (often) 4 (always/almost always).

[b] Item scoring was reversed.

C Experiment Instructions

The experiment was fully computerized. The following shows the computer screens with instructions, choice tasks, and questionnaires as presented to the subjects.





Determination of Earnings

At the end of the experiment, one of all decisions you made will be randomly selected by the computer program. Each of the decision situations has the same chance to be selected. Your earnings will be determined by the decision you made in the chosen decision situation.

Some of the decision situations involve money that you will receive not today, but after a certain number of days.

All immediate earnings will be paid out in cash individually and confidentially at the end of the experiment. All future earnings will be paid by bank transfer at the time specified. If you receive future earnings, you will be given a statement that specifies the amount and the time of earnings at the end of the experiment. In addition, you will receive an Email to remind you that the transaction was made.

All payments will be carried out by another person than the experimenter. I as an experimenter will not get to know your earnings. Your anonymity is thus secured.

Please click on CONTINUE.

Continue

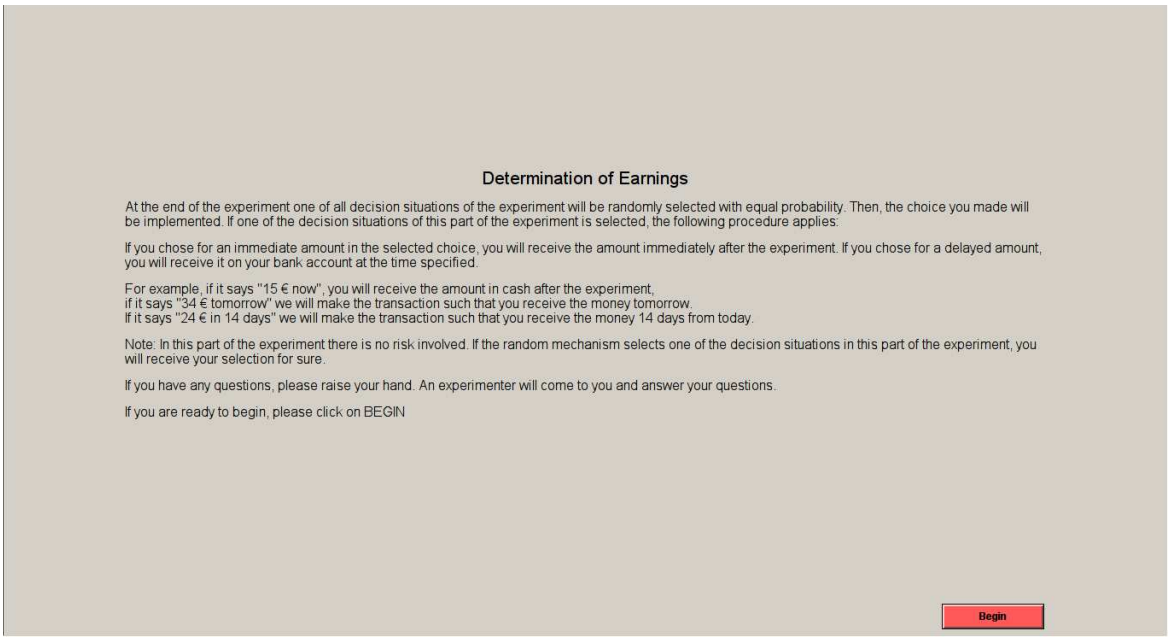
Part 1

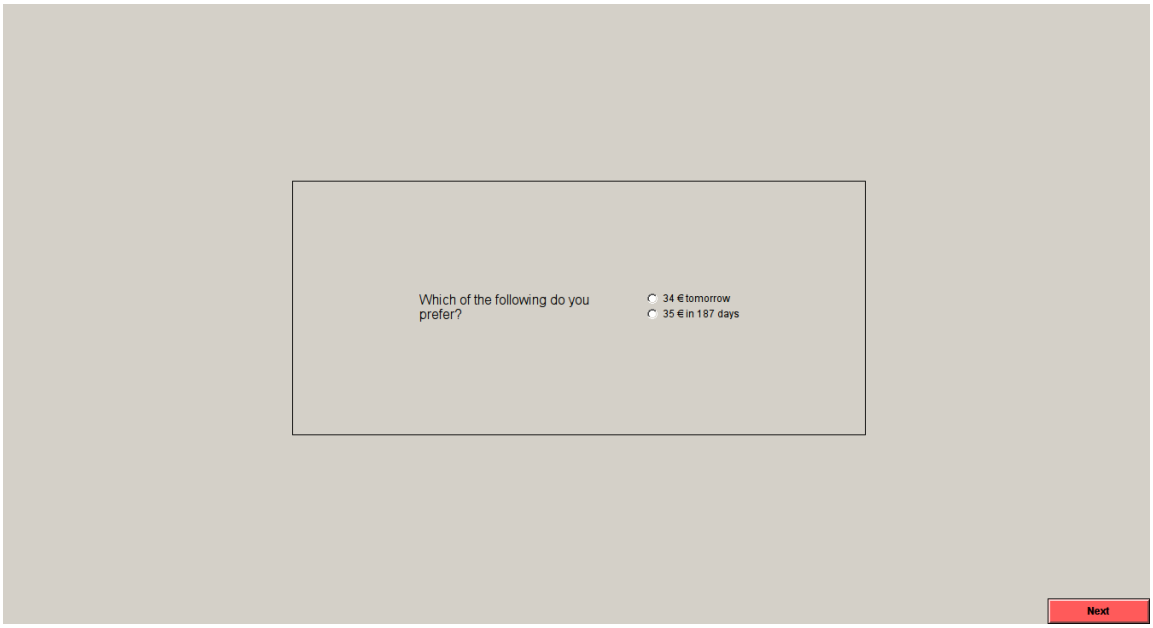
In the first part of the experiment you will face a series of decision situations where you have to make choices between two different amounts of money. In each decision situation there is always a smaller amount, which you will obtain sooner, and a larger amount, which you will receive later.

Please note that each decision situation has the same likelihood to be the one that is relevant for your earnings. All these decision situations are completely independent of each other. A choice you made in one decision situation does not affect any of the other following decision situations. Therefore, you should **view each decision independently** and consider all your choices carefully.

Please click on CONTINUE.

Continue





In total subjects went through 27 of such intertemporal choice tasks. The other tasks and the order of tasks can be found in Table 1 of the main text.

Part 2

In the second part of the experiment you will face another series of choices. These choices will not influence the earnings you can make from the choices in the first part of the experiment, nor will your earlier choices influence the earnings you can make with your choices in this part of the experiment.

In the following, you will be confronted with a series of 20 decision situations, that will appear in random order on the screen. All these decision situations are completely independent of each other. A choice you make in one decision situation does not affect any of the other following decision situations.

In each decision situation, there is always the choice between a lottery and varying secure options. Each decision situation is displayed on a screen. The screen consists of 20 rows. For **every** row you have to decide whether you prefer **option A** or **option B**.

Option A is a lottery and is the same for every row in a given decision situation. Option B is a secure option and takes 20 different values, one for each row.

Please note that each decision situation has the same likelihood to be the one that is relevant for your earnings. Therefore, you should **view each decision independently** and consider all of your choices carefully.

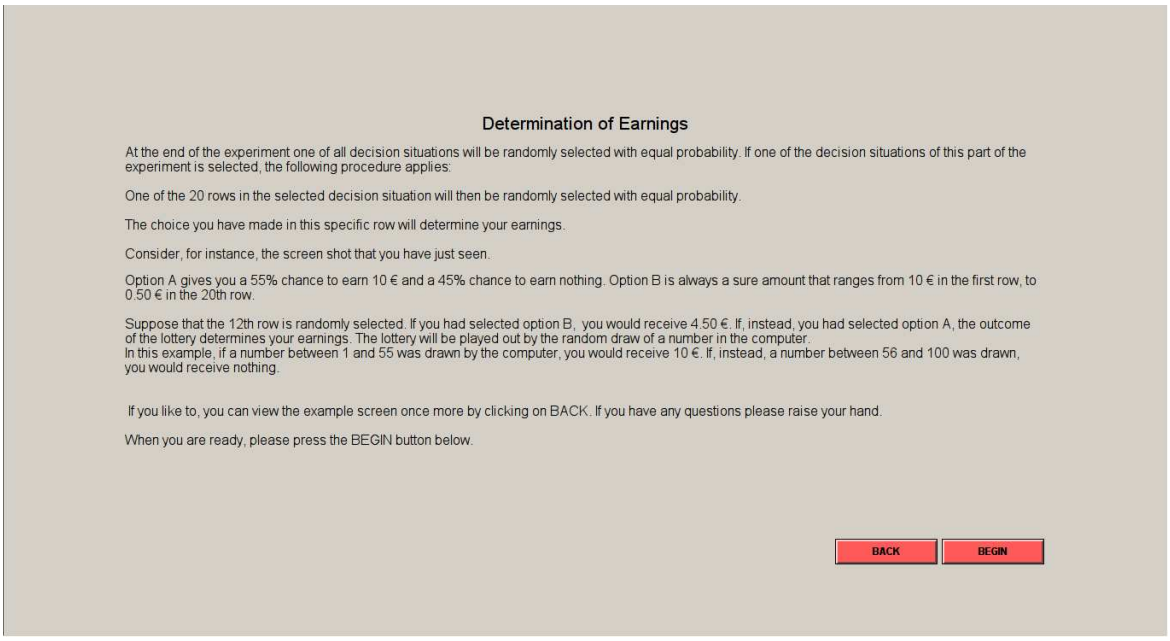
By clicking on NEXT you will see an example screen of a decision situation.

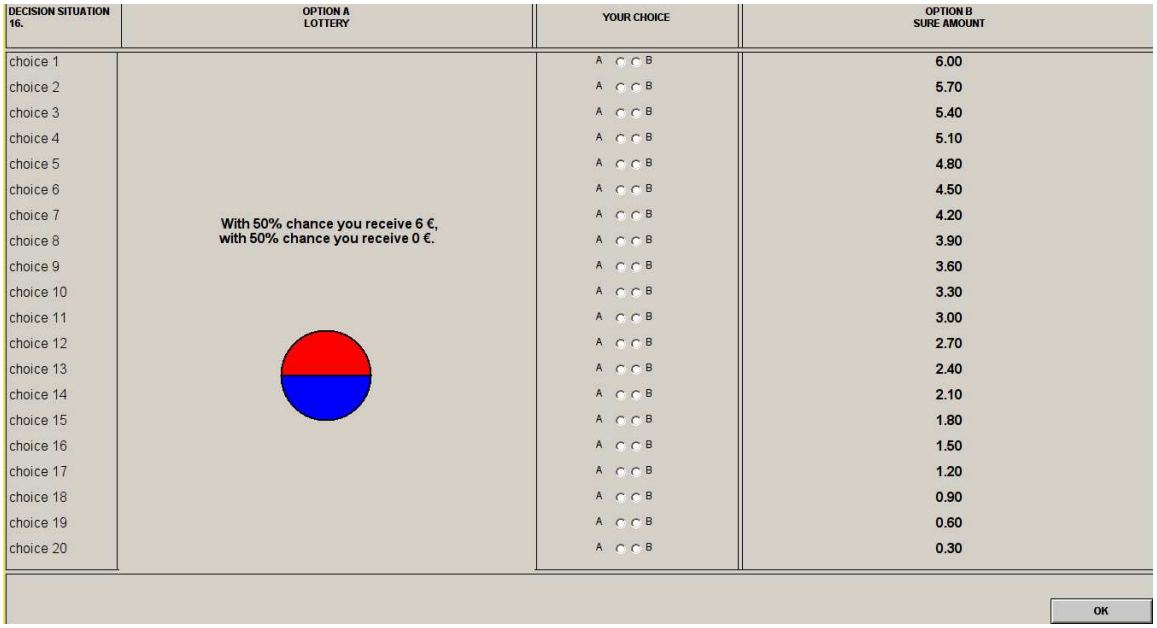
NEXT

This is a **screen shot** of one decision situation you are going to face. You are not asked to make choices now! Please have a careful look. Thereafter click on NEXT to procede.

DECISION SITUATION 13.	OPTION A LOTTERY	YOUR CHOICE	OPTION B SURE AMOUNT
choice 1	With 55% chance you receive 10.- Euro, with 45% chance you receive nothing.	A B	10.-
choice 2		A B	9.50
choice 3		A B	9.-
choice 4		A B	8.50
choice 5		A B	8.-
choice 6		A B	7.50
choice 7		A B	7.-
choice 8		A B	6.50
choice 9		A B	6.-
choice 10		A B	5.50
choice 11		A B	5.-
choice 12		A B	4.50
choice 13		A B	4.-
choice 14		A B	3.50
choice 15		A B	3.-
choice 16		A B	2.50
choice 17		A B	2.-
choice 18		A B	1.50
choice 19		A B	1.-
choice 20		A B	0.50

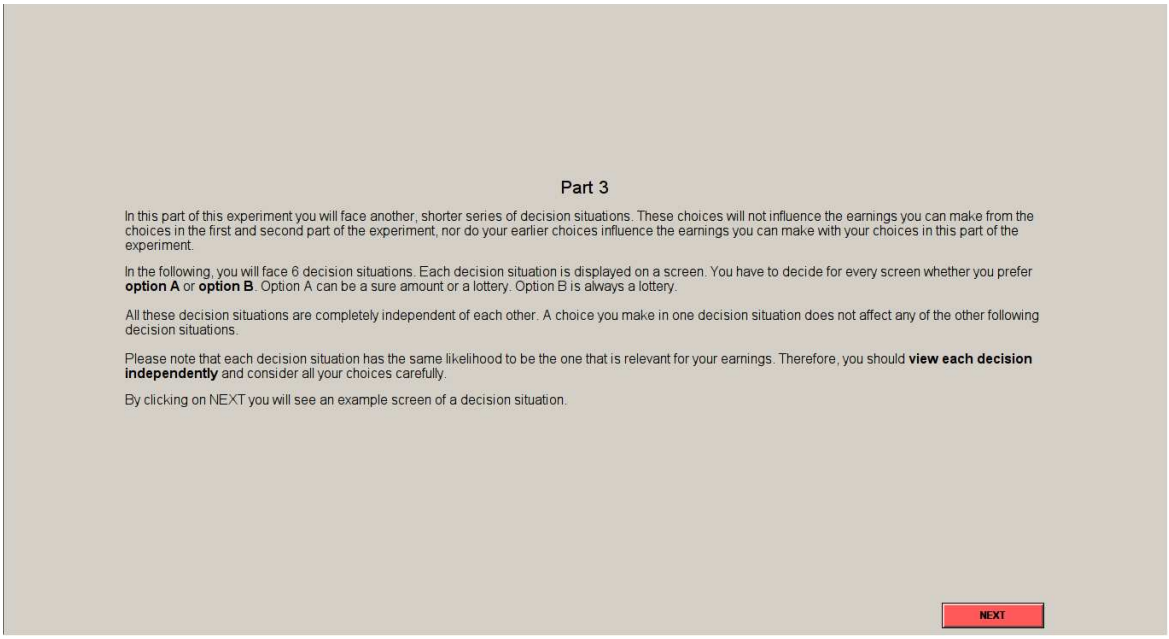
NEXT

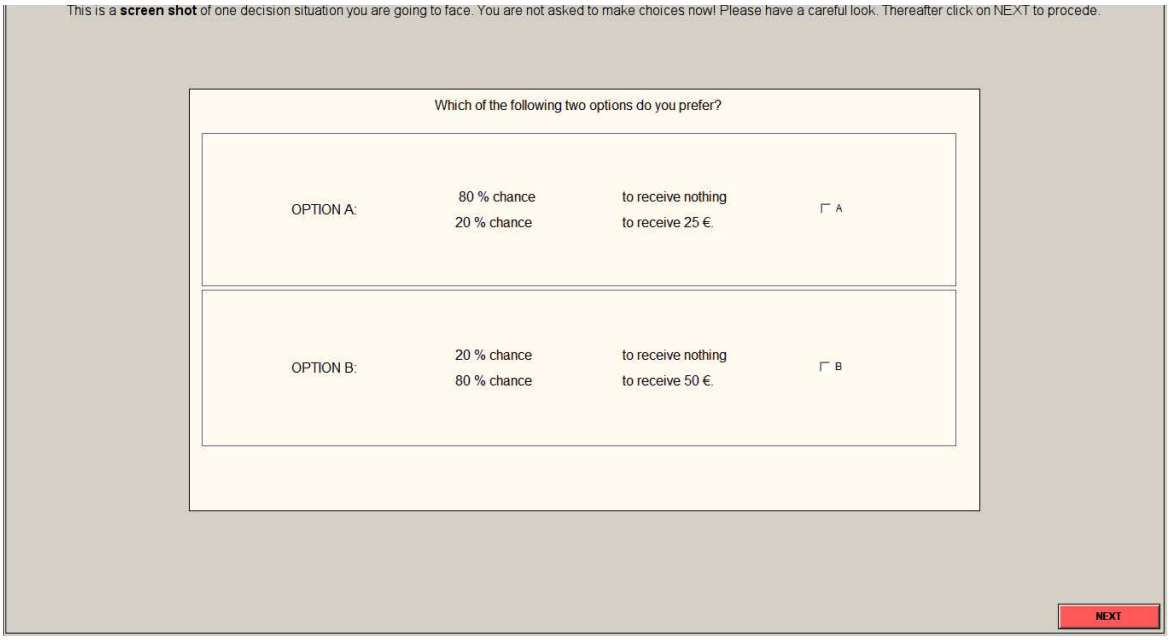


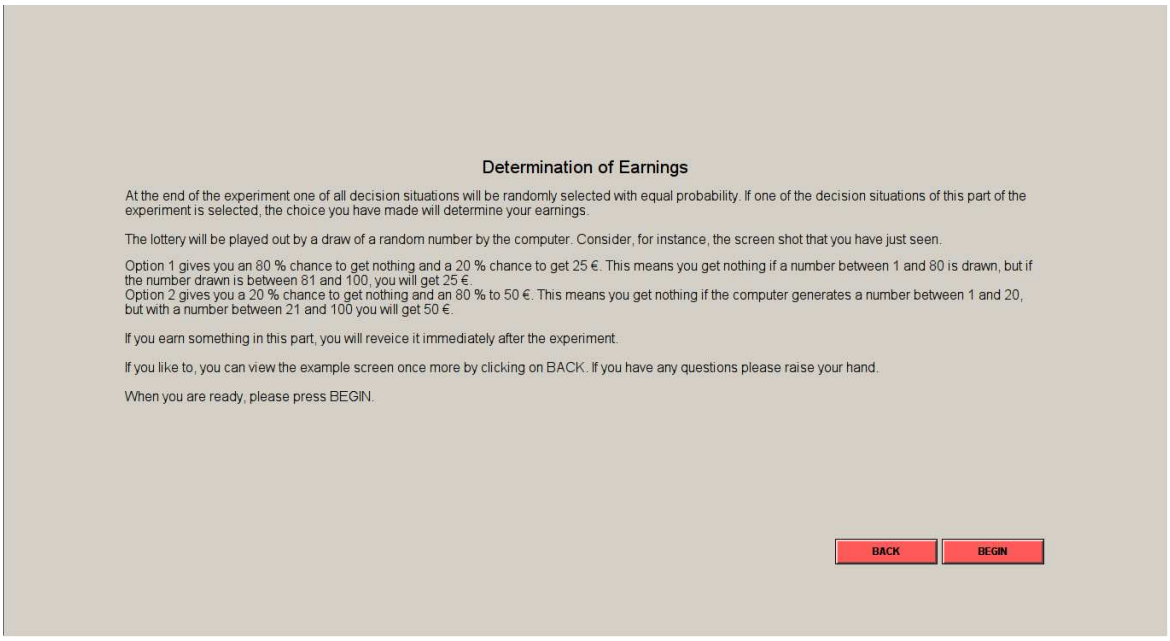


In total subjects went through 20 of such lottery choice tasks. The other tasks and the order of tasks can be found in Table 2 of the main text.

Next to the intertemporal and lottery choice tasks, subjects also participated in the independent choice tasks described as Part 3 and Part 4 below. We do not report on data gathered in these parts in the current paper.

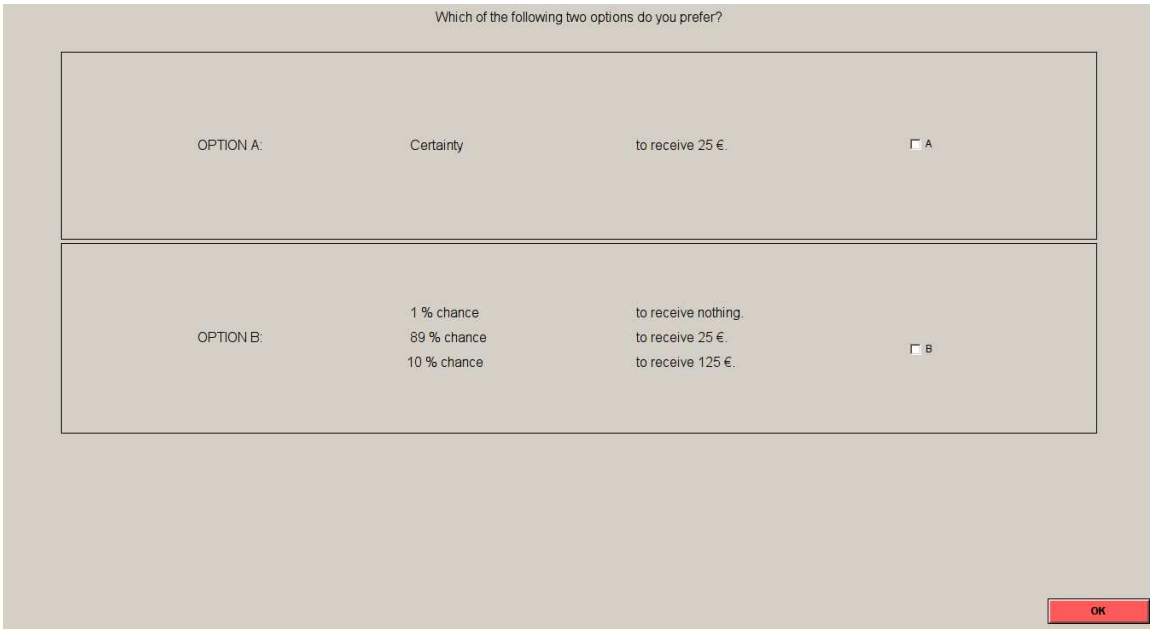





Determination of Earnings
At the end of the experiment one of all decision situations will be randomly selected with equal probability. If one of the decision situations of this part of the experiment is selected, the choice you have made will determine your earnings.
The lottery will be played out by a draw of a random number by the computer. Consider, for instance, the screen shot that you have just seen.
Option 1 gives you an 80 % chance to get nothing and a 20 % chance to get 25 €. This means you get nothing if a number between 1 and 80 is drawn, but if the number drawn is between 81 and 100, you will get 25 €.
Option 2 gives you a 20 % chance to get nothing and an 80 % to 50 €. This means you get nothing if the computer generates a number between 1 and 20, but with a number between 21 and 100 you will get 50 €.
If you earn something in this part, you will reveice it immediately after the experiment.
If you like to, you can view the example screen once more by clicking on BACK. If you have any questions please raise your hand.
When you are ready, please press BEGIN.
BACK
BEGIN

Which of the following two options do you prefer?
OPTION A:
Certainty
to receive 25 €.
A
OPTION B:
5 % chance
89 % chance
6 % chance.
to receive nothing.
to receive 25 €.
to receive 125 €.
B
OK

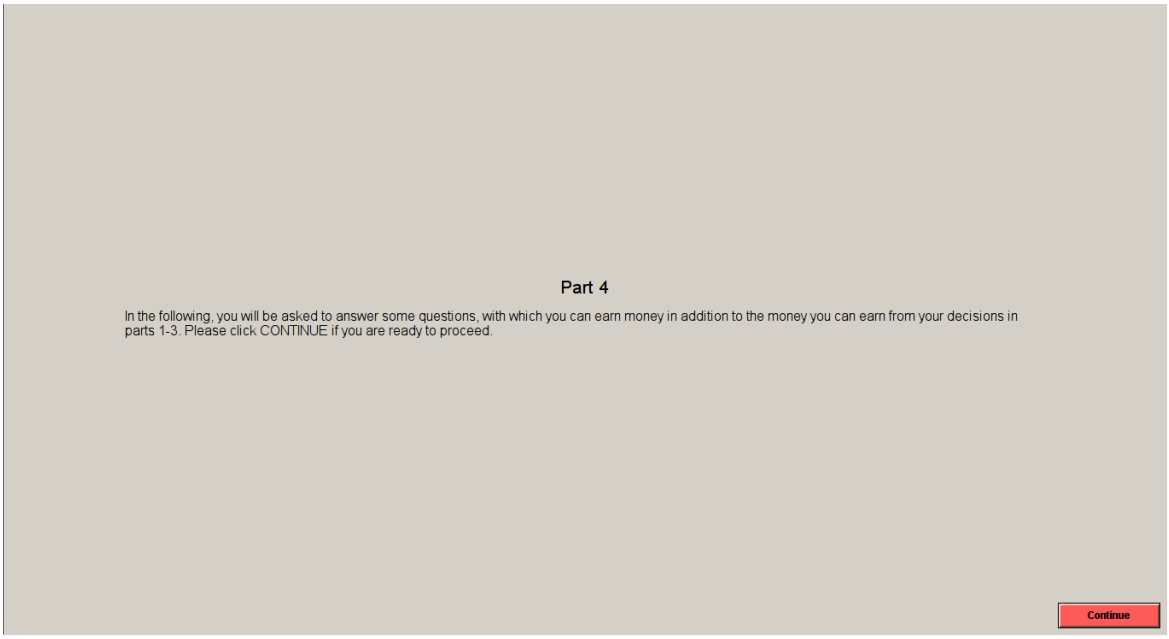
Which of the following two options do you prefer?
OPTION A:
89 % chance
11 % chance
to receive nothing
to receive 25 €.
A
OPTION B:
94 % chance
6 % chance
to receive nothing
to receive 125 €.
B
OK

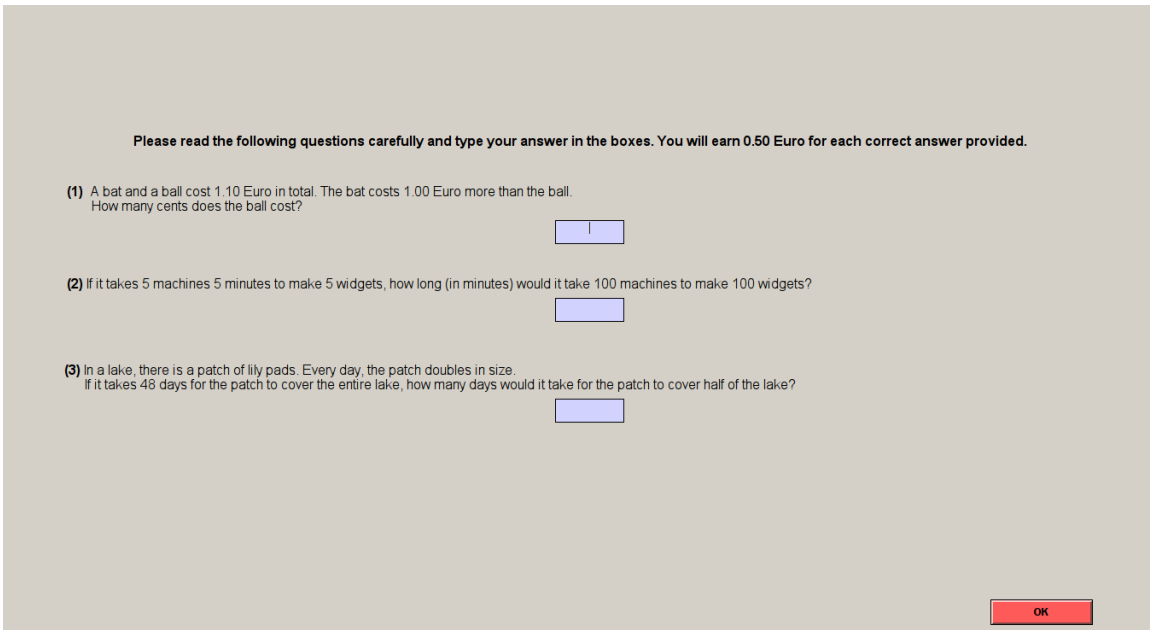

Which of the following two options do you prefer?
OPTION A:
Certainty
to receive 25 €.
A
OPTION B:
1 % chance
89 % chance
10 % chance
to receive nothing.
to receive 25 €.
to receive 125 €.
B
OK

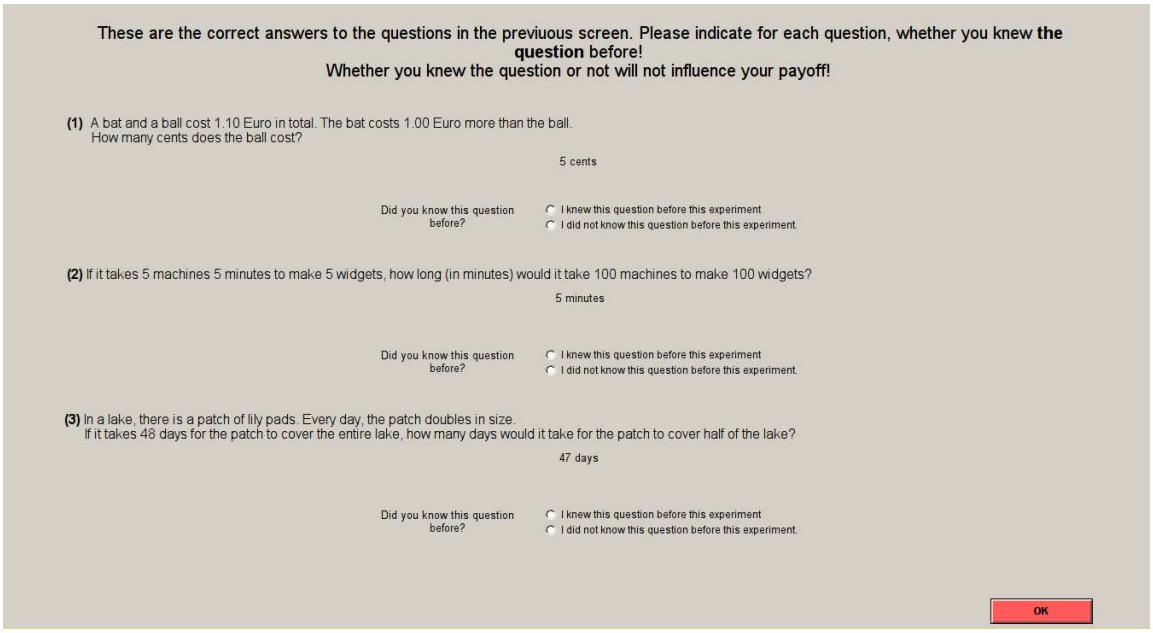
Which of the following two options do you prefer?
OPTION A:
89 % chance
11 % chance
to receive nothing
to receive 25 €.
A
OPTION B:
90 % chance
10 % chance
to receive nothing
to receive 125 €.
B
OK

Which of the following two options do you prefer?
OPTION A:
Certainty
to receive 25 €.
A
OPTION B:
0,5 % chance
89 % chance
10,5 % chance
to receive nothing.
to receive 25 €.
to receive 125 €.
B
OK

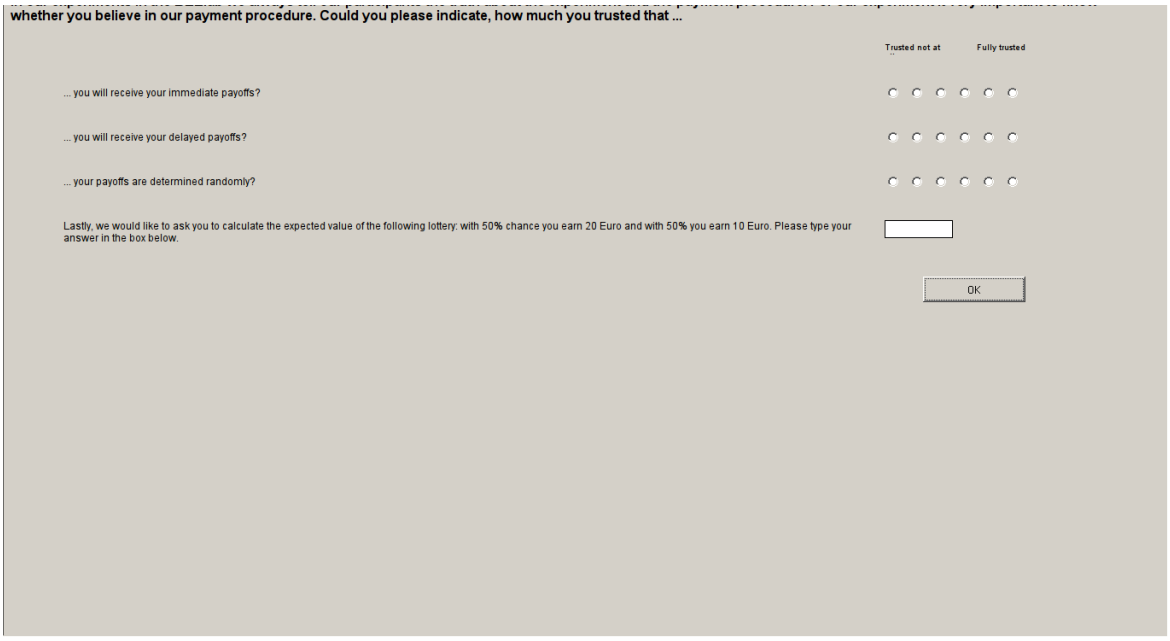
Which of the following two options do you prefer?
OPTION A:
89 % chance
11 % chance
to receive nothing
to receive 25 €.
A
OPTION B:
89,5 % chance
10,5 % chance
to receive nothing
to receive 125 €.
B
OK


Part 4
In the following, you will be asked to answer some questions, with which you can earn money in addition to the money you can earn from your decisions in parts 1-3. Please click CONTINUE if you are ready to proceed.
Continue

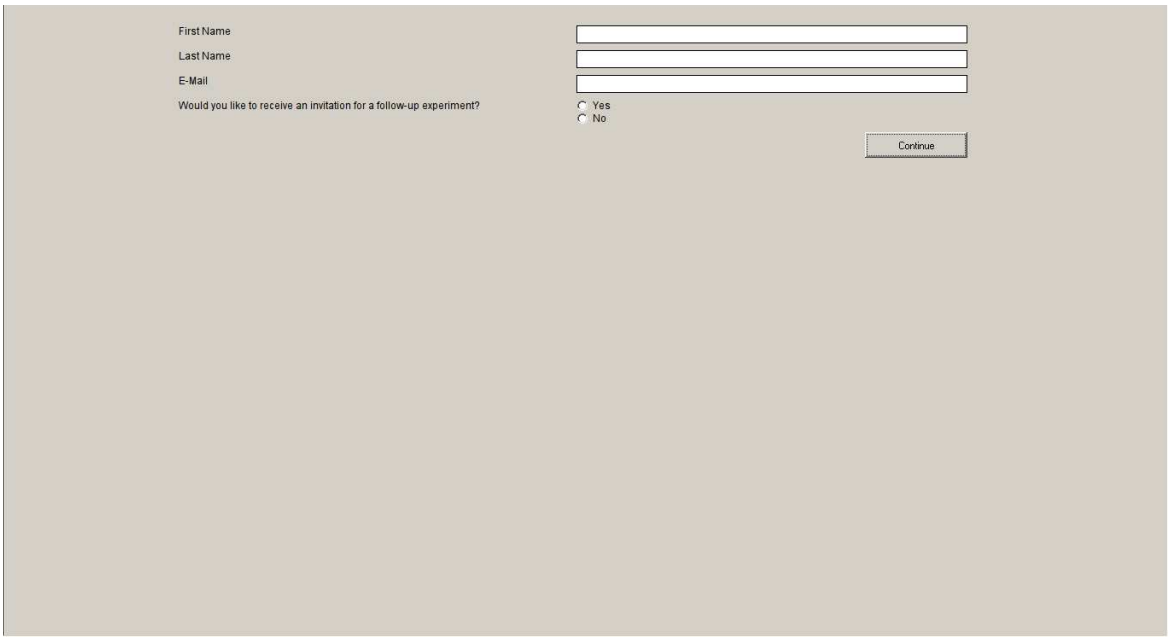


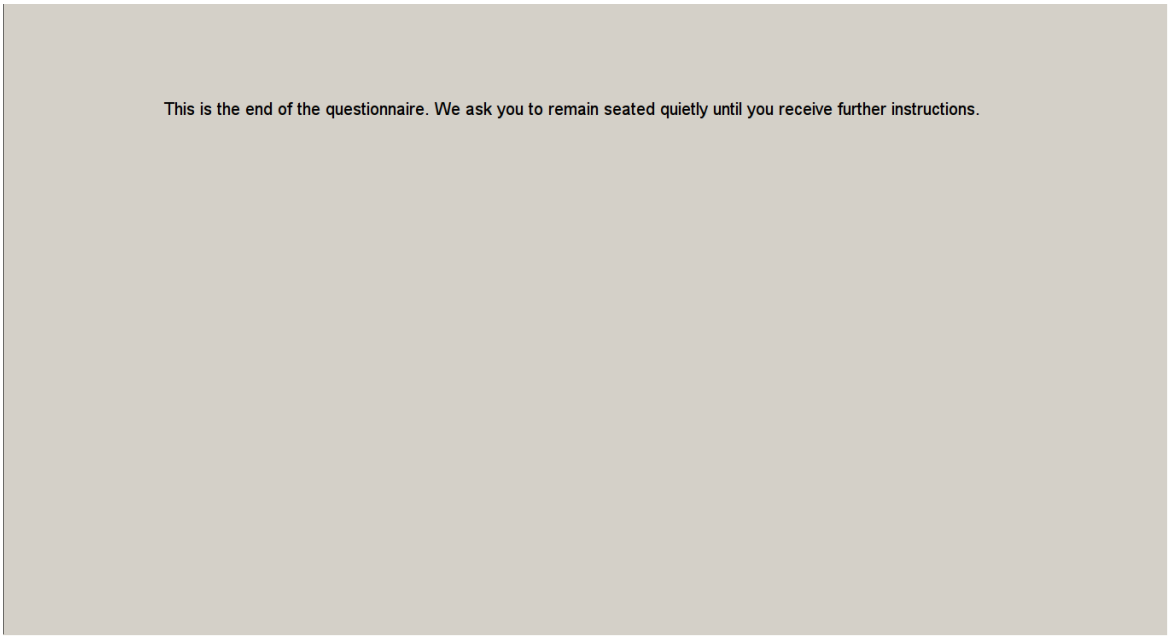


At this point subjects first answered the "Barratt Impulsiveness" questionnaire (see Table B.1 in Appendix for the list of questions) followed by the "Attitudes Towards Risk" questionnaire (see Table 3 in the main text for the list of questions).


whether you believe in our payment procedure. Could you please indicate, how much you trusted that ...
Trusted not at all
Fully trusted
... you will receive your immediate payoffs?
... you will receive your delayed payoffs?
... your payoffs are determined randomly?
Lastly, we would like to ask you to calculate the expected value of the following lottery: with 50% chance you earn 20 Euro and with 50% you earn 10 Euro. Please type your answer in the box below.
OK

Gender
male
female
Age
Which subject do you study?
OK


First Name
Last Name
E-Mail
Would you like to receive an invitation for a follow-up experiment?
Yes
No
Continue


This is the end of the questionnaire. We ask you to remain seated quietly until you receive further instructions.

Your Earnings

The decision situation that was selected for payout by the random mechanism asked you to decide between varying sure amounts and a lottery of 90 € with probability 10 % and 0 € with probability 90 %.
In the selected line you chose for the sure amount of **4.50 €** .

. You will be payed this amount.

In addition, your earnings for correctly answered questions are: **0.00 €**
(0,50 € per correct answer)

In addition, you receive a showup fee of **3 €.**

Your total earnings are **7.50 €.**

We will now carry out the payments individually.

Before the payment we would like to make a digital photocopy of your right hand. We will use it to measure the length of your fingers, which is indicative of prenatal testosterone levels. Your anonymity is guaranteed.

Please remain seated quietly until you are called for payment. When you are called for payment, please bring along the card with your number and the pens.

Thank you very much for your participation.