Supplemental Appendix to Risk Preferences and Field Behavior: The Relevance of Higher-Order Risk Preferences

By Sebastian O. Schneider & Matthias Sutter

A Online Appendix: Method and Implementation

A.1 Measurement of Time Preferences

To measure time preferences, participants repeatedly faced a two-option decision task. One option consists of a fixed amount on the day of the experiment, and the other option consists of a larger amount with a three-week delay that varies from decision to decision. Depending on the choice, the amount paid with a three-week delay is either increased or decreased, and the decision is repeated. We iterate this step four times to approximate the future equivalent, i.e., the value making participants indifferent between the (fixed) earlier payment and the future payment, as the mid-point of the interval in which it must lie (see Appendix A.3 below for an illustration). As our measure of time preference (i.e., impatience), we compute the ratio of the (approximated) future equivalent to the earlier payoff.

A.2 Measurement of Cognitive Abilities

Concerning the measures for cognitive abilities, we focus on fluid intelligence. Our first task, a commonly used matrix test, aims at reasoning, while our second task, a symbol-digit-correspondence task (Dohmen et al., 2010), aims at processing speed. For the first task, participants had five minutes (300 seconds) to complete ten test items, whereas for the second task subjects were given 90 seconds to complete as many symbol-digit-correspondences as possible. We compute a single measure of cognitive ability from these tasks by weighting the successfully completed items in each task, rescaled to be on comparable scales, with the time given for a task, i.e., (the number of matrices solved * 300 + the (rescaled) number of correct symbol-digit pairs * 90)/(300 + 90). Finally, for comparison reasons, we center and standardize this measure.

A.3 Illustrating the Bisection Approach

This example illustrates the bisection approach for elicitation of certainty equivalents. The same procedure applies to the elicitation of time preferences (future equivalents), where the "sure amount" S_i, $i = 1, \ldots, 3$, from the risk task now corresponds to the future equivalent, which can be computed approximately by replacing the parameters L_1 in Table A-1 with 100 (immediate payment, the lowest possible future equivalent), and H_1 with 140 (highest possible future equivalent). Participants then decide between the future equivalent, which is varied from decision to decision, and the immediate payment. The immediate payment stays 100 Taler throughout the four iterations.

Table A-1: Illustration of the Bisection Approach

		Lottery Outcomes		
Iteration	Sure Amount	Low	High	Choice
1	$S_1 = L_1 + (H_1 - L_1)/2 = 70$	$L_1 : 0$	$H_1 : 140$	Lottery
2	$S_2 = S_1 + (H_1 - L_1)/4 = 105$	$L_1 : 0$	$H_1 : 140$	Sure Amount
3	$S_3 = S_2 - (H_1 - L_1)/8 = 87.5$	$L_1 : 0$	$H_1 : 140$	Sure Amount
Result	$CE = S_3 - (H_1 - L_1)/16 = 78.75$			

Legend: S_i denotes the sure amount, and H_i and L_i denote the high and low outcomes of the lottery in iteration $i = 1, \ldots, 3$. CE is the resulting certainty equivalent. See Section I in the main text for details.

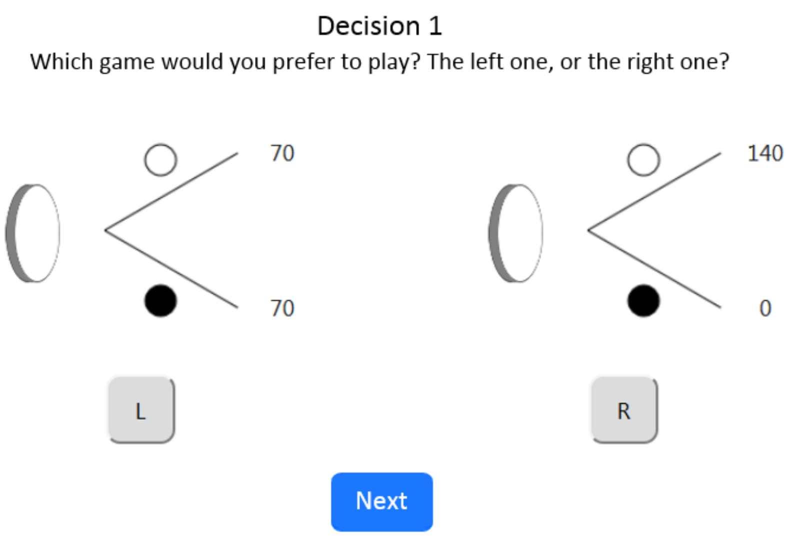


Figure A-1: Elicitation of (Higher-Order) Risk Preferences via Certainty Equivalents: Exemplary Decision Screen with an Animated Rotating Coin With a Black and a White Side

A.4 Choice Behavior in the Certainty Equivalent Tasks

Across the three bisection iterations used for elicitation of a certainty equivalent, participants who are indifferent between receiving the initially offered amount and the lottery might want to increase their expected payoff by choosing "lottery - safe - safe" ("gambling path") instead of "safe - lottery - lottery", which ultimately results in the same difference to their true certainty equivalent. We find no evidence that this has happened systematically. Aggregated over all lotteries, and robust to analyzing grades in isolation, students chose the safe option significantly more often in the first iteration compared to the second or third iteration, even if we control for the number of the lottery played to capture possible learning effects. At the individual level, we can analyze choice paths. The share of individuals choosing the "gambling path" 3 or more times is just as high as the share of individuals choosing the path "safe - lottery - safe" 3 or more times, which, just as the "gambling path", consists of two safe choices. For elicitation of the last certainty equivalent, we observe that the "gambling path" is taken significantly less often than a uniform distribution would suggest (t-test, p-value=0.079), and that for the last two certainty equivalents, when possible learning could be assumed completed, this path is pursued significantly less often than for the first two certainty equivalents (t-test, p-value<0.001). In fact, after elicitation of the first certainty equivalent, the path is chosen less often with every additional certainty equivalent that is elicited (Cuzick non-parametric trend test, p-value<0.001). Only one of the 658 individuals chose the gambling path throughout all six certainty equivalents (compared to 28 who always choose the safe path).

A.5 Details on the Adapted Spline-Regression Approach to Predict Utility Functions (Schneider et al., 2025)

Spline regression can be seen as a generalized interpolation approach, where the fitted curve does not need to meet every point exactly, but where some degree of smoothing is allowed. Compared to ordinary least squares regression or polynomial regression, the key idea of spline regression is to overcome over- and underfitting by regressing on so-called basis functions (e.g., $x, x^2, x^3, \ldots$) that are each locally defined, i.e., only on a sub-interval of the function's support. That is: Instead of regressing on only one linear, squared, cubic, ... term of wealth $(x, x^2, x^3, \ldots)$ on the whole interval from 0 to 140 Taler as in a polynomial regression, spline regression regresses on several of these terms that are only piece-wise defined (and zero otherwise). In particular, here, we use so-called B-splines as basis functions.[17] B-splines of degree s can be computed by means of differences of truncated power functions of the form

$$f(x|b,s) = (x-b)^s_+ = (x-b)^s \mathbb{1}(x>b).$$

Quadratic B-splines, for example, can be computed as $f(x|b_1,2) - 3f(x|b_2,2) + 3f(x|b_3,2) - f(x|b_4,2)$, for some $b_1 < b_2 < b_3 < b_4$ with $b_n - b_{n-1} = b_{n-1} - b_{n-2}$ for $n = 3, 4$. In order to allow for a non-linear fourth derivative of the utility function, which characterizes temperance, we use truncated power functions, and consequently B-splines, of degree $s = 6$.

Using local basis functions allows us to obtain a different functional form in one part of the interval than in the other(s). Recalling a typical utility function illustrates the necessity for this: Some parts might be rather flat, whereas in other parts, marginal utility is high, and correspondingly, utility is increasing rather steeply (see, for example, the exemplary estimated utility function depicted in Figure A-2). In fact, this is what higher-order risk preferences are all about: To characterize the increase of the curvature of marginal utility (and increase(s) thereof). Increasing the order of polynomials while sticking to globally

[17]B-spline basis functions look like Gaussian functions, see Figure 2 in Schneider et al. (2025) for an illustration.

defined linear, squared, cubic, ... basis functions as in polynomial regression may result in a terrible fit to the data due to over- and underfitting, as Runge (1901) pointed out.

Penalized spline regression now solves the issue of subjectively deciding about the number of sub-intervals as well as the position of their boundaries b_i. The idea is to use an abundant number of piece-wise defined local B-spline basis functions that are placed at equally spaced boundaries b_i of the sub-intervals. Overfitting is then tackled in a data-driven way by using a penalization term λP to balance the trade-off between smoothness and fit to the data, such that just the right amount of flexibility is used. Specifically, $\lambda \geq 0$ is the penalty weight, and P is a measure of the "roughness" – the variation of neighboring regression coefficients. For example, roughness can be measured using the variation of coefficients β_i of directly adjacent basis functions as $\sum_{i>2}(\beta_i - \beta_{i-1})^2$. The penalty is simply added to the objective function (loss function) that is to be minimized,

$$S = ||y - X\beta|| + \lambda P,$$

where the first part, $||y - X\beta||$ is the objective function of ordinary least squares regression, with the design matrix X consisting of the B-spline basis functions (evaluated at the measurement points), β the vector of coefficients, and y the vector of the response variable at the measurement points. In that way, the coefficients of two adjacent basis functions will only differ if it favors the fidelity/smoothness-trade-off controlled by λ (see below for how λ is determined).

The penalty In practice, penalties on directly neighboring coefficients are less suited, as the resulting functions would approach a horizontal line already for a relatively low amount of smoothness. Instead, higher-order differences are used. In our case, a penalty on variation of "second-order neighbors" might have unwanted consequences, too: The resulting function would still be biased towards a linear function too strongly, which could potentially result in the disappearance of the fourth derivative for already low degrees of smoothness. A combination of penalties on the variation of third- and fourth-order neighbors works well (as simulations on bias and measurement error discussed in Schneider et al., 2025, and below in Appendix B.2 underline). Specifically, we set

$$P = \eta P_3 + P_4,$$

where $P_3 = \sum_{i>3} \beta_i - 3\beta_{i-1} + 3\beta_{i-2} - \beta_{i-3}$, $P_4 = \sum_{i>4} \beta_i - 4\beta_{i-1} + 6\beta_{i-2} - 4\beta_{i-3} + \beta_{i-4}$, and $\eta \approx P_4/(5P_3)$. The idea is to govern the smoothing process by P_4, and add P_3 with a lower weight to avoid "wiggles" in the third and lower-order derivatives: While P_4 not only smoothes the utility function, but also the fourth derivative, the third and lower-order derivatives might in some cases contain unwanted fluctuations at the boundaries that are not explained by the data. P_3 explicitly addresses these, and the combination of a penalty on an uneven order with a penalty on an even order is also able to address potentially remaining "wiggles" in the lower-order derivatives. Thus, P_3 receives a data-dependent weight P_4/P_3 resulting from a first initial fit where penalization can already be expected to be governed by P_4. The multiplication with the factor 1/5 finally ensures that $\eta P_3 < P_4$, so that P_4 governs the smoothing process. See Schneider et al. (2025) for a more detailed discussion including similar approaches in the literature.[18]

Choice of the Penalty Weight $\boldsymbol{\lambda}$ Finally, λ, the weight of the penalty term, i.e., the solution to the trade-off between fit to the data and smoothness illustrated in Figure 2 in the main text, is determined by optimizing the function's predictive quality via cross-validation. To operationalize this, several choices are possible (see Schneider et al., 2025, for a detailed discussion and approaches taken in the literature). We opt for a leave-k-out version to minimize overfitting, which is equivalent to avoiding "giving too much weight" to noise, following our interpretation of noise in the main text (variation that deviates from the pattern suggested by all other points). Specifically, motivated by response error as explained in the main text, we choose $k = 2$, and exhaustively use all sub-sets of inner utility points excluding 2 points to fit the curve, predict the points left out, and compute the average mean squared prediction error for a given value of λ. Minimization is finally performed by choosing the λ that minimizes this mean squared prediction error.

Further parameters In this setup, using 20 sub-intervals (with corresponding number of B-spline basis functions) works well (as again the simulations on bias and measurement error in Schneider et al., 2025, and below in Appendix B.2 underline). Using 15 sub-intervals does not yield the required flexibility, and using 25 sub-intervals results in overfitting in many cases; see Schneider et al. (2025) for a discussion.

Derivatives and Utility-Based Preference Measures Finally, derivatives can be calculated analytically with a closed-form solution without the need for additional numerical computation (De Boor, 1987, Ch. 10, Equations (12) & (16)). This is another crucial advantage of the spline approach for our purpose. Figure A-2 depicts an exemplary estimated utility function with derivatives. Based on the derivatives of the utility function, attitude measures can be calculated as described in Section II.B in the main text.

[18] Besides penalties P on variation of neighboring coefficients that are controlled by λ, we penalize non-monotonic variation as well as deviations from the origin and from (1,1); for those cases, the weight is just chosen high enough to make sure the conditions are met.

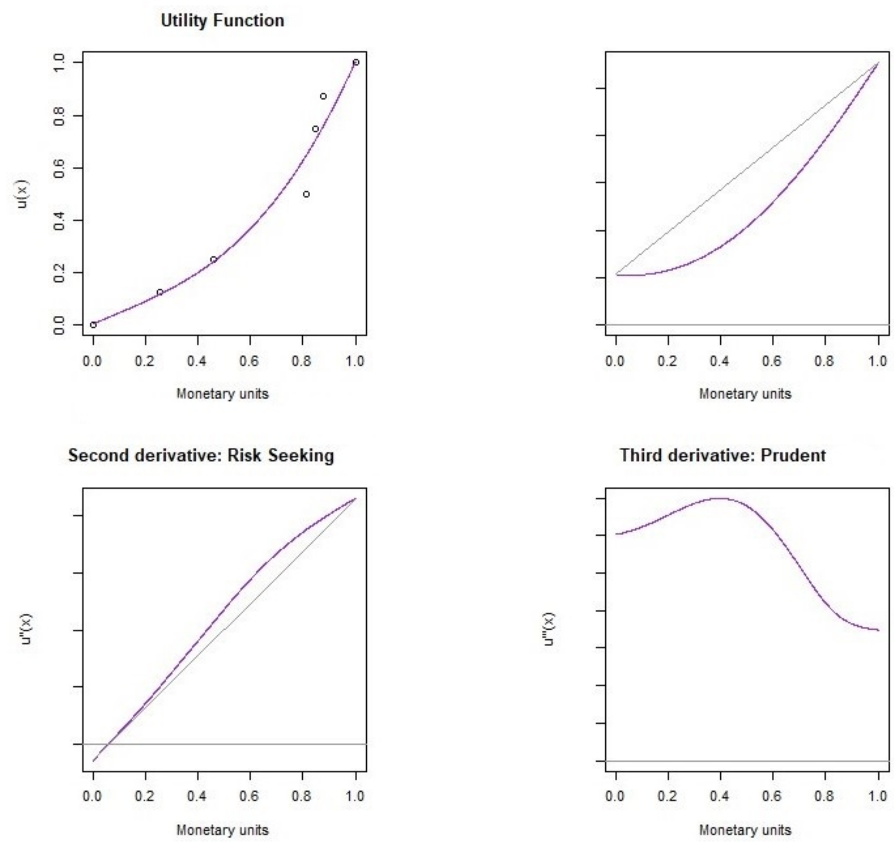


Figure A-2: Estimation of Utility Functions (Including Derivatives) from Utility Points: Example

References

De Boor, C. (1987). *A practical guide to splines* (Applied Mathematical Sciences ed.), Volume 27. New York, NY: Springer.

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

Runge, C. (1901). Über empirische Funktionen und die Interpolation zwischen äquidistanten Ordinaten. *Zeitschrift für Mathematik und Physik 46*, 224–243.

Schneider, S. O., M. Ibanez, and G. Riener (2025). Measuring utility – An application to higher-order risk. Working Paper, MPI Bonn.

B Online Appendix: Considering Challenges to Identification

B.1 Overview

Here, we explore whether alternative explanations could rationalize the relationships that we find between higher-order risk preferences and field behavior. In the main text, we have seen that our sophisticated method increases predictive quality compared to the raw measures, but does not lead to different qualitative results. We therefore focus on the sophisticated measures in the empirical investigations of this appendix. We start with an overview in this first subsection, including a discussion of our results in light of Rabin's (2000) calibration theorem. In the following subsections, we investigate (and rule out) several alternative explanations for our results summarized in this overview in more detail, and finally provide a formal rationalization (confirmed by simulations) of our results regarding the relation between survey measures and field behavior in the last subsection B.5 of this appendix.

Bias and Measurement Error *Bias in Estimates of (Higher-Order) Risk Preferences.* Could our results be driven by bias in estimates? Here, we refer to Schneider et al. (2025): To rule out biased estimates, they compare the measures from Section II.B as derived from the sophisticated method with analytically computed measures. They assume that the true utility curve is given by the family of concave curves $u_c(x) = -c(x-1)^4 + (1-c)x + c$, and its convex counterpart $v_c(x) = cx^4 + (1-c)x$, where $c \in (0, 1]$ and x, the monetary amounts, are rescaled to the unit interval I.[19] The scalar c controls the curvature: For $c = 0.01$, the function visually cannot be distinguished from a linear function, while for $c = 1$, we observe substantial curvature.[20]

They use 10,000 iterations: 100 for each value of c that they increase in steps of 0.01. They start each iteration by randomly sampling x-values in a way that reflects our elicitation procedure. Then, they evaluate these with u_c and v_c, and use the sophisticated method to fit utility curves and derive the measures of (higher-order) risk preferences. For all three measures (r, p, and t) the correlation with the true measure is above .99 (p-values < 0.001 in all cases). See Schneider et al. (2025) for corresponding figures and a detailed description and discussion of the simulation. From these results, we conclude that bias in our estimates is of no concern in our study. In particular, there is no "substitution" of one order of risk preference for another.

[19] Although parametric functions cannot be used to describe empirically observed combinations of higher-order risk preferences with full accuracy (see the discussion in Noussair et al., 2014), they can of course still be used to show that our sophisticated method produces unbiased measures.

[20] This implies an Arrow-Pratt measure corresponding to the 83% quantile of absolute values of r that we observe in our study, and to the 87% quantile when excluding simple patterns of decision-making as discussed below.

Risk Aversion Measured with Measurement Error. Gillen et al. (2019) provide a recent illustration of how taking into account measurement error in experimental variables may cause previously "found" relations between other variables to disappear. To account for measurement error, Gillen et al. (2019) propose to follow suggestions from the old literature on errors in variables (see, e.g., Durbin, 1954): Combining repeated measurement of a characteristic with an instrumental variable approach by instrumenting the repeated measures with each other. They (re-)label this approach *Obviously Related Instrumental Variables (ORIV)*. We apply this technique with our two measures r and r^{CE} to show that the relation between field behavior and higher-order risk preferences that we document is not just a result of risk aversion measured with measurement error (see Table D-23 in Appendix D).[21] Results are robust, as we observe the same pattern as we see in Table 2 in the main text for all considered domains of behavior.

Measurement Error and Noise. Could it be that we mistake noise or measurement error for prudence and temperance? While we have explained theoretically in Section II why this is not to be expected, we examine this further. Using a simulation design based on our own data, we investigate the impact of "noise" and of interpolation error due to approximating the certainty equivalents on our measures. In addition, Schneider et al. (2025) have used the simulation design outlined above (*"Bias in Estimates"*) to investigate the impact of "noise" and interpolation error.

Irrespective of the simulation design and of how exactly noise and interpolation error is operationalized, the sophisticated method is able to recover the true degree of risk aversion, prudence, and temperance very well in the sense that the correlation with the "true" measures remain high (e.g., above .95 in our "main specification"). Moreover, as theoretically derived, arguably exaggerated levels of noise induce a mild downward bias of less than 0.04 SD on our sophisticated measures. For our raw measures, we observe a mild noise-induced upward bias. Hence, our two sets of measures react in opposite ways to noise, allowing for a "sandwich approach" to address the influence of noise on our results. Appendix B.2 discusses the simulations in detail, including further results, e.g., on correlation in noise-induced measurement error. Given the simulation results and the fact that results are the same with our two sets of measures, we can dismiss the worry of measurement error or noise driving our results as unsubstantial, following the idea of a sandwich estimator.

[21] Simulations presented in Appendix B.2 show that the correlation in measurement error between r and r^{CE} is small enough to eliminate more than 98% of measurement error in each measure using this instrumental variable approach to address measurement error.

Simple Patterns of Decision-Making Could it be that simple patterns of decision-making in our tasks are being wrongly interpreted as prudence and temperance? To examine this, we exclude observations that show patterns of constant switching between the options, or of always choosing the safe amount or always choosing the lottery (or, in other words, a "corner solution"). In addition, we also exclude subjects that show a more general tendency to simplified decision-making, that is, those who always choose either the lottery or the safe payment for a given lottery (e.g., first lottery-lottery-lottery, then safe-safe-safe, then lottery-lottery-lottery or safe-safe-safe, and so on). Additionally, coding for example choosing the lottery as 0, and opting for the safe payment as 1, repetitions of possible patterns of length 2 (10, 01, 11, 00), length 3 (001, 010, 011, 100, ...), and length 4 (0001, 0010, 0011, 0100, ...) are excluded.

Choices of 31 observations (about 5% of our sample) match these patterns, with 28 always choosing the safe option and 2 always choosing the lottery. Excluding these observations keeps the distributions of risk aversion, prudence, and temperance in Figure D-2 (Appendix D) fairly similar to before exclusion (Figure D-1 in Appendix D) – with one exception: the peak of the second mode in all three distributions vanishes. Yet, despite discarding about 5% of our sample, the main results about the relation between higher-order risk preferences and field behavior stay the same (see Table D-24 in Appendix D).

Deviations from Expected Utility *Probability Weighting.* What if more sophisticated patterns of decision-making than those just discussed are at play, such as probability weighting? Probability weighting is one of the key modifications of expected-utility theory in e.g., prospect theory. It could potentially matter because numerous studies have shown over- and underweighting for prospects with small or large probabilities. For intermediate probabilities as used in our tasks, however, the empirical importance of probability weighting is less clear (see Appendix B.3 and Table B-2 for an overview). On average, the empirical evidence suggests at best a small influence of probability weighting on our measures.

Nevertheless, non-linear processing of probabilities could theoretically challenge our results in two ways. First, accounting for probability weighting on the individual level in our elicitation technique may affect our estimates of risk aversion, prudence, and temperance. Second, accounting for probability weighting when predicting field behavior by inclusion of an individual's probability weight may affect our results. In Appendix B.3, we discuss both possibilities in detail and show empirically in Tables D-25 and D-26 in Appendix D that these possibilities leave our results practically unaffected.

Reference Dependence. Besides probability weighting, the reference-dependent evaluation of choice options has been proposed to address "anomalies" in decision-making, i.e., behavior that cannot be explained by expected-utility theory (see, e.g., Rabin, 2000). Reference dependence is the crucial difference between rank-dependent models (covered above by addressing probability weighting) and (cumulative) prospect theory (Tversky and Kahneman, 1992), for example.

As there is no (standard) theory of reference points (Wakker, 2010), we consider various possibilities for their location, following Baillon et al. (2020). They empirically study which (broad) reference-point rules are used in decision-making under risk. The considered rules are: Status Quo/initial wealth, "MaxMin", "MinMax", "x at max P", the expected value of a lottery, and the prospect itself (à la Kőszegi and Rabin, 2006).[22] In Appendix B.4, we apply each of these rules to our method and our data. Using standard assumptions, we can show that these rules do not explain the choice behavior of our sample in the experiment. We thus conclude that accounting for reference dependence does not affect our main insight that higher-order risk preferences are key when relating risk preferences to field behavior.

Consistency in and Identification of Utility Curvature *Consistency of Utility Across Domains.* Our approach is based on the assumption that an individual's shape of the utility function is (largely) consistent across domains. We think that this assumption is uncontroversial in light of our results, but that it deserves to be spelled out.

Stake Sizes and Identification of Utility Curvature. Rabin's (2000) calibration theorem shows that a utility function with enough curvature to explain commonly observed small stakes decisions (that all involve losses) cannot plausibly explain larger scale decisions, implying that utility is close to linear for small stakes. In this light – how can our measures plausibly be related to field behavior – empirically, and in line with models based on expected-utility theory?

Our first argument relates to stake sizes: In absolute terms, the gambles we apply involve what is often referred to as small stakes, that is, negligible compared to "final wealth" – of (rich) adults. Yet, our experimental participants still go to school, 99% of them are between 10 and 18.5 years old, and they have a more diverse background than the common university student pool. Without any doubt, money has a different value for them than the same amounts for adults. Final payments in Euro were adjusted to correspond to 120% of the (suggested) weekly pocket money in any given age cohort (see "General Experimental Setup" in Section I.A). Baillon et al. (2020), for example, denote their experiment with payments up

[22]MaxMin refers to comparing the lowest outcomes of the given options, and taking the highest one as a reference point (i.e., the maximum of the minimal outcomes). MinMax refers to comparing the highest outcomes of all options, and taking the lowest one as a reference point (i.e., the minimum of the maximal outcomes). x at max P refers to the payoff with the highest probability.

to a week's salary as "high-stakes experiment". We don't want to argue that our experiment is a high-stakes experiment, but given the still limited accumulated funds of teenagers, our stakes are certainly substantial for them. Therefore, we find it plausible that our stakes are large enough to observe non-linearity in utility when assessing these amounts, and that this curvature is informative for the field behavior under study.

The second point is related: Our measures – or higher-order risk preferences more generally – do not "need" much curvature (in fact, on average, the curvature in our study is close to risk neutrality). Instead, the concept is about the *change* in this curvature, and for that, a non-linear function suffices. If we think about the risk-apportionment tasks for prudence classification due to Eeckhoudt and Schlesinger (2006), even someone who does not strongly object to receiving a mean-zero risk in isolation (compared to receiving just nothing) might still prefer to receive the risk in a state of higher wealth. Intuitively, being the type of person who wants to be prepared to handle a risk does not require strong objection to risk. This is corroborated by studies relying on the risk-apportionment tasks (e.g., Noussair et al., 2014), where prudence is prevalent even absent of risk aversion.

While we find that these arguments can plausibly explain our non-zero measures that relate to field behavior, they open an interesting question: What to expect with small stakes? Given the empirical support for the assumption about consistent shapes of utility functions across domains that we have discussed above, we expect a somewhat related consistency here: With lower stakes than the ones used here, our approach might identify the curvature of, say, a "gain-loss utility" function with the features of prospect theory's value function that governs small-scale decision-making. Such a function is used, for example, in the framework of Kőszegi and Rabin (2006), and can be seen as a remedy to Rabin's paradox. In such a framework, we would expect that the *change* of curvature of this "gain-loss utility" is predictive for that of an individual's consumption utility which describes decision-making for larger scales – just as the identified utility in our setting is predictive of utility in domains with outcomes that differ from monetary ones.

In the following, we provide further details on the considerations and results highlighted in this overview.

B.2 Discussing (Correlated) Measurement Error and Noise as Alternative Explanations for our Results

To match characteristics of the real data regarding the correlation structure of measurement error, we consider a simulation setting based on our true data. In particular, we assume that the estimated, smooth spline functions are the true utility functions. Hence, we sample these smooth individual utility functions from our data, and use the corresponding measures as the true measures. That way, with the undistorted (assumed true) data, we achieve the actual correlations between the different orders of risk preferences that we see in our data (also see Figure B-2b below).

We then follow our elicitation procedure to elicit certainty equivalents according to these assumed true utility functions, and add noise, where we take into account noisiness in experimental participants' choices and the (linear interpolation) approximation of the certainty equivalents from the bisection approach. In particular, of the possible certainty equivalents for a lottery, we always choose the certainty equivalent option that comes closest to the true preferences as implied by the sampled true utility function, but we "distort" the computation by normally distributed noise. Even when there is no distortion, there will be interpolation error, as the bisection approach will typically not offer a certainty equivalent option that lies directly on the smoothened spline function (the assumed true utility function).

While there is no need to calibrate interpolation error, as it just follows from the procedure, we have to make assumptions regarding the decision error.[23] We add a vector of i.i.d. noise to all data points.[24] The noise is drawn from a normal distribution as the standard choice for modeling errors, and calibrated based on the length of a decision "bracket" in our procedure, i.e., the distance between two neighboring data points.

This distance between two neighboring certainty equivalents differs, as the differences between highest and lowest outcome of our lotteries differ. It is largest for the first lottery for which we elicit a certainty equivalent, and it is smallest for the very last (6th) lottery. We take the average length of a decision bracket for elicitation of the second and the third certainty equivalent. This corresponds to 8.75 "Taler" (or 0.0625 when rescaled to the unit interval [0,1]).[25] We hence pick the variance of our normally distributed error such that with a 99% probability, it lies within this range. If the error exceeds this range, we truncate it accordingly.

The idea behind our calibration is that if the true utility function would imply a certainty equivalent that lies right in the middle of two possible certainty equivalent choices, then the error term reaches both boarders, effectively allowing both possible choices. On the other hand, if the true utility function would imply exactly one of the possible certainty equivalent

[23]Adding a uniformly distributed error to each data point instead of using the (characteristic) interpolation error inherent in our elicitation procedure (bisection approach) leaves our results qualitatively unaffected and leads to the same conclusions below. Specifically, we added a uniformly distributed error that is symmetrical around the elicited data points (i.e., the linearly approximated certainty equivalents calculated as the mid-points of the elicited interval). These errors are calibrated such that both bounds of the elicited certainty equivalent intervals may be reached, or the next closest data point, whichever is closest.

[24]Adding noise to only one data point leads to the same qualitative results below, and leaves our conclusions unaffected. The same holds true for adding a vector of mean zero noise to all data points, as well as a vector of the same noise to all data points.

[25]Doubling this amount leads to the same qualitative results below, and leaves our conclusions unaffected.

choices, the error still can be large enough so that the smaller or larger choice can be made. For the first lottery equivalent, the error is not large enough to move away from "a perfect match" in terms of certainty equivalent option, while already for the fourth (and the following) certainty equivalents to elicit, it might well include more than three certainty equivalent options.[26]

That way, on average, we distort the certainty equivalents of any given individual by more than 14 Taler in sum. To put that in perspective: The median distance between any two certainty equivalents in our data is about 16.6 Taler. That is, in total we distort an individual's certainty equivalents, on average, by almost a median distance between two points. If we consider the distortion per point, we distort by about 15% of the median distance per data point for any data point and any individual on average. We think that this degree of distortion exceeds what can plausibly be expected (but note that we obtain the same qualitative results leading to the same conclusions if we increased the error even further, as noted above).

Results

Despite the arguably large errors that we assume, our sophisticated method is able to recover the true data generating process almost perfectly, and we obtain a correlation coefficient of .99 between the "noisy" measure of risk aversion and the "true" measure ($p < 0.001$). For prudence, we obtain a correlation of .98, and for temperance, the correlation is .96 ($p < 0.001$ in both cases). For the simple measure of risk aversion, the correlation between the noisy measure and the true measure is .99, and for the simple prudence measure it is .85, underlining the effectiveness of error correction inherent in our sophisticated method (where the correlation between noisy and true prudence measure is still .98). From this we already see that measurement error cannot play a large role in our measures, as noise leaves in particular our sophisticated measures largely unaffected.

As discussed theoretically in Section II, our sophisticated measures "react", on average, with a downward bias to noise, whereas the simple estimators "react", on average, with an upward bias. This is confirmed in our simulation (see Table B-1). These biases amount to between 1% and 7% of a SD of the underlying measure in absolute terms, where the highest number corresponds to the simple measure of prudence, and the smallest to the sophisticated measure of risk aversion, yet with reversed signs. Note that in all our regressions, we express

[26]Note that fine-tuning the error to the decision bracket is possible, but that might introduce a correlation between curvature and noise. We do not want to add this by design, and it also seems unrealistic to assume that a risk-averse individual makes more mistakes than a risk-neutral individual. We hence opt for errors that are independent of the data points to which they are added (but note that these errors arguably correspond to the upper end of what we might assume realistically).

Table B-1: The Impact of Noise on Bias and Standard Deviations in Measures of (Higher-Order) Risk Preferences

	Sophisticated estimators			Simple estimators	
	Risk Aversion	Prudence	Temperance	Risk Aversion	Prudence
Bias (in SD (Noisy data))	−0.01	−0.04	−0.03	0.01	0.07
Std. Dev. (Noisy data)/ Std. Dev. (Truth)	0.997	0.976	0.959	1.019	1.151

these measures in their SDs (i.e., we divide them by their standard deviation), and hence, Table B-1 reports the relevant measures to assess a potential impact of measurement error on our results.

The differential impact of noise on our sophisticated and simple measures is also reflected in the SD of the measures. For the sophisticated measures, the standard deviations decrease by the introduction of noise (see the last row of Table B-1), consistent with a (very mild) downward bias. For the simple measures, we see an increase.

Next, we investigate correlation in this arguably small amount of measurement error, amounting to between 1% and 7% (in absolute terms) of the variation in the respective measures on average. While one might worry that these (small) errors in our measures have a correlation coefficient close to 1 because of resulting from the same utility function and the same choices, our simulation shows that this is clearly not the case (see Figure B-1). The correlation in measurement error of the prudence (p) and the other two measures is about .35, while that of the risk aversion (r) and temperance (t) measures is .6. Theoretically, this is so because different decisions and certainty equivalents influence e.g., the curvature (that would capture utility-based risk aversion) and the change in curvature (that would capture utility-based prudence) to different degrees. Hence, it is to be expected that an error in any given point influences our measures of risk aversion and prudence to different degrees. When the share of risk-averse individuals equals the share of risk-seeking individuals, we even observe zero correlation of measurement error between the prudence and the risk aversion measures.

Looking at measurement error between our raw (simple) measures r^{CE} and p^{CE}, we see a different picture compared to the sophisticated measures: While we observe correlation in measurement error also in those measures, it is a magnitude smaller, and, most importantly, it is *negative* (whereas the measures themselves are of course positively correlated, see Figures B-2a and B-2b).

Before we put these figures into context, we discuss the correlation between measurement error in the raw (r^{CE}) and the sophisticated risk aversion measure (r) (correlation coefficient of .13, see Figure B-1) with respect to the usefulness of the instrumental variable approach to address errors in variables (e.g., Durbin, 1954; Gillen et al., 2019). It is certainly reassuring to see that the correlation of about .9 between the raw and the sophisticated measure of risk aversion is hardly affected by noise (.88 without noise, .89 with noise, see Figures B-2a and B-2b). This point notwithstanding, with the arguably low correlation in measurement error between the raw and the sophisticated risk aversion measure (.13, see Figure B-1), using an IV approach, we can expect to eliminate the 98.31% share of measurement error in the risk aversion measures that cannot be explained by the other measure (since with one explanatory variable, we have $R^2 = r^2 = 0.13^2 = 0.0169$, i.e., 1.69% of the variation of measurement error in one measure can be explained by the variation of measurement error in the other measure). We think that this makes our IV results (Table D-23 in Appendix D) very useful, keeping in mind that our assumptions arguably match the upper limit of error that we may expect in reality.[27] Hence, we think that this exercise helps to build trust in the conclusion from our (OR)IV results: Our key results cannot be explained by risk aversion measured with measurement error.

Could our results regarding the importance of prudence and temperance then simply reflect our measures of prudence and temperance capturing measurement error that is correlated with our left-hand side variables, because these measures are similarly affected by noise than our indices of field behavior (once we have removed 98.31% of the (small) measurement error from the risk aversion measure with the IV approach to address measurement errors)? We think that this is implausible in light of our IV results (Table D-23 in Appendix D), and the small impact of noise on our measures to begin with, and we also deem this to be a rather unrealistic scenario of data-generating processes for our indices of field behavior and our experimental measures (essentially both producing noise only, but correlated noise). Yet, our results here also rule out such a case: To see this, consider the degree to which measurement error affects the correlations between our measures of risk preferences: In Figure B-2a we report correlation coefficients of our noisy measures, while in Figure B-2b we report correlation coefficients from our true measures. For the sophisticated measures, we see an increase in the (unadjusted, raw Pearson) correlations of prudence with risk aversion and temperance of 0.02 and 0.03, respectively, or about 4-5%. Hence, even if our left-hand

[27] In all the robustness checks mentioned above, including when doubling the error term, the correlation coefficient in measurement error between the sophisticated and simple measures never exceeds .22, meaning that we can still eliminate more than 95% of measurement error in these measures using this instrumental variable technique to address errors in variables.

side indices would capture risk aversion directly (instead of a correlated behavior) and even if it would be affected by (a perfectly correlated) measurement error, i.e., just as much as our risk aversion measure (which we consider to be an upper bound for the impact of correlated measurement error on our results), noise would affect the *coefficient* by an *increase* of less than 5%. However, the correlation between our raw prudence and risk aversion measure *decreases* by 0.048 (or 9%) from 0.508 to 0.461 by the introduction of noise, and, in the scenario above, would hence affect the coefficient by only 9% - but now it would *decrease*. Even though we could argue that, in absolute terms, these effects are small, ultimately their size does not even matter for our conclusion: As our results are the same with the raw and the sophisticated measures, they cannot be driven by correlated measurement error, as the sign of the correlation in measurement error differs and this would either under- or overestimate the truth. Thus, as the results are the same, the only conclusion can be that correlated measurement error is not driving our results.

Next, could it be that measurement error drives our result (independently of the correlation of the error component in our measures)? As laid out above, while our sophisticated measures react to noise with an underestimation of the truth on average, our raw measures react in exactly the opposite way, mirroring the ranking of the orders (smallest degree of under- and overestimation in risk aversion, largest degree of under- and overestimation in prudence, see Table B-1). Hence, if all our results were driven by measurement error, we would overestimate the importance in higher-order risk preference with one set of measures, but underestimate the importance with the other. As the results are the same with both sets of measures, however, here again, the only conclusion can be that measurement error is not driving our results.

To sum up, not only our measures themselves (raw and sophisticated) react differently to noise, also the correlation structure of measurement error is of different sign in these two sets of measures. Yet, we see the same key results with our raw and our sophisticated measures. Hence, as we have explained, the only conclusion from these simulations is: Measurement error, or correlation in measurement error, cannot drive our results, as it either under- or overestimates the truth. As the results coincide, there is hence no room for measurement error. Since qualitatively, none of these results actually depends on the exact calibration of the noise added, we can also rule out that these results depend on specific characteristics of our simulation design. We would come to the same conclusion if we increased or decreased the amount of noise added, if we relaxed the mean-zero error assumption and added independent errors to all data points, if we added the same amount of error to all data points, if we added all noise to only one data point instead of to several data points, or if we used uniformly distributed errors instead of those that naturally arise from our elicitation procedure (the bisection approach) to model interpolation errors.

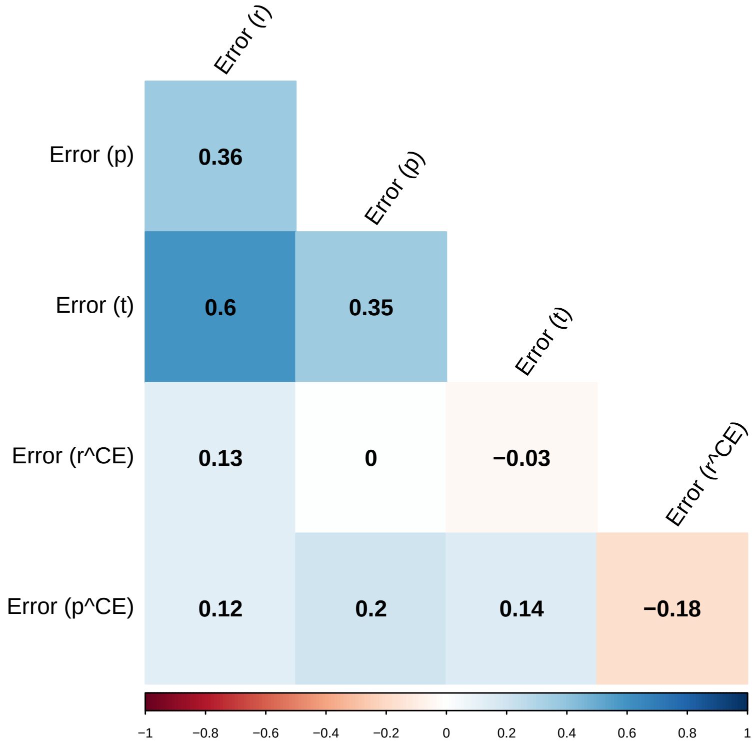
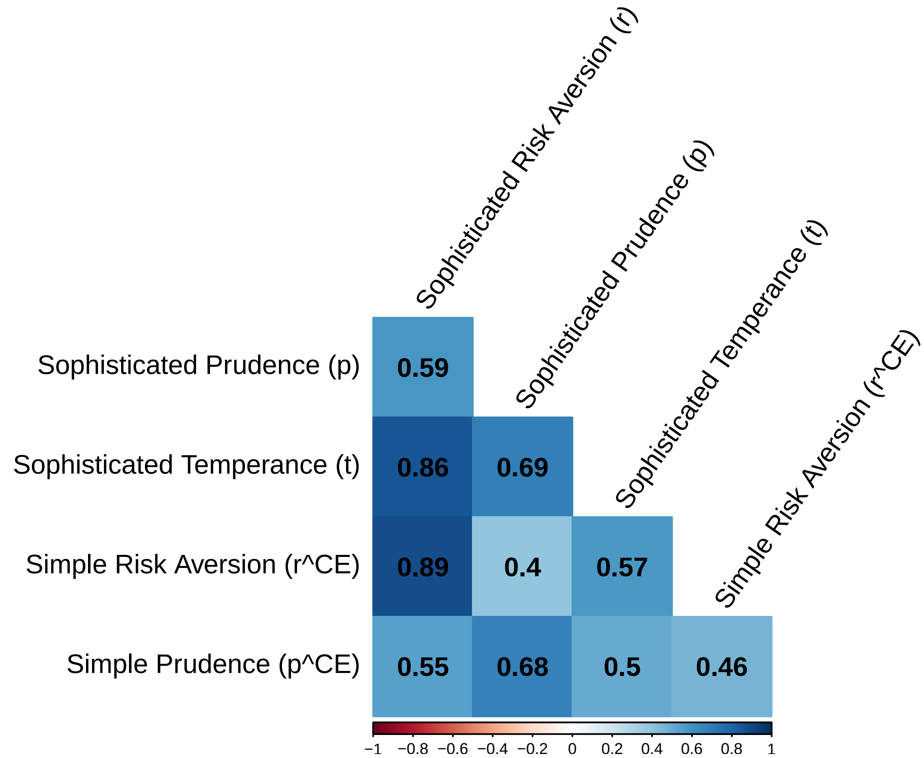
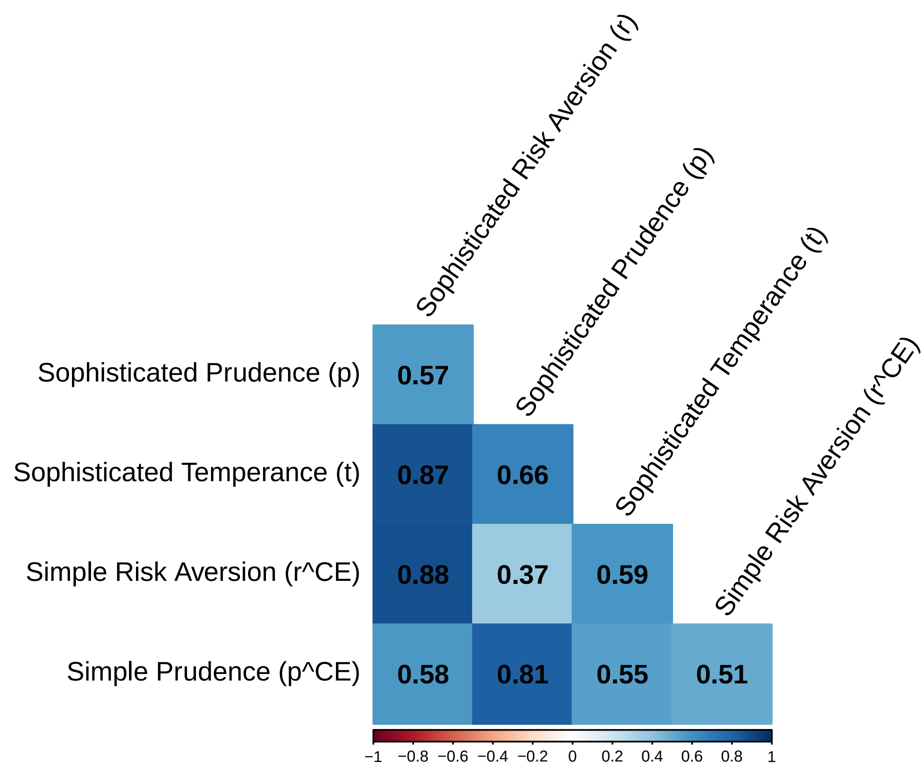


Figure B-1: Correlation in Measurement Error



(a) Correlation in Measures Derived from Noisy Data



(b) Correlation in Measures Derived from Undistorted Data

Figure B-2: The Impact of Noise on the Correlations of (Higher-Order) Risk Preference Measures

B.3 Discussing Probability Weighting as an Alternative Explanation for our Results

In the overview of our robustness tests in Appendix Section B.1, we have already argued that probability weighting is of the highest relevance for prospects with small or large probabilities, but of much less importance for 50:50 gambles as we have used them in our paper. In fact, several studies have documented the absence of probability weighting for such gambles (e.g., Abdellaoui, 2000; Bleichrodt and Pinto, 2000; Abdellaoui et al., 2011). Summarizing results on the weight attached to a 50% probability from 23 parameters of the four most common parametric weighting functions reported in 19 studies over the last 20 years yields a median value of $w(.5) = .483$ or, if only using one weight per study, .481 (where w denotes the probability weighting function; own calculation, see Table B-2 at the end of this subsection).

The first mentioned median value $w(.5) = .483$ is reported by l'Haridon and Vieider (2019), which is probably one of the most comprehensive studies on probability weighting with data from almost 3000 participants across 30 countries. They assume no utility curvature, thus attribute all risk-averse behavior to probability weighting, and report estimates for the 2-parameter probability weighting function proposed by Prelec (1998), which indicate an absence of probability weighting for a probability of .456 such that $w(.456) = .456$. The second mentioned median value, $w(.5) = .481$, is derived from parameters of the 2-parameter probability-weighting function proposed by Goldstein and Einhorn (1987) and Lattimore et al. (1992), reported by Bruhin et al. (2010) on the basis of more than 400 observations. The parameter estimates for the assumed probability weighting function imply the absence of probability-weighting for a probability of .469, such that $w(.469) = .469$.

We now continue to discuss in detail how non-linear processing of probabilities could theoretically challenge our results in two ways. First, accounting for probability weighting in our elicitation technique may affect our estimates of risk aversion, prudence, and temperance. Second, accounting for probability weighting when predicting field behavior by inclusion of an individual's probability weight may affect our results. In the following, we discuss both possibilities and provide empirical results showing that they leave our results unaffected.

Theoretical Considerations It is well known that accounting for probability weighting reduces the curvature of utility (although not necessarily significantly so when using certainty equivalents; see, e.g., Wakker and Deneffe, 1996). Naturally, this decreases measures of utility curvature, such as the Arrow-Pratt measure of risk aversion. A glance at Equation (1) in the main paper shows that this extends to higher-order risk preferences. "Flattening" the utility curve thus reduces the variance in these measures. Consequently, the precision of estimates (measured by their standard error and, consequently, their p-value) linked to these measures when predicting field behavior is reduced *mechanically* without a change in observed behavior, as the precision is proportional to the reciprocal of variance in measurement (less variation, less precision in estimates).[28] That is, accounting for probability weighting in elicitation of measures of utility curvature might yield more plausible estimates for utility curvature in general (although, given the results discussed above, a drastic change is not to be expected for binary equiprobable lotteries as used here). However, this will yield less precisely estimated coefficients of these measures when predicting field behavior. Naturally, the same applies to measures that rely on derivatives of the curvature, as a more linear utility function results in less variation in the derivatives. This however, does not change the way in which participants react differently to different stakes as measured by higher-order risk preferences. Yet, at least theoretically, ceteris paribus it reduces the predictive quality of its measurement.

Empirical Investigation In what follows, we examine whether the theoretical considerations above are of empirical relevance in the setting of our paper. Although maybe obvious, this is a good place to note that a failure in explaining field behavior exclusively by utility curvature when accounting for probability weighting should by no means be interpreted as a failure of risk aversion to explain field behavior: When accounting for probability weighting, risk-averse behavior is described by both parameters, utility curvature, *and* probability weighting; the same applies analogously to prudence and temperance.[29] It is therefore inappropriate to label measures relying on the utility function as risk aversion, prudence, and temperance when accounting for probability weighting – we do so nevertheless to keep the exposition clear.

In Section II, we have described how utility points are derived from the elicitation of certainty equivalents of two-outcome lotteries with equally likely outcomes. The first certainty equivalent, for example, has a utility of 0.5 due to normalization of utility such that the highest outcome has utility of 1, and the lowest has utility of 0, that is, $U(CE_1) = 0.5u(x_{\max}) + 0.5u(x_{\min}) = 0.5$. However, if participants over- or underweigh probabilities of 0.5, then the utility of the first iteration, for example, becomes $U(CE_1) = w(0.5)u(x_{\max}) +$

[28]To keep things simple, consider a linear regression of an overall measure of field behavior with only individual risk aversion r_i as predictor, i.e., $y = \beta_0 + r_i\beta_1 + \varepsilon_i$. The standard error of the estimate $\hat{\beta_1}$, our estimate for risk aversion, is $SE(\hat{\beta_1}) = \sqrt{\hat{\sigma}/\sum_{i=1}^{n}(x_i - \bar{x})^2}$.

[29]To see this, consider a linear utility function, and probability weighting such that $w(.5) = .49$ for the highest outcome of a lottery with two equally likely outcomes 50 and 100. A decision-maker with these preferences will prefer the sure outcome of 75, as $0.5 \times (0.49 \times 100 + 0.51 \times 50) = 0.5 \times (49 + 25.5) < 75$, and will thus exhibit risk-averse behavior, even though their utility function is linear, indicating – in EUT framework – risk-neutral behavior. Depending on the assumptions on probability weighting, one can fnd examples that illustrate that the curvature of utility is not enough to predict risk-averse or risk-seeking behavior unambiguously when accounting for probability weighting.

$(1 - w(0.5))u(x_{\min}) = w(0.5)$ under rank-dependent probability weighting as incorporated in cumulative prospect theory (Tversky and Kahneman, 1992) or rank-dependent utility theory (e.g., Quiggin, 1982, for the general case including non-binary outcomes). The calculation of utility in subsequent iterations works analogously.

To determine an individual's level of probability weighting, we assume that $0.4 < w(0.5) < 0.6$, informed by the empirical findings on probability weighting discussed above. Specifically, using the specification and data reported in l'Haridon and Vieider (2019), in the aggregate, there is no probability weighting around probabilities of 45%, i.e., probabilities of around 5 percentage points below the objective probability of 50%, which we use in our lotteries; this seems like a good starting point for a lower boundary (where we take the more conservative number compared to Bruhin et al. (2010), who report absence of probability weighting around probabilities of 47%). To allow for individual heterogeneity, we double this distance, and set $w(.5) = .4$ as lower boundary, which is about six times as much underweighting than on the aggregate level. We think that this is a conservative assumption, given the introductory discussion on probability weighting around 50%, and given that we use binary outcomes only with the same 50:50 probability that we hold constant (much in contrast to most of the surveyed studies in Table B-2, for which reason one can expect a larger impact of probability weighting in these studies compared to our setting). We then proceed by "diverting" the risk aversion that an individual expresses into probability weighting in the following way: For every individual, we compute the relative risk premia r^{CE} (see Section II.A) for the lottery with expected utility 0.5, as well as those corresponding to expected utility of 0.25 and 0.75 in an analogous way. These three relative risk premia range from -0.5 to 0.5, where we take, if all three have the same sign, the one with the lowest absolute value, and otherwise the mean of the three, and "translate" it linearly to the interval from 0.40 to 0.60, such that risk premia of 0 (i.e., risk neutrality) translate in no over-/underweighting, i.e., $w(0.5) = 0.5$. We use the lowest absolute value of these premia, as probability weighting is assumed to be unaffected by absolute changes in outcomes, as long as the ranks stay constant (see, e.g., Tversky and Kahneman, 1992). We thus still attribute changes in behavior over the range of outcomes to higher-order risk preferences – as this is what they are supposed to capture by definition.

The obtained values of probability weighting (mean value for $w(0.5) = 0.483$, std. dev. 0.029, for the full sample) match those reported by l'Haridon and Vieider (2019) and Bruhin et al. (2010) almost exactly, where all risk-averse behavior was attributed to probability weighting or EUT candidates were excluded from the estimation process, respectively (meaning that their estimates can be seen as rather extreme deviation from objective probabilities, i.e., $w(0.5) = 0.5$). For the subsample of risk-averse participants, we obtain $w(0.5) = 0.470$

with a standard deviation of 0.024.

Tables D-25 presents the results when conducting our whole analysis with the utility-based measures defined in Section II after having accounted for probability weighting in the elicitation of utility points. We first note smaller coefficients for our measures in absolute terms compared to those in Table 2, in line with our theoretical considerations above. Yet, the changes in coefficients are moderate, mostly limited to the third digit, once more underlining the moderate role that probability weighting plays in our setting. Regarding the relationship with field behavior, two coefficients, however, are no longer significant, both of them previously significant at the 10% level: the coefficient of risk aversion for short-term prevention, and the coefficient of prudence for saving. Nevertheless, the conclusion is unchanged: For the relationship between risk preferences and field behavior, it matters substantially to account for higher-order risk preferences.

Controlling for Probability Weighting When Predicting Field Behavior Finally, one might wonder whether including a measure for probability weighting in our regressions would affect the predictive power of higher-order risk preferences. Together with other experimental measures for risk attitudes, Charness et al. (2020) investigate the predictive power of probability weighting for field behavior in a representative sample of the Dutch population. They find that probability weighting, elicited as a parameter of the one-parameter Prelec probability weighting function, using the method by Tanaka et al. (2010), has neither any predictive power for the financial decisions performed in the laboratory, nor for the field behavior under study (investment and insurance decisions as well as occupational choice).

While these results suggest that it is unlikely that our results will actually be affected by accounting for probability weighting in our regressions (because, for example, our measure of prudence could measure probability weighting despite having accounted for it in the estimation process), the discussed null findings may of course have many reasons.[30] We thus include the obtained measure of probability weighting as described above in our regressions – in addition to accounting for probability weighting in the estimation process of our utility-based coefficients.

Table D-26 presents the corresponding results when conducting the analysis with the utility-based measures defined in Section II after having accounted for probability weighting in the elicitation of utility points and when controlling for the weight attached to the probability of 50%. The results with respect to utility curvature (i.e., the Arrow-Pratt measure under EUT) and probability weighting should be treated with a bit of caution (see our com-

[30] We also note that Dimmock et al. (2021) have successfully related probability weighting to investment behaviors, i.e., one of the field behavior investigated in Charness et al. (2020).

ment above). Therefore, one should not necessarily understand significant coefficients for the probability weight as evidence in favor of its predictive quality over risk aversion; rather, one should understand significant coefficients for p and t (i.e., the prudence and temperance measures under EUT) as being robust against accounting for probability weighting.

Regarding the measure p (prudence under EUT), no coefficient has lost its significance, and in general, coefficients rather have increased, although only moderately. With respect to the measure t (temperance under EUT), the coefficients for long-term prevention, cautious planning, and saving are still significant, while the coefficient on risky investment is no longer significant. Yet, the conclusion with respect to the role of higher-order risk preferences for the relation of risk preferences to field behavior is unchanged: For the relationship between risk preferences and the domains of field behavior considered here, it matters substantially to account for higher-order risk preferences.

Table B-2: Literature Overview: Probability Weighting

Weighting Function	Study	N	Parameter Values	$w^+(0.5)$	Note
GE-87†	Abdellaoui et al. (2005)	41	$\gamma = 0.832$, $\delta = 0.975$	0.494	Table 9, power approximation
	Fehr-Duda et al. (2006)	181	$\gamma = 0.515$, $\delta = 0.835$	0.455	Table 3
	Stott (2006)	96	$\gamma = 0.96$, $\delta = 1.40$	0.583	Table 5
	Booij et al. (2009)	1935	$\gamma = 0.618$, $\delta = 0.772$	0.436	Section 5.3
	Bruhin et al. (2010)	448	$\gamma = 0.377$, $\delta = 0.926$	0.481	Table 5, CPT type, pooled data
	Murad et al. (2016)	86	$\gamma = 0.477$, $\delta = 0.729$	0.422	Figure 4
TK-92†	Stott (2006)	96	$\gamma = 0.96$	0.499	Table 5
	Harrison and Rutström (2008a)	158	$\gamma = 0.911$	0.497	Table 1
	Rieskamp (2008)	30	$\gamma = 0.77$	0.477	Study 2, Results
	Zeisberger et al. (2010)	86	$\gamma = 0.88$	0.494	Table 4, pooled data
	Brooks et al. (2013)	90	$\gamma = 0.83$	0.488	Table 4
	Drichoutis and Lusk (2016)	100	$\gamma = 0.690$	0.454	Footnote 14, composite model
Prelec I†	Tu (2005)	1743	$\gamma = 1.00$	0.500	Table 4.6
	Stott (2006)	96	$\gamma = 0.94$	0.492	Table 5
	Tanaka et al. (2010)	181	$\gamma = 0.74$	0.467	Section II.C
	Toubia et al. (2013)	125	$\gamma = 0.534$	0.439	Figure 7, DEEP model
Prelec II†	Stott (2006)	96	$\gamma = 1.00$, $\delta = 1.00$	0.500	Table 5
	Booij et al. (2009)	1935	$\gamma = 0.618$, $\delta = 1.052$	0.432	Footnote 11
	van de Kuilen and Wakker (2011)	78	$\gamma = 1.454$, $\delta = 1.578$	0.396	Figure 6
	Abdellaoui et al. (2011)	67	$\gamma = 0.85$, $\delta = 0.93$	0.506	Figure 3, solid line
	Epper et al. (2011)	332	$\gamma = 0.367$, $\delta = 0.838$	0.481	Table 6
	Murphy and ten Brincke (2018)	142	$\gamma = 0.665$, $\delta = 0.860$	0.510	Section 6.1
	l'Haridon and Vieider (2019)	2939	$\gamma = 0.602$, $\delta = 0.908$	0.483	Section 5.1
			Simple median	0.483	
			Median of medians per study	0.481	

Notes: $w^+(0.5)$ denotes the probability weight associated with the probability of 0.5 that is implied by the reported parameters.

† The parameters correspond to the weighting function by Goldstein and Einhorn (1987)/Lattimore et al. (1992), denoted GE-87, with $w(p) = \delta p^\gamma/(\delta p^\gamma + (1-p)^\gamma)$, Tversky and Kahneman (1992), denoted TK-92, $w(p) = p^\gamma/(p^\gamma + (1-p)^\gamma)^{1/\gamma}$, or Prelec (1998), either in its one-parameter form Prelec I, $w(p) = \exp(-(-\ln(p))^\gamma)$, or the two-parameter version Prelec II, $w(p) = \exp(-\delta(-\ln(p))^\gamma)$, respectively.

The reported results are rounded to the third digit, unless provided with a lower precision. If papers report individual data, we use median values here, if reported. If several parameters for aggregate data are reported (e.g., for different experimental conditions or several samples), we take the median of the reported parameters, unless the functions have been fit to pooled data, in which case we take the thus obtained values from the full sample. If, in addition to own data, historical data from other authors has been used for comparison reasons, we omit these. 'Median of medians per study' means that for studies reporting estimates for several weighting functions (e.g., Stott, 2006), we first build the median weight using the weights implied by the different weighting functions, and then build the median across all studies.

B.4 Discussing Reference Dependence as an Alternative Explanation for our Results

In the overview of our robustness tests above, we have introduced the reference-point rules considered by Baillon et al. (2020) as our starting point to investigate the potential role of reference dependence. Recall that they consider the following (broad) reference point rules:

Status Quo/initial wealth, "MaxMin", "MinMax", "x at max P", and the two stochastic reference points (also see Sprenger, 2015), the expected value of a lottery, and the prospect itself (à la Kőszegi and Rabin, 2006). "MaxMin" refers to comparing the lowest outcomes of the given options, and taking the highest of those as a reference point (i.e., the maximum of the minimal outcomes). MinMax in turn refers to comparing the highest outcomes of the given options, and taking the lowest one as a reference point (thus, the minimum of the maximal outcomes). x at max P, finally, refers to the payoff with the highest probability.

Using out-of-sample predictions, Baillon et al. (2020) find that the considered rules reflect the preferences of the sample well, and their results indicate that the Status Quo (initial wealth in our case, see below) and the "MaxMin" reference (which in our tasks corresponds to the amount of the safe payment, see below) are by far the most relevant ones, describing behavior of more than 60% of their sample. Nevertheless, following Baillon et al. (2020), we also consider the other reference-point rules below, and from that we will conclude that none of these rules can explain the behavior of our sample in our experiment; thus, much less can they explain our measures of higher-order risk preferences, as we demonstrate here.

Initial Wealth (Status Quo) as Reference Point In reference-dependent models such as prospect theory (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992), it is common practice to evaluate outcomes against the "status quo" (Wakker, 2010). Empirically, this is the most relevant reference point according to Baillon et al. (2020). In our study, for working without any (announced or displayed) endowment and only with positive outcomes (there is not a single minus sign appearing in any of our tasks), and for realizing payments only after all decisions have been made, individuals who evaluate prospects by comparing them with the status quo are facing gains only, without any endowment being deduced, and they are thus fully described by the approaches discussed so far (i.e., the results in the main text, or those accounting for probability weighting). For this reason, we can conclude that our measures of (higher-order) risk preferences cannot be affected by reference dependence and loss aversion due to evaluating outcomes based on the status quo (as there can be no loss compared to the status quo in our study).

Sure Outcome as Reference Point Instead of using their initial wealth as a reference point, participants could evaluate outcomes in comparison to the offered sure outcome (see, e.g., Wakker, 2010, Example 8.7.1, and the references therein for studies following this approach). In the classification of reference-point rules by Baillon et al. (2020), using the sure outcome as a reference point is consistent with the three reference rules "MaxMin", "MinMax", as well as "x at max P", and empirically, those are the second-most important reference-point rules after the status quo.[31] To study choice behavior described by such preferences, we allow for probability weighting for gains (w^+) and losses (w^-), and follow Baillon et al. (2020) in setting

$$U(x) = \begin{cases} (x - S_i)^\alpha & x \geq S_i \\ -\lambda(-(x - S_i))^\alpha & x < S_i \end{cases} \tag{3}$$

for utility of outcome x, which depends on the sure amount S_i offered in a given decision i, and the loss-aversion parameter $\lambda > 0$, where $\lambda > 1$ corresponds to loss-averse behavior. Utility curvature is determined by α for gains and losses.[32]

Due to the specifics of the bisection approach that we apply for the determination of indifference values (see Appendix A.3 for details on the bisection approach), the loss-aversion parameter which makes participants indifferent (according to (3)) between the lottery and the safe payment in a binary-choice task is the same for any given individual for every first (second, ...) decision to elicit the certainty equivalent of a lottery, respectively (see the end of this section for the calculations). More specifically, in the first binary decision task of a sequence of decisions to elicit the certainty equivalent of a lottery, the suggested safe amount always corresponds to the expected value of the lottery. Any individual described by Equation (3) would in the first task of a lottery sequence to determine a certainty equivalent (independently of searching for the certainty equivalent of the first, second, third, ... lottery) always choose the safe payment (the lottery) if $\lambda > (<) \frac{w^+(.5)}{w^-(.5)}$. In the second binary-decision

[31] In our tasks, the sure outcome by construction is always higher than the minimal outcome, always lower than the maximal outcome, and always the one with the highest probability.

[32] The power (CRRA) utility function is arguably the most widely-used parametric form in combination with reference dependence, and it has repeatedly been found to fit utility curvature best when allowing for a reference point and probability weighting among standard parametric alternatives (e.g., Stott, 2006; Abdellaoui et al., 2007). Applying the same power (CRRA) utility function for the utility of gains and losses is common practice when empirically studying reference-dependent behavior (see, e.g., Harrison and Rutström, 2008b, Post et al., 2008, Tanaka et al., 2010, or Baillon et al., 2020). It is in line with common findings in the literature, failing to reject the hypothesis of equal curvature of gains and losses (e.g., Abdellaoui et al., 2007) or obtaining literally the same parameters when parametrically fitting the power family (e.g., Tversky and Kahneman, 1992). Such a formulation implicitly includes linear utility for $\alpha = 1$ (as applied by, e.g., l'Haridon and Vieider, 2019). These assumptions thus match empirical results very well and can be considered adequate common practice, which we follow. Reflection has, moreover, also been reported in the context of higher-order risk preferences (Bleichrodt and van Bruggen, 2020).

task, a similar pattern applies. Thus, the evaluation of the prospects remains constant for all second binary-choice tasks in all the six sequences used to elicit the six certainty equivalents. Consequently, if the sure-outcome reference-dependent model describes choice behavior in our tasks, decision-makers should consistently choose the same option across all six sequences in the second choice task as well.

Continuing with this logic, we identify 31 individuals who express a choice pattern consistent with such a model. All these participants belong to the group excluded when analyzing simple patterns of decision-making, and in our robustness test overview in Appendix B.1 above, we have summarized that this does not change our results. Given those results we hence conclude that this kind of reference dependence, i.e., evaluating the outcomes against the sure outcome as described in (3), cannot explain our results.

Deterministic Expectation-Based Reference Dependence Although in the study by Baillon et al. (2020) at most 20% of decision-makers behave consistently with an expectation-based reference point (i.e., expected value or the prospect itself), we consider these rules nonetheless. The expectation of a lottery has been used as reference point for example by Deck and Schlesinger (2010): They show that the pattern of behavior that they observe (risk aversion, prudence, and intemperance) in their risk-apportionment tasks with unresolved compound lotteries (Eeckhoudt and Schlesinger, 2006) is consistent with interpreting the mean of the lotteries as the reference point for evaluation of the choices in their study. For this result, they rely on the parametrization reported by Tversky and Kahneman (1992), i.e., a curvature parameter $\alpha = 0.88$ for gains and for losses, and a parameter of .65 for the probability-weighting function due to Tversky and Kahneman (1992). Although our choice tasks are completely different, and in particular do not involve any negative numbers, and although empirically this rule seems to be less relevant (Baillon et al., 2020), we evaluate the possibility that participants evaluate the outcomes against the mean of the lottery for which a certainty equivalent is elicited. To do so, similarly to before, we allow for probability weighting, and follow Baillon et al. (2020) in setting

$$U(x) = \begin{cases} (x - M_i)^\alpha & x \geq M_i \\ -\lambda(-(x - M_i))^\alpha & x < M_i \end{cases} \tag{4}$$

for utility of outcome x, which depends on the mean M_i of the lottery i, and the loss-aversion parameter $\lambda > 0$, where $\lambda > 1$ again corresponds to loss-averse behavior, and utility curvature is again determined by α for gains and losses as in (3).

Also for this reference-point rule, the loss-aversion parameter which makes participants indifferent (according to (4)) between the lottery and the safe payment in a binary-choice task is the same for any given individual for every first (second, ...) decision to elicit the certainty equivalent of a lottery, respectively (see the end of this section for the calculations).

Although the calculations differ, the result is exactly the same, and we thus identify the same 31 individuals who express a choice pattern which could be consistent with such a model. These participants belong to the group excluded when analyzing simple patterns of decision-making, and in our robustness test overview in Appendix B.1 above we have summarized that this does not change our results. Hence, we can also here conclude that this kind of reference dependence, i.e., evaluating the outcomes against the expected value of the lottery as described in (4), cannot explain our results.

Prospect-Specific Expectation-Based Reference Dependence The last category of reference point rules studied by Baillon et al. (2020) consists of prospect-specific reference points. That is: Every outcome is evaluated against the prospect itself – if a prospect has two equally likely outcomes a and b, then with a 50% probability, the reference becomes a, and with 50% probability, the reference becomes b. Probably the most well-known example of this kind is the core of the framework by Kőszegi and Rabin (2006; 2007):[33] Formally,

$$U(F|F) = \int\int u(c|r)\, dF(r)dF(c) = \int\int m(c) + \mu((m(c) - m(r))\, dF(r)dF(c),$$

where F is the probability measure from which consumption c and the reference level of consumption, r, respectively, are drawn, and u is the agent's utility, which can be further decomposed into "consumption utility" m and "gain-loss" utility μ (Kőszegi and Rabin, 2006). We follow Kőszegi and Rabin (2006) in assuming $m(x) = x$, i.e., linear consumption utility, and in setting

$$\mu(x) = \begin{cases} \eta x & x > 0 \\ \eta\lambda x & x \leq 0, \end{cases} \tag{5}$$

where $\eta > 0$ is the weight attached to gain-loss utility, and $\lambda > 0$ is the coefficient of loss aversion.[34] Moreover, following Baillon et al. (2020), we equate outcomes with final wealth, thus omitting initial wealth. Just as for the reference-point rules discussed thus far, calculations differ (see below), but yield the same result. The same 31 individuals as before express a choice pattern which could be consistent with preferences according to (5), and their exclusion from our analysis does not alter our results qualitatively (see the summary in our robustness test overview in Appendix B.1 above).

[33] To be precise, it is the choice-acclimating personal equilibrium (CPE) introduced in Kőszegi and Rabin (2007); the choice-specific version is called preferred personal equilibrium (PPE), is recursively defined, and hence "cannot be estimated" (Baillon et al., 2020).

[34] Kőszegi and Rabin (2006) assume $\lambda > 1$, but we can be more general here and allow for individuals to behave gain-seeking as well.

Sure Outcome as Reference Point – Theoretical Analysis (with Calculations) We now present the theoretical analysis of choice behavior in the elicitation tasks when assuming that the sure outcome is considered as reference point when evaluating the lottery, that is, when utility is given by Equation (3) as discussed above. Recall that for every lottery that we consider, we elicit the corresponding certainty equivalent using three decisions (see Appendix A.3). Thus, the lottery stays the same in three subsequent decision tasks, where the safe amounts differ.

For the theoretical analysis, we always proceed in the same way. First, we derive the loss-aversion parameter that corresponds to indifference between the two choice options. From this, we can derive how decision makers choose, based on their loss-aversion parameter. As it turns out, the so derived decision patterns will be independent of whether we elicit a certainty equivalent for the first, second, ... lottery. That is, they will be the same for all six lotteries.

- First Decision(s): Lottery (outcomes $L_1 = x_{\text{min}} = 0$ and $H_1 = x_{\text{max}} = 140$ with equal probability) vs. safe amount ($S_1 = (L_1 + H_1)/2 = L_1 + (H_1 - L_1)/2$).

  $$\begin{aligned} \text{Indifference} \;&\Leftrightarrow\; w^+(.5)(H_1 - S_1)^\alpha - w^-(.5)\lambda(S_1 - L_1)^\alpha = (S_1 - S_1)^\alpha \\ &\Leftrightarrow\; \lambda = \frac{w^+(.5)(H_1 - S_1)^\alpha}{w^-(.5)(S_1 - L_1)^\alpha} = \frac{w^+(.5)}{w^-(.5)}\left(\frac{(H_1 - L_1)/2}{(H_1 - L_1)/2}\right)^\alpha = \frac{w^+(.5)}{w^-(.5)} \end{aligned}$$

  As the safe amount for all first decision tasks within a sequence of tasks to determine a given lottery's certainty equivalent is always determined in the same way (i.e., $S_i = (H_i + L_i)/2$), all equations, particularly the last two, hold independently of the value of H_1 and L_1; in particular, they hold for $(H_2, L_2) = (CE_{.5}, 0)$ as in the fourth ($= (2-1)3+1$) decision, $(H_3, L_3) = (140, CE_{.5})$ as in the seventh ($= (3-1)3+1$) decision, and so on.

  Therefore, any decision-maker with $\lambda < \frac{w^+(.5)}{w^-(.5)}$ will always opt for the lottery in these first tasks, while those with $\lambda > \frac{w^+(.5)}{w^-(.5)}$ will opt for the safe outcome.

- Second Decision(s):

  (a) Case $\lambda > \frac{w^+(.5)}{w^-(.5)}$: Safe amount decreased compared to first decision task ($S_2 = S_1 - (H_1 - L_1)/4 = (H_1 + 3L_1)/4$); lottery unchanged $(H_1, 0.5; L_1)$.

  $$\begin{aligned} \text{Indifference} \;&\Leftrightarrow\; w^+(.5)(H_1 - S_2)^\alpha - w^-(.5)\lambda(S_2 - L_1)^\alpha = (S_2 - S_2)^\alpha \\ &\Leftrightarrow\; \lambda = \frac{w^+(.5)(H_1 - S_2)^\alpha}{w^-(.5)(S_2 - L_1)^\alpha} = \frac{w^+(.5)}{w^-(.5)}\left(\frac{3/4(H_1 - L_1)}{1/4(H_1 - L_1)}\right)^\alpha \\ &= \frac{w^+(.5)}{w^-(.5)}3^\alpha \end{aligned}$$

(b) Case $\lambda < \frac{w^+(.5)}{w^-(.5)}$: Safe amount increased compared to first decision task ($S_2 = S_1 + (H_1 - L_1)/4 = (3H_1 + L_1)/4$); lottery unchanged ($H_1, 0.5; L_1$).

$$\begin{aligned}\text{Indifference} \;\Leftrightarrow\;& w^+(.5)(H_1 - S_2)^\alpha - w^-(.5)\lambda(S_2 - L_1)^\alpha = (S_2 - S_2)^\alpha \\ \Leftrightarrow\;& \lambda = \frac{w^+(.5)(H_1 - S_2)^\alpha}{w^-(.5)(S_2 - L_1)^\alpha} = \frac{w^+(.5)}{w^-(.5)}\left(\frac{1/4(H_1 - L_1)}{3/4(H_1 - L_1)}\right)^\alpha \\ &= \frac{w^+(.5)}{w^-(.5)}(1/3)^\alpha\end{aligned}$$

Here again, we see that the difference between H_1 and L_1 cancels out, and thus, the degree of loss aversion λ that makes an individual indifferent between the lottery and the safe payment is independent of the values of H_i and L_i; in particular the equations thus also hold for $(H_2, L_2) = (CE_{.5}, 0)$ as in the fifth $(= (2 - 1)3 + 2)$ decision, $(H_3, L_3) = (140, CE_{.5})$ as in the eigth $(= (3 - 1)3 + 2)$ decision, and so on.

Therefore, any decision maker with $\frac{w^+(.5)}{w^-(.5)} < \lambda < \frac{w^+(.5)}{w^-(.5)}3^\alpha$ will, in the second decision of a sequence of tasks to determine a given lottery's certainty equivalent, always face the choice where the safe amount is decreased compared to the first decision (case a) here), and will, due to the specifics of the bisection approach as implemented by us, always opt for the lottery. In case $\frac{w^+(.5)}{w^-(.5)} < \frac{w^+(.5)}{w^-(.5)}3^\alpha < \lambda$, the decision maker will always opt for the safe amount in these second decisions. For the case b), the pattern follows by analogy.

- Third Decision(s):

(a) Case $\lambda > \frac{w^+(.5)}{w^-(.5)}$, $\lambda > \frac{w^+(.5)}{w^-(.5)}3^\alpha$: Safe amount decreased twice ($S_3 = S_2 - (H_1 - L_1)/8 = (H_1 + 7L_1)/8$); lottery unchanged ($H_1, 0.5; L_1$).

$$\begin{aligned}\text{Indifference} \;\Leftrightarrow\;& w^+(.5)(H_1 - S_3)^\alpha - w^-(.5)\lambda(S_3 - L_1)^\alpha = (S_3 - S_3)^\alpha \\ \Leftrightarrow\;& \lambda = \frac{w^+(.5)(H_1 - S_3)^\alpha}{w^-(.5)(S_3 - L_1)^\alpha} = \frac{w^+(.5)}{w^-(.5)}\left(\frac{7/8(H_1 - L_1)}{1/8(H_1 - L_1)}\right)^\alpha \\ &= \frac{w^+(.5)}{w^-(.5)}7^\alpha\end{aligned}$$

(b) Case $\lambda > \frac{w^+(.5)}{w^-(.5)}$, $\lambda < \frac{w^+(.5)}{w^-(.5)}3^\alpha$: Safe amount decreased, then increased ($S_3 = S_2 + (H_1 - L_1)/8 = (3H_1 + 5L_1)/8$); lottery unchanged ($H_1, 0.5; L_1$).

$$\begin{aligned}\text{Indifference} \;\Leftrightarrow\;& w^+(.5)(H_1 - S_3)^\alpha - w^-(.5)\lambda(S_3 - L_1)^\alpha = (S_3 - S_3)^\alpha \\ \Leftrightarrow\;& \lambda = \frac{w^+(.5)(H_1 - S_3)^\alpha}{w^-(.5)(S_3 - L_1)^\alpha} = \frac{w^+(.5)}{w^-(.5)}\left(\frac{5/8(H_1 - L_1)}{3/8(H_1 - L_1)}\right)^\alpha \\ &= \frac{w^+(.5)}{w^-(.5)}(5/3)^\alpha\end{aligned}$$

(c) Case $\lambda < \frac{w^+(.5)}{w^-(.5)}$, $\lambda > \frac{w^+(.5)}{w^-(.5)}(1/3)^\alpha$: Safe amount increased, then decreased ($S_3 = S_2 - (H_1 - L_1)/8 = (5H_1 + 3L_1)/8$); lottery unchanged $(H_1, 0.5; L_1)$.

$$\begin{aligned}\text{Indifference} \quad \Leftrightarrow \quad & w^+(.5)(H_1 - S_3)^\alpha - w^-(.5)\lambda(S_3 - L_1)^\alpha = (S_3 - S_3)^\alpha \\ \Leftrightarrow \quad & \lambda = \frac{w^+(.5)(H_1 - S_3)^\alpha}{w^-(.5)(S_3 - L_1)^\alpha} = \frac{w^+(.5)}{w^-(.5)}\left(\frac{3/8(H_1 - L_1)}{5/8(H_1 - L_1)}\right)^\alpha \\ & = \frac{w^+(.5)}{w^-(.5)}(3/5)^\alpha\end{aligned}$$

(d) Case $\lambda < \frac{w^+(.5)}{w^-(.5)}$, $\lambda < \frac{w^+(.5)}{w^-(.5)}(1/3)^\alpha$: Safe amount increased twice ($S_3 = S_2 + (H_1 - L_1)/8 = (7H_1 + L_1)/8$); lottery unchanged $(H_1, 0.5; L_1)$.

$$\begin{aligned}\text{Indifference} \quad \Leftrightarrow \quad & w^+(.5)(H_1 - S_3)^\alpha - w^-(.5)\lambda(S_3 - L_1)^\alpha = (S_3 - S_3)^\alpha \\ \Leftrightarrow \quad & \lambda = \frac{w^+(.5)(H_1 - S_3)^\alpha}{w^-(.5)(S_3 - L_1)^\alpha} = \frac{w^+(.5)}{w^-(.5)}\left(\frac{1/8(H_1 - L_1)}{7/8(H_1 - L_1)}\right)^\alpha \\ & = \frac{w^+(.5)}{w^-(.5)}(1/7)^\alpha\end{aligned}$$

Unsurprisingly, also for the third decisions, the difference between H_1 and L_1 cancels out, and again the degree of loss aversion λ, making individuals indifferent between the lottery and the safe payment, is independent of H_i and L_i and the equations thus also hold for $(H_2, L_2) = (CE_{.5}, 0)$ as in the sixth ($= (2-1) * 3 + 3$) decision, $(H_3, L_3) = (140, CE_{.5})$ as in the nineth ($= (3-1) * 3 + 3$) decision, and so on.

Thus, for a decision maker with utility given by (3), decision-making in all our decision tasks can be fully and exclusively described by the degree of loss aversion. In particular, it is independent of the exact values of H_i and L_i, therefore also independent of our starting values, and will thus be exactly the same for any sequence of lotteries that are built in this manner. As a trivial result, all decision makers described by such preferences can be identified by repeating patterns of decisions of length 3 in our tasks (since for every lottery, we use three iterations to elicit its indifference value).

Expected Value as Reference Point – Theoretical Analysis (with Calculations)

In this theoretical analysis of choice behavior in the elicitation tasks, we assume that the expected value of the lottery is considered as reference point when evaluating the lottery, that is, when utility is given by Equation (4) as discussed above. Recall that for every lottery that we consider, we elicit the corresponding certainty equivalent using three decisions (see Appendix A.3). Thus, the lottery stays the same in three subsequent decision tasks, where the safe amounts differ.

Again, we first derive the loss-aversion parameter that corresponds to indifference between the two choice options. We then derive how decision makers choose, based on their loss-aversion parameter. And, here again, the so derived decision patterns will be independent of whether we elicit a certainty equivalent for the first, second, ... lottery so that they will be the same for all six lotteries.

- First Decision(s): Lottery (outcomes $L_1 = x_{\min} = 0$ and $H_1 = x_{\max} = 140$ with equal probability) vs. safe amount ($S_1 = (L_1 + H_1)/2 = L_1 + (H_1 - L_1)/2 = M_1$).

$$\begin{aligned} \text{Indifference} \;&\Leftrightarrow\; w^+(.5)(H_1 - M_1)^\alpha - w^-(.5)\lambda(M_1 - L_1)^\alpha = (S_1 - M_1)^\alpha \\ &\Leftrightarrow\; \lambda = \frac{w^+(.5)(H_1 - M_1)^\alpha}{w^-(.5)(M_1 - L_1)^\alpha} = \frac{w^+(.5)}{w^-(.5)}\left(\frac{(H_1 - L_1)/2}{(H_1 - L_1)/2}\right)^\alpha \\ &\qquad = \frac{w^+(.5)}{w^-(.5)} \end{aligned}$$

  As in the first decision(s), the sure outcome is the expectation of the lottery, decision-making in these first tasks for the case of the sure-outcome reference dependence coincides with the case of expectation-based reference dependence, and has thus been fully analyzed and discussed above: Any decision maker with $\lambda < \frac{w^+(.5)}{w^-(.5)}$ will always opt for the lottery in these first tasks, while those with $\lambda > \frac{w^+(.5)}{w^-(.5)}$ will opt for the safe outcome.

- Second Decision(s):

  (a) Case $\lambda > \frac{w^+(.5)}{w^-(.5)}$: Safe amount decreased compared to first decision task ($S_2 = S_1 - (H_1 - L_1)/4 = (H_1 + 3L_1)/4$); lottery unchanged ($H_1, 0.5; L_1$).

$$\begin{aligned} \text{Indifference} \;&\Leftrightarrow\; w^+(.5)(H_1 - M_1)^\alpha - w^-(.5)\lambda(M_1 - L_1)^\alpha = -\lambda(M_1 - S_2)^\alpha \\ &\Leftrightarrow\; \lambda \overset{*}{=} \frac{w^+(.5)(H_1 - M_1)^\alpha}{w^-(.5)(M_1 - L_1)^\alpha - (M_1 - S_2)^\alpha} \\ &\qquad = \frac{w^+(.5)}{w^-(.5) - (1/2)^\alpha}\left(\frac{(H_1 - L_1)/2}{(H_1 - L_1)/2}\right)^\alpha \\ &\qquad = \frac{w^+(.5)}{w^-(.5) - (1/2)^\alpha} \end{aligned}$$

  The equality labelled with * is only defined if the denominator is different from zero. Thus, we see that for a decision maker without probability weighting on the loss domain and without utility curvature, there is no indifference, and the lottery is always preferred. For all others, the lottery is preferred if $\lambda < \frac{w^+(.5)}{w^-(.5) - (1/2)^\alpha}$. In case the reverse inequality holds, decision makers choose the safe amount.

(b) Case $\lambda < \frac{w^+(.5)}{w^-(.5)}$: Safe amount increased compared to first decision task ($S_2 = S_1 + (H_1 - L_1)/4 = (3H_1 + L_1)/4$); lottery unchanged $(H_1, 0.5; L_1)$.

$$\begin{aligned}\text{Indifference} \;&\Leftrightarrow\; w^+(.5)(H_1 - M_1)^\alpha - w^-(.5)\lambda(M_1 - L_1)^\alpha = (S_2 - M_1)^\alpha \\ &\Leftrightarrow\; \lambda = \frac{w^+(.5)(H_1 - M_1)^\alpha - (S_2 - M_1)^\alpha}{w^-(.5)(M_1 - L_1)^\alpha} \\ &= \frac{w^+(.5) - (1/2)^\alpha}{w^-(.5)} \left(\frac{(H_1 - L_1)/2}{(H_1 - L_1)/2}\right)^\alpha \\ &= \frac{w^+(.5) - (1/2)^\alpha}{w^-(.5)}\end{aligned}$$

Also in this case, we see that the difference between H_1 and L_1 cancels out, and thus, the degree of loss aversion λ that makes an individual indifferent between the lottery and the safe payment is independent of the values of H_i and L_i; in particular the equations thus also hold for $(H_2, L_2) = (CE_{.5}, 0)$ as in the fifth $(= (2-1)3+2)$ decision, $(H_3, L_3) = (140, CE_{.5})$ as in the eigth $(= (3-1)3+2)$ decision, and so on.

Therefore, any decision maker with $\frac{w^+(.5)}{w^-(.5)} < \lambda$ and $\lambda < \frac{w^+(.5)}{w^-(.5)-(1/2)^\alpha}$ will, in the second decision of a sequence of tasks to determine a given lottery's certainty equivalent, always face the choice where the safe amount is decreased compared to the first decision (case a) here), and will, due to the specifics of the bisection approach as implemented by us, always opt for the lottery. In case $\frac{w^+(.5)}{w^-(.5)} < \lambda$ and $\frac{w^+(.5)}{w^-(.5)-(1/2)^\alpha} < \lambda$, the decision maker will always opt for the safe amount in these second decisions. For the case b), the pattern follows by analogy.

- Third Decision(s):

  (a) Case $\lambda > \frac{w^+(.5)}{w^-(.5)}$, $\lambda > \frac{w^+(.5)}{w^-(.5)-(1/2)^\alpha}$: Safe amount decreased twice ($S_3 = S_2 - (H_1 - L_1)/8 = (H_1 + 7L_1)/8$); lottery unchanged $(H_1, 0.5; L_1)$.

$$\begin{aligned}\text{Indifference} \;&\Leftrightarrow\; w^+(.5)(H_1 - M_1)^\alpha - w^-(.5)\lambda(M_1 - L_1)^\alpha = -\lambda(M_1 - S_3)^\alpha \\ &\Leftrightarrow\; \lambda \overset{*}{=} \frac{w^+(.5)(H_1 - M_1)^\alpha}{w^-(.5)(M_1 - L_1)^\alpha - (M_1 - S_3)^\alpha} \\ &= \frac{w^+(.5)}{w^-(.5) - (3/4)^\alpha} \left(\frac{(H_1 - L_1)/2}{(H_1 - L_1)/2}\right)^\alpha \\ &= \frac{w^+(.5)}{w^-(.5) - (3/4)^\alpha}\end{aligned}$$

  Also in this case, the equality labelled with * is only defined if the denominator is different from zero, which here is the case if $w^-(.5) \neq (3/4)^\alpha$.

(b) Case $\lambda > \frac{w^+(.5)}{w^-(.5)}$, $\lambda < \frac{w^+(.5)}{w^-(.5)-(1/2)^\alpha}$: Safe amount decreased, then increased ($S_3 = S_2 + (H_1 - L_1)/8 = (3H_1 + 5L_1)/8$); lottery unchanged $(H_1, 0.5; L_1)$.

$$\begin{aligned} \text{Indifference} \;\Leftrightarrow\;& w^+(.5)(H_1 - M_1)^\alpha - w^-(.5)\lambda(M_1 - L_1)^\alpha = -\lambda(M_1 - S_3)^\alpha \\ \Leftrightarrow\;& \lambda \overset{*}{=} \frac{w^+(.5)(H_1 - M_1)^\alpha}{w^-(.5)(M_1 - L_1)^\alpha - (M_1 - S_3)^\alpha} \\ &= \frac{w^+(.5)}{w^-(.5) - (1/4)^\alpha}\left(\frac{(H_1 - L_1)/2}{(H_1 - L_1)/2}\right)^\alpha \\ &= \frac{w^+(.5)}{w^-(.5) - (1/4)^\alpha} \end{aligned}$$

Here the equality labelled with * is only defined if the denominator is different from zero, which here is the case if $w^-(.5) \neq (1/4)^\alpha$.

(c) Case $\lambda < \frac{w^+(.5)}{w^-(.5)}$, $\lambda > \frac{w^+(.5)-(1/2)^\alpha}{w^-(.5)}$: Safe amount increased, then decreased ($S_3 = S_2 - (H_1 - L_1)/8 = (5H_1 + 3L_1)/8$); lottery unchanged $(H_1, 0.5; L_1)$.

$$\begin{aligned} \text{Indifference} \;\Leftrightarrow\;& w^+(.5)(H_1 - M_1)^\alpha - w^-(.5)\lambda(M_1 - L_1)^\alpha = (S_3 - M_1)^\alpha \\ \Leftrightarrow\;& \lambda = \frac{w^+(.5)(H_1 - M_1)^\alpha - (S_3 - M_1)^\alpha}{w^-(.5)(M_1 - L_1)^\alpha} \\ &= \frac{w^+(.5) - (1/4)^\alpha}{w^-(.5)}\left(\frac{(H_1 - L_1)/2}{(H_1 - L_1)/2}\right)^\alpha \\ &= \frac{w^+(.5) - (1/4)^\alpha}{w^-(.5)} \end{aligned}$$

(d) Case $\lambda < \frac{w^+(.5)}{w^-(.5)}$, $\lambda < \frac{w^+(.5)-(1/2)^\alpha}{w^-(.5)}$: Safe amount increased twice ($S_3 = S_2 + (H_1 - L_1)/8 = (7H_1 + L_1)/8$); lottery unchanged $(H_1, 0.5; L_1)$.

$$\begin{aligned} \text{Indifference} \;\Leftrightarrow\;& w^+(.5)(H_1 - M_1)^\alpha - w^-(.5)\lambda(M_1 - L_1)^\alpha = (S_3 - M_1)^\alpha \\ \Leftrightarrow\;& \lambda = \frac{w^+(.5)(H_1 - M_1)^\alpha - (S_3 - M_1)^\alpha}{w^-(.5)(M_1 - L_1)^\alpha} \\ &= \frac{w^+(.5) - (3/4)^\alpha}{w^-(.5)}\left(\frac{(H_1 - L_1)/2}{(H_1 - L_1)/2}\right)^\alpha \\ &= \frac{w^+(.5) - (3/4)^\alpha}{w^-(.5)} \end{aligned}$$

Again, for all the possible third decisions, the difference between H_1 and L_1 cancels out, and again the degree of loss aversion λ making individuals indifferent between the lottery and the safe payment is independent of H_i and L_i and the equations also hold for $(H_2, L_2) = (CE_{.5}, 0)$ as in the sixth $(= (2-1)*3+3)$ decision, $(H_3, L_3) = (140, CE_{.5})$ as in the nineth $(= (3-1)*3+3)$ decision, and so on.

Thus, also for a decision maker with utility given by (4), decision-making in all our decision tasks can be fully and exclusively described by the degree of loss aversion. In particular, it is independent of the exact values of H_i and L_i in this case, too, and therefore also independent of our starting values, and will thus be exactly the same for any sequence of lotteries that are built in this manner. As a trivial result, here again, all decision makers described by such preferences can be identified by repeating patterns of decisions of length 3 in our tasks (since for every lottery, we use three iterations to elicit its indifference value).

Prospect as Reference Point – Theoretical Analysis (with Calculations) As a last step, we theoretically analyze choice behavior in the elicitation tasks when assuming that the lottery itself is considered as reference point when evaluating the lottery, that is, when utility is given by Equation (5). As we elicit the corresponding certainty equivalent using three decisions (see Appendix A.3), the lottery stays the same in three subsequent decision tasks, and the safe amounts differ.

Just as before, we first derive the loss-aversion parameter that corresponds to indifference between the two choice options in a decision task. From that we derive the choice of decision makers based on their loss-aversion parameter. The so derived decision patterns in the three subsequent choice tasks to elicit a lottery's certainty equivalent will again be independent of whether we elicit the certainty equivalent for the first, second, ... lottery that we present, so that these patterns will be the same for all six lotteries.

- First Decision(s): Lottery (outcomes $L_1 = x_{\min} = 0$ and $H_1 = x_{\max} = 140$ with equal probability) vs. safe amount ($S_1 = (L_1 + H_1)/2 = L_1 + (H_1 - L_1)/2$).

$$\begin{aligned} \text{Indifference} \;&\Leftrightarrow\; (H_1 + L_1)/2 + .25\eta(H_1 - L_1) - .25\eta\lambda(H_1 - L_1) = S_1 \\ &\Leftrightarrow\; \lambda = 1 \end{aligned}$$

  That is: Any decision maker with $\lambda < 1$ will always opt for the lottery in these first tasks (as the result is independent of H_1 and L_1), while those with $\lambda > 1$ will opt for the safe outcome.

- Second Decision(s):

  (a) Case $\lambda > 1$: Safe amount decreased compared to first decision task ($S_2 = S_1 - (H_1 - L_1)/4 = (H_1 + 3L_1)/4$); lottery unchanged $(H_1, 0.5; L_1)$.

$$\begin{aligned} \text{Indifference} \;&\Leftrightarrow\; (H_1 + L_1)/2 + .25\eta(H_1 - L_1) - .25\eta\lambda(H_1 - L_1) = S_2 \\ &\Leftrightarrow\; \lambda = \frac{(H_1 + L_1)/2 - S_2}{.25\eta(H_1 - L_1)} + 1 = \frac{(H_1 - L_1)/4}{\eta(H_1 - L_1)/4} + 1 \\ &\quad\;\; = 1 + 1/\eta \end{aligned}$$

(b) Case $\lambda < 1$: Safe amount increased compared to first decision task ($S_2 = S_1 + (H_1 - L_1)/4 = (3H_1 + L_1)/4$); lottery unchanged $(H_1, 0.5; L_1)$.

$$\begin{aligned} \text{Indifference} \;&\Leftrightarrow\; (H_1 + L_1)/2 + .25\eta(H_1 - L_1) - .25\eta\lambda(H_1 - L_1) = S_2 \\ &\Leftrightarrow\; \lambda = \frac{(H_1 + L_1)/2 - S_2}{.25\eta(H_1 - L_1)} + 1 = -\frac{(H_1 - L_1)/4}{\eta(H_1 - L_1)/4} + 1 \\ &\quad\;\; = 1 - 1/\eta \end{aligned}$$

The difference between H_1 and L_1 cancels out, and thus, the degree of loss aversion λ that makes an individual indifferent between the lottery and the safe payment is independent of the values of H_i and L_i; in particular the equations thus also hold for $(H_2, L_2) = (CE_{.5}, 0)$ as in the fifth $(= (2-1)3+2)$ decision, $(H_3, L_3) = (140, CE_{.5})$ as in the eigth $(= (3-1)3+2)$ decision, etc.

Therefore, any decision maker with $\lambda > 1$ and $\lambda < 1 + 1/\eta$ will, in the second decision of a sequence of tasks to determine a given lottery's certainty equivalent, always face the choice where the safe amount is decreased compared to the first decision (case a) here), and will, due to the specifics of the bisection approach as implemented by us, always opt for the lottery. In case $\lambda > 1$ and $\lambda > 1 + 1/\eta$, the decision maker will always opt for the safe amount in these second decisions. For the case b), the pattern follows by analogy.

- Third Decision(s):

(a) Case $\lambda > 1$, $\lambda > 1 + 1/\eta$: Safe amount decreased twice ($S_3 = S_2 - (H_1 - L_1)/8 = (H_1 + 7L_1)/8$); lottery unchanged $(H_1, 0.5; L_1)$.

$$\begin{aligned} \text{Indifference} \;&\Leftrightarrow\; (H_1 + L_1)/2 + .25\eta(H_1 - L_1) - .25\eta\lambda(H_1 - L_1) = S_3 \\ &\Leftrightarrow\; \lambda = \frac{(H_1 + L_1)/2 - S_3}{.25\eta(H_1 - L_1)} + 1 = \frac{3/2(H_1 - L_1)/4}{\eta(H_1 - L_1)/4} + 1 \\ &\quad\;\; = 1 + 3/(2\eta) \end{aligned}$$

(b) Case $\lambda > 1$, $\lambda < 1 + 1/\eta$: Safe amount decreased, then increased ($S_3 = S_2 + (H_1 - L_1)/8 = (3H_1 + 5L_1)/8$); lottery unchanged $(H_1, 0.5; L_1)$.

$$\begin{aligned} \text{Indifference} \;&\Leftrightarrow\; (H_1 + L_1)/2 + .25\eta(H_1 - L_1) - .25\eta\lambda(H_1 - L_1) = S_3 \\ &\Leftrightarrow\; \lambda = \frac{(H_1 + L_1)/2 - S_3}{.25\eta(H_1 - L_1)} + 1 = \frac{1/2(H_1 - L_1)/4}{\eta(H_1 - L_1)/4} + 1 \\ &\quad\;\; = 1 + 1/(2\eta) \end{aligned}$$

(c) Case $\lambda < 1$, $\lambda > 1 - 1/\eta$: Safe amount increased, then decreased ($S_3 = S_2 - (H_1 -$

$L_1)/8 = (5H_1 + 3L_1)/8)$; lottery unchanged $(H_1, 0.5; L_1)$.

$$\begin{aligned}\text{Indifference} \;\Leftrightarrow\;& (H_1 + L_1)/2 + .25\eta(H_1 - L_1) - .25\eta\lambda(H_1 - L_1) = S_3 \\ \Leftrightarrow\;& \lambda = \frac{(H_1 + L_1)/2 - S_3}{.25\eta(H_1 - L_1)} + 1 = -\frac{1/2(H_1 - L_1)/4}{\eta(H_1 - L_1)/4} + 1 \\ & = 1 - 1/(2\eta)\end{aligned}$$

(d) Case $\lambda < 1$, $\lambda < 1 - 1/\eta$: Safe amount increased twice ($S_3 = S_2 + (H_1 - L_1)/8 = (7H_1 + L_1)/8)$; lottery unchanged $(H_1, 0.5; L_1)$.

$$\begin{aligned}\text{Indifference} \;\Leftrightarrow\;& (H_1 + L_1)/2 + .25\eta(H_1 - L_1) - .25\eta\lambda(H_1 - L_1) = S_3 \\ \Leftrightarrow\;& \lambda = \frac{(H_1 + L_1)/2 - S_3}{.25\eta(H_1 - L_1)} + 1 = -\frac{3/2(H_1 - L_1)/4}{\eta(H_1 - L_1)/4} + 1 \\ & = 1 - 3/(2\eta)\end{aligned}$$

Also here, for all the possible third decisions, the difference between H_1 and L_1 cancels out, thus the degree of loss aversion λ making individuals indifferent between the lottery and the safe payment is independent of H_i and L_i and the equations also hold for $(H_2, L_2) = (CE_{.5}, 0)$ as in the sixth $(= (2-1)*3+3)$ decision, $(H_3, L_3) = (140, CE_{.5})$ as in the nineth $(= (3 - 1) * 3 + 3)$ decision, and so on.

Thus, also for a decision maker with utility given by (5), decision-making in all our decision tasks can be fully and exclusively described by the degree of loss aversion. In particular, it is independent of the exact values of H_i and L_i, therefore independent of our starting values, and will thus be exactly the same for any sequence of lotteries that are built in this manner. As a trivial result, also here, all decision makers described by such preferences can be identified by repeating patterns of decisions of length 3 in our tasks (since for every lottery, we use three iterations to elicit its indifference value).

B.5 Discussing the Predictive Quality of the Survey Measure for Field Behavior compared to Experimental Measures

In this subsection, we discuss the perhaps puzzling fact that not only the survey-based risk aversion measure is correlated with higher-order risk preferences, but the experimental risk-aversion measure is so, too. Moreover, the correlation between the experimental risk-aversion measure and higher-order risk preferences is the higher one. Against this background, why should risk aversion not be related to field behavior because of its relation to, e.g., prudence?

The key to this puzzle is that no empirical measure will ever be a "pure" measure, and that all the correlations are way below 1. This holds for the correlation between the survey measures and the experimental measures, as well as for the correlations between the experimental measures. This implies that there is a large share of variation in all of these measures that cannot be explained by these other measures alone. In consequence, all of these measures could explain something that the other measures cannot (unless the remaining variation was only noise).

One could think about the survey measure capturing various dimensions, whereas the experimental measures only capture the dimension they are "designed" to capture, plus some common variation. We illustrate this a bit more formally: Let c be the common variation in risk measures (due to a general risk preference, or gender, age, cognitive abilities, and so on), r^* be pure risk aversion, and p^* denote pure prudence. Assume all are uniformly distributed random variables (where the length of the interval does not matter here), $c \sim \mathcal{U}(0,1)$, $r^* \sim \mathcal{U}(0,1)$, and $p^* \sim \mathcal{U}(0,1)$. Then, suppose our experimental measures of prudence and risk aversion are given by

$$p = c + p^* \quad \text{and} \quad r = c + r^*.$$

In this scenario, using 500 independently simulated observations, we have $\rho(r,p) = .53$ at the 1% significance level, mirroring what we see in our data. Now suppose that our survey risk measure s is able to capture both, pure prudence as well as pure risk aversion (and possibly other dimensions plus noise and measurement error). Depending on the set of experimental risk preference measures (raw or sophisticated measures), the raw (Pearson) correlations with the survey question in our data are even up to the same magnitude (about .27). We conservatively set a slightly higher relation with risk aversion than with prudence, and add a uniformly distributed component $\varepsilon \sim \mathcal{U}(0,1)$ to reflect measurement error and other components:

$$s = .4(.55r^* + .45p^*) + .6\varepsilon$$

With 500 independently simulated observations, we have $\rho(p,s) = .24$ and $\rho(r,s) = .29$, slightly exaggerating the relation with risk aversion and slightly underestimating the relation with prudence compared to our data (as a conservative approach). Finally consider behavior b_p that is driven by prudence, and, as we consider real-world outcomes, a plenitude of other factors, such as income, norms, and so on. We model this with a uniformly distributed component $\mu \sim \mathcal{U}(0,1)$:

$$b_p = .2p^* + .8\mu.$$

With the 500 independently simulated observations, we obtain a correlation between prudence-driven behavior and the survey measure of $\rho(b_p, s) = .08$ and a correlation with the prudence measure of $\rho(b_p, p) = .17$ at the 10% and 1% significance levels, again closely mirroring what we see in our data for the case of addictive behavior, for example. The relation with the risk-aversion measure is small and insignificant, and we have $\rho(b_p, r) = -.01$ with a p-value of about .8, again reflecting our results on addictive behavior, for example.

This answers our initial question: If the joint variation (c in this example) in the experimental measures is due to factors that are not specific to one of the dimensions risk aversion or prudence (examples could be IQ, or culture), the experimental measures may be highly correlated, but still have low explanatory power for the other dimension of risk attitude (i.e., the prudence measure has a low explanatory power for pure risk aversion and vice versa). The survey measure, as a "capture-it-all measure", does not have the precision of the prudence measure, but it still captures enough prudence to establish a (significant) relationship.

References

Abdellaoui, M. (2000). Parameter-free elicitation of utility and probability weighting functions. *Management Science 46*(11), 1497–1512.

Abdellaoui, M., A. Baillon, L. Placido, and P. P. Wakker (2011). The rich domain of uncertainty: Source functions and their experimental implementation. *American Economic Review 101*(2), 695–723.

Abdellaoui, M., H. Bleichrodt, and C. Paraschiv (2007). Loss aversion under prospect theory: A parameter-free measurement. *Management Science 53*(10), 1659–1674.

Abdellaoui, M., F. Vossmann, and M. Weber (2005). Choice-based elicitation and decomposition of decision weights for gains and losses under uncertainty. *Management Science 51*(9), 1384–1399.

Baillon, A., H. Bleichrodt, and V. Spinu (2020). Searching for the reference point. *Management Science 66*(1), 93–112.

Bleichrodt, H. and J. L. Pinto (2000). A parameter-free elicitation of the probability weighting function in medical decision analysis. *Management Science 46*(11), 1485–1496.

Bleichrodt, H. and P. van Bruggen (2020). The reflection effect for higher order risk preferences. *The Review of Economics and Statistics*, 1–45.

Booij, A. S., B. M. S. van Praag, and G. van de Kuilen (2009). A parametric analysis of prospect theory's functionals for the general population. *Theory and Decision 68*(1-2), 115–148.

Brooks, P., S. Peters, and H. Zank (2013). Risk behavior for gain, loss, and mixed prospects. *Theory and Decision 77*(2), 153–182.

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

Charness, G., T. Garcia, T. Offerman, and M. C. Villeval (2020). Do measures of risk attitude in the laboratory predict behavior under risk in and outside of the laboratory? *Journal of Risk and Uncertainty 60*(2), 99–123.

Deck, C. and H. Schlesinger (2010). Exploring higher order risk effects. *The Review of Economic Studies 77*(4), 1403–1420.

Dimmock, S. G., R. Kouwenberg, O. S. Mitchell, and K. Peijnenburg (2021). Household portfolio underdiversification and probability weighting: Evidence from the field. *The Review of Financial Studies 34*(9), 4524–4563.

Drichoutis, A. C. and J. L. Lusk (2016). What can multiple price lists really tell us about risk preferences? *Journal of Risk and Uncertainty 53*(2-3), 89–106.

Durbin, J. (1954). Errors in variables. *Revue de l'Institut International de Statistique / Review of the International Statistical Institute 22*(1/3), 23.

Eeckhoudt, L. and H. Schlesinger (2006). Putting risk in its proper place. *American Economic Review 96*(1), 280–289.

Epper, T., H. Fehr-Duda, and R. Schubert (2011). Energy-using durables: The role of time discounting in investment decisions. Swiss Federal Office of Energy (SFOE) Publication No. 290417.

Fehr-Duda, H., M. de Gennaro, and R. Schubert (2006). Gender, financial risk, and probability weights. *Theory and Decision 60*(2-3), 283–313.

Gillen, B., E. Snowberg, and L. Yariv (2019). Experimenting with measurement error: Techniques with applications to the Caltech Cohort Study. *Journal of Political Economy 127*(4), 1826–1863.

Goldstein, W. M. and H. J. Einhorn (1987). Expression theory and the preference reversal phenomena. *Psychological Review 94*(2), 236–254.

Harrison, G. W. and E. E. Rutström (2008a). Expected utility theory and prospect theory: one wedding and a decent funeral. *Experimental Economics 12*(2), 133–158.

Harrison, G. W. and E. E. Rutström (2008b). Risk aversion in the laboratory. In G. W. Harrison and J. C. Cox (Eds.), *Risk Aversion in Experiments (Research in Experimental Economics, Vol. 12)*, pp. 41–196. Bingley: Emerald Group Publishing Limited.

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

Kőszegi, B. and M. Rabin (2006). A model of reference-dependent preferences. *Quarterly Journal of Economics 121*(4), 1133–1165.

Kőszegi, B. and M. Rabin (2007). Reference-dependent risk attitudes. *American Economic Review 97*(4), 1047–1073.

Lattimore, P. K., J. R. Baker, and A. D. Witte (1992). The influence of probability on risky choice. *Journal of Economic Behavior & Organization 17*(3), 377–400.

l'Haridon, O. and F. M. Vieider (2019). All over the map: A worldwide comparison of risk preferences. *Quantitative Economics 10*(1), 185–215.

Murad, Z., M. Sefton, and C. Starmer (2016). How do risk attitudes affect measured confidence? *Journal of Risk and Uncertainty 52*(1), 21–46.

Murphy, R. O. and R. H. W. ten Brincke (2018). Hierarchical maximum likelihood parameter estimation for cumulative prospect theory: Improving the reliability of individual risk parameter estimates. *Management Science 64*(1), 308–326.

Noussair, C. N., S. T. Trautmann, and G. van de Kuilen (2014). Higher order risk attitudes, demographics, and financial decisions. *The Review of Economic Studies 81*(1), 325–355.

Post, T., M. J. van den Assem, G. Baltussen, and R. H. Thaler (2008). Deal or no deal? Decision making under risk in a large-payoff game show. *American Economic Review 98*(1), 38–71.

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

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior & Organization 3*(4), 323–343.

Rabin, M. (2000). Risk aversion and expected-utility theory: A calibration theorem. *Econometrica 68*(5), 1281–1292.

Rieskamp, J. (2008). The probabilistic nature of preferential choice. *Journal of Experimental Psychology: Learning, Memory, and Cognition 34*(6), 1446–1465.

Schneider, S. O., M. Ibanez, and G. Riener (2025). Measuring utility – An application to higher-order risk. Working Paper, MPI Bonn.

Sprenger, C. (2015). An endowment effect for risk: Experimental tests of stochastic reference points. *Journal of Political Economy 123*(6), 1456–1499.

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

Tanaka, T., C. F. Camerer, and Q. Nguyen (2010). Risk and time preferences: Linking experimental and household survey data from Vietnam. *American Economic Review 100*(1), 557–571.

Toubia, O., E. Johnson, T. Evgeniou, and P. Delquié (2013). Dynamic experiments for estimating preferences: An adaptive method of eliciting time and risk parameters. *Management Science 59*(3), 613–640.

Tu, Q. (2005). *Empirical analysis of time preferences and risk aversion*. Ph. D. thesis, Tilburg University.

Tversky, A. and D. Kahneman (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty 5*(4), 297–323.

van de Kuilen, G. and P. P. Wakker (2011). The midweight method to measure attitudes toward risk and ambiguity. *Management Science 57*(3), 582–598.

Wakker, P. and D. Deneffe (1996). Eliciting von Neumann-Morgenstern utilities when probabilities are distorted or unknown. *Management Science 42*(8), 1131–1150.

Wakker, P. P. (2010). *Prospect Theory: For Risk and Ambiguity*. Cambridge, UK: Cambridge University Press.

Zeisberger, S., D. Vrecko, and T. Langer (2010). Measuring the time stability of Prospect Theory preferences. *Theory and Decision 72*(3), 359–386.

C Online Appendix: Risk Aversion and Field Behavior: Literature Overview

In this appendix, we provide an overview of the evidence about how (and whether) risk aversion is related to field behavior (previous results on higher-order risk preferences are limited to Noussair et al. (2014) and a handful of laboratory studies that, however, ignore field behavior, i.e., any behavior outside the lab; see Table C-2). As the following text as well as Table C-1 will show, the evidence on a relation between risk aversion and field behavior depends very much on whether risk aversion was measured experimentally – in which case the evidence is in a large majority not significant – or measured with survey questions – which have much better predictive power, but which capture much more than just risk aversion, as we show in the results section V of the main paper.

For this literature overview, we have started with all the articles that we have collected on the topic over the years and the references therein, many of which have been pointed to us in seminars and conference presentations. We have also used those papers as starting points for a search with connectedpapers.com, and used Google Scholar as well as elicit.org with keywords like "risk preferences", "risk aversion", "field behavior", or "experiment". Finally, we screened the articles published in the American Economic Review, Econometrica, the Journal of the European Economic Association, the Journal of Political Economy, the Quarterly Journal of Economics, the Review of Economic Studies, and the Journal of Risk and Uncertainty in the 10 years before mid-2022, and again followed the references therein. This way, we have tried to address publication bias, to have a more representative basis.

We then kept all those papers that relate a measure of risk aversion to one of the domains of field behavior that we study (i.e., comparable to at least one of the questions that we include in our questionnaires for the respective domain). That means that, e.g., studies investigating the link between migration and risk preferences (Jaeger et al., 2010), or voting behavior and risk preferences (Pástor and Veronesi, 2020) are not considered. As the selection of field behavior that we study here is based on theoretical considerations, this effectively means that the types of field behavior excluded in this overview are unrelated to higher-order risk preferences (and/or that the behavior under study is irrelevant to the everyday behavior of adolescents, such as technology adoption of small scale farmers).

Regarding risk measures, we focus on the survey measure on willingness to take risk in general (Dohmen et al., 2011), and experimental methods that allow the derivation of an economic risk aversion measure. We do so for three reasons: a) to keep this review concise and informative (as most results are available for these measures), b) to mirror our setting in Table 2, and c) to have a clear separation between the linguistically holistic concept of willingness to take risk in general and experimentally obtained measures that can be related to economic theory.[35]

If a study contains several data sets or uses several measures, we treat each of them as a separate observation. In case several specifications are reported, we opt for the most simple one that controls for at least gender and age, if reported.

To ease readability, we have reversed the SOEP-Survey question (Dohmen et al., 2011) – actually a question on the willingness to take risk –, and report it as willingness to avoid risk (WAR) so that the sign of the relationship is in line with the experimental measures of risk aversion. Moreover, we only consider the general willingness-to-take-risks question and refrain from the differentiation into domains (leisure, finance, health, ...) in line with the approach we used in Section V.

Health-Related Behavior No study has looked at the relationship between risk preferences and obsessive smartphone usage before, but for addictive behavior more generally, several studies report results. As can be seen in Table C-1, for smoking and excessive drinking, the evidence regarding the relationship to experimental measures of risk aversion is rather weak: Besides the significant estimations at the 10% level reported by Anderson and Mellor (2008) with a Holt-Laury setup, significant results were reported only in subsamples (Harrison et al., 2010; Galizzi and Miraldo, 2017), and in the latter study only very few actually exhibit the behavior under study (Galizzi and Miraldo, 2017, report two smokers in their female sample). Similarly, results in Anderson and Mellor (2008) are sensitive to the definitions of smoking and drinking.[36]

These points notwithstanding, only in 4 out of 25 (16%) cases is behavior with respect to smoking or drinking significantly predicted by experimental measures of risk aversion, and those significant cases do not even have the same sign. The survey measure in turn predicts smoking and excessive or problematic drinking in 8 out of 9 (89%) cases significantly, and always with the same sign, thus providing a much clearer (and more predictive) picture.

Turning to high BMI values and exercising, we note that experimental measures seem to be more successful than the survey measures, even though the two results by Galizzi et al.

[35] For this reason we omit the Balloon Analogue Risk Task (BART, Lejuez et al., 2002), which can be considered an experimental method. Inclusion would, however, not change the picture here (Szrek et al., 2012). Similarly, we do not report results on survey lottery questions (Barsky et al., 1997) as they are somewhere in between survey measure and experimental measure, and as we have not collected such measures in our study. Unfortunately, studies more and more rely on the aggregation of the survey measure and the experimental measure following Falk et al. (2018), and refrain from reporting coefficients on the risk measures in isolation (e.g., Falk et al., 2018; Lades et al., 2021). As a result, we have excluded them here.

[36] Anderson and Mellor (2008) write that once they treat only those who smoke every day as a smoker (as opposed to those who smoke "some days"), the reported significant relation can no longer be found. Similarly, once excessive drinking is defined as having 2 instead of 5 drinks every day of the week (1 instead of 4 for females), the significant relationship between risk and drinking vanishes.

(2016) are at odds with those of previous studies. Nevertheless, in 5 out of 6 cases (83%), the experimental measures yield significant results when predicting the BMI or exercising. However, it has again to be noted that the signs are partly of opposite nature (three negative, two positive), which makes the evidence difficult to interpret. The survey measure yields significant results in 50% of all cases relating BMI or exercising to risk preferences; all with the same sign.

Summarizing this subsection, we note in this domain that experimentally elicited risk aversion measures are much less often significantly related to health-related behavior than survey measures are (with the latter showing a significant relation in 10 out of 13 cases, i.e., in 77%, see Table C-1).

Financial Decision-Making The field behavior in the domain of financial decision-making that is addressed in our questionnaire and its relation to risk measures have received considerable attention (e.g., Noussair et al., 2014; Charness et al., 2020; Galizzi et al., 2016). Several studies address saving behavior, debt, risky investment, and insurance coverage. Yet, among all the studies that we have identified for this overview (see Table C-1), the only experimental risk measure that has significantly been related to any financial decision-making is reported by Jacobson and Petrie (2009): Using the elicitation procedure for risk aversion due to Binswanger (1980), they successfully predict informal loan take-up in Rwanda (at the 10% level). For the same data, the measure resulting from lotteries à la Holt and Laury (2002) yields no relation, however, nor do these risk measures relate to insurance coverage or saving.

Survey measures have been related to financial decision-making with more success – the only study reporting a null finding is Charness et al. (2020), whereas Dohmen et al. (2011) and others report significant estimations.

In sum, in 1 out of 30 cases (3%), a significant relation (at the 10% level) between financial decision-making and an experimental measure of risk aversion has been established, whereas for the survey measure, this was possible in 5 out of 8 cases (63%), see Table C-1.

General Prevention We could not find any study linking measures of risk aversion to field behavior in the domain of (short- and long-term) prevention (which is not too surprising given the absence of a clear-cut theoretical prediction for the relation with risk aversion in both the short- and the long-term cases, contrarily to the case of prudence).

Eco-Friendly Behavior Inspired by our questionnaire and preliminary results, Fuhrmann-Riebel et al. (2021) confirm our finding with respect to risk aversion and eco-friendly behavior: Their measure of risk aversion – our simple measure of risk r^{CE} – predicts eco-friendly behavior as measured by sustainable plastics use. Actual energy consumption and sustainable energy use (here measured as adhering to energy saving measures), however, are unrelated to risk in their data. This is in line with Ziegler (2020), who fails to explain consumption of sustainable (green) energy using the survey measure.[37] Schleich et al. (2019) and Volland (2017), in contrast, significantly relate energy-efficient technology use to a risk aversion measure (building on Holt-Laury), and energy consumption to the survey measure, respectively.

Mirroring our results, for eco-friendly behavior, experimental risk aversion measures are relatively successful in predicting field behavior (significant in 2 of 4 (=50%) cases) – and here even more successful than the survey measure (significant in 1 of 3 (= 33%) cases).

Planning Behavior We have found only one study – Ding et al. (2010) – that linked planning behavior (exam preparation in their case) to a measure of risk aversion or the survey question on willingness to take risk. While the survey measure successfully predicts this kind of field behavior, only one experimental measure yields a significant estimation at the 10% level. The second experimental risk aversion measure is unrelated with planning behavior.

Preference for Competitive Income Studies investigating field behavior that is related to a preference for a competitive income as measured here either address self-employment (e.g., Caliendo et al., 2007) or measures that describe individuals' willingness to accept a non-stable income (e.g., Bonin et al., 2007; Dohmen and Falk, 2011). These latter studies either relate earnings risk to measures of risk aversion, or predict choosing a performance-based payment scheme in experiments. Although we focus on field behavior here, these choices might arguably be treated as field behavior, as they might determine students' earnings for more than a day's consumption.

While the experimental measures of risk aversion seem to predict choice in these kinds of experiments relatively well (5/8 = 63%), they are less successful in predicting self-employment (1/9 = 11%) compared to the survey measure (4/5 = 80%). Thus, altogether, the survey measure yields significant results in 11 of 12 cases here (92%), whereas in 6 of 17 (35%) cases the experimental measures leads to significant estimations, see Table C-1.

[37] Lades et al. (2021), measuring risk aversion with the combination of survey and experimental measure following Falk et al. (2018), fail to predict their index of eco-friendly behavior, too.

Overall assessment All in all, for all the everyday field behavior that we address in this study, we conclude that the survey measure asking for the general willingness to take risks is more successful in predicting field behavior than experimental measures of risk aversion, except for the case of eco-friendly behavior. With respect to eco-friendly behavior, experimental measures yield significant results more often, even though the number of studies is rather small in this domain. The difference in the shares of significant results amounts to roughly 17 percentage points for this type of field behavior (see Table C-1). For all the other types of behavior, particularly health-related behavior and financial decision making, we observe a stark contrast: The survey measure performs way superior in relating risk preferences to field behavior – the difference in the shares of significant estimations even exceeds 50 percentage points in these two areas, and the absolute shares of significant results when using the experimental measures are rather close to what one has to expect by pure chance (e.g., 3% significant relations obtained with the experimental measures in the case of financial decision making, see Table C-1). For addictive behavior and financial decision-making, the performance of the experimental measures is particularly weak; the survey measures are relatively successful in those two areas.

This picture reflects our results from Table 2 in the main paper very well, where the survey measure is significantly predicting health-related behavior, but the simple measure of risk aversion is not. Similarly, the survey measure predicts financial decision-making, whereas the simple measure of risk aversion does not. Also in line with the observation here, the simple measure of risk aversion predicts eco-friendly behavior significantly in Table 2, as well as cautious planning, although this latter behavior not quite with the precision of the survey measure.

In summary, all this supports our conclusion from Section V: The survey measure used in the previous studies summarized here seems to capture something in addition to risk aversion, for which reason it shows a significant relation to field behavior much more often than the experimental measures of risk aversion. This insight from our literature review nicely mirrors the general pattern emerging from our Table 2 in the main paper. There we show that our survey measure predicts field behavior significantly, but mainly so because it is related to our estimates of higher-order risk preferences, not because the survey measure captures risk aversion so well. This is in line with the theoretical predictions for higher-order risk preferences for all the behavior considered here (see Section I.C and the theoretical considerations in Appendix E), and corroborates our main message from Section V.

Table C-1: Literature Overview: Risk Aversion and Field Behavior

Domain	Outcome	Experimental Measure of Risk Aversion[a]			Survey Measure (WAR = 10 - WTR)[a]		
		Significant Relation	Sign[b]	**No Relation**	**Significant Relation**	Sign[b]	**No Relation**
Health	Smoking	Anderson and Mellor (2008): HL Harrison et al. (2010, female subsample): HL Galizzi and Miraldo (2017, female subsample): HL	− + − − −	Ding et al. (2010): CE/MPL Ding et al. (2010): CE/MPL (Prob.) Harrison et al. (2010, male subsample): HL Szrek et al. (2012): HL Sutter et al. (2013): CE/MPL Galizzi et al. (2016): B/EG Galizzi et al. (2016): HL Galizzi et al. (2016): HL w./ financial outcomes Galizzi et al. (2016): HL w./ health outcomes Galizzi and Miraldo (2017, male subsample): HL Harrison et al. (2018): HO/LS Beauchamp et al. (2019): CE/Categorization	Ding et al. (2010) Dohmen et al. (2011) Szrek et al. (2012) Galizzi et al. (2016) Beauchamp et al. (2019)	− − − − − − − − − −	
	Alcohol	Anderson and Mellor (2008): HL	−	Ding et al. (2010): CE/MPL Ding et al. (2010): CE/MPL (Prob.) Szrek et al. (2012): HL Sutter et al. (2013): CE/MPL Galizzi et al. (2016): HL Galizzi et al. (2016): B/EG Galizzi et al. (2016): HL w./ financial outcomes Galizzi et al. (2016): HL w./ health outcomes Beauchamp et al. (2019): CE/Categorization)	Ding et al. (2010) Szrek et al. (2012) Beauchamp et al. (2019)	− − − − − −	Galizzi et al. (2016)
	High BMI	Anderson and Mellor (2008): HL Sutter et al. (2013): CE/MPL de Oliveira et al. (2016): B/EG Galizzi et al. (2016): HL Galizzi et al. (2016): B/EG	− − − − − − − ++ ++	Samek et al. (2021): CE/MPL	Samek et al. (2021)	−	Galizzi et al. (2016)
	Exercising	Leonard et al. (2013): B/EG	− −	Galizzi et al. (2016): HL w./ financial outcomes Galizzi et al. (2016): HL w./ health outcomes Samek et al. (2021): CE/MPL	Dohmen et al. (2011)	− − −	Samek et al. (2021)
		Share significant (and in line with majority of results): 7/29 = 24%			Share significant: 10/13 = 77%		

[a] CE refers to studies that elicit a certain amount of money for a fixed lottery (e.g., Sutter et al., 2013), or alternatively elicit the high outcome of a lottery for which the certainty equivalent stays fixed (e.g., Samek et al., 2021, following Benjamin et al., 2013). Moreover, B/EG: Binswanger (1980)/Eckel and Grossman (2008), GP: Gneezy and Potters (1997), HL: Holt and Laury (2002), HO: Hey and Orme (1994), and TCN: Tanaka et al. (2010). WTR refers to the one-item survey question on general willingness to take risks (Dohmen et al., 2011).

[b] −/− −/− − − and +/++/+ + + indicate significant relations at the 10 / 5 / 1 percent level.

Table C-1: Literature Overview: Risk Aversion and Field Behavior *(continued)*

DOMAIN	OUTCOME	EXPERIMENTAL MEASURE OF RISK AVERSION[a]			SURVEY MEASURE (WAR = 10 - WTR)[a]		
		Significant Relation	Sign[b]	**No Relation**	**Significant Relation**	Sign[b]	**No Relation**
Financial Decision-Making	Savings			Jacobson and Petrie (2009): HL Jacobson and Petrie (2009): EG/B Sutter et al. (2013): CE/MPL Noussair et al. (2014): CE/MPL Galizzi et al. (2016): HL Galizzi et al. (2016): B/EG Charness et al. (2020): HL Charness et al. (2020): GP Charness et al. (2020): TCN Charness et al. (2020): B/EG	Galizzi et al. (2016):	++	Charness et al. (2020)
	Debt[c]	Jacobson and Petrie (2009): EG/B	−	Jacobson and Petrie (2009): HL Noussair et al. (2014): CE/MPL Fairley and Weitzel (2017): CE/MPL	Oosterbeek and van den Broek (2009)	− − −	
	Risky Investment[d]			Ding et al. (2010): CE/MPL Ding et al. (2010): CE/MPL (Prob.) Noussair et al. (2014): CE/MPL Beauchamp et al. (2019): CE/Categorization Charness et al. (2020): HL Charness et al. (2020): GP Charness et al. (2020): TCN Charness et al. (2020): B/EG	Ding et al. (2010) Dohmen et al. (2011) Beauchamp et al. (2019) Nicolaou and Shane (2019)	− − − − − − − − − − − −	Charness et al. (2020)
	Financial Insurance			Jacobson and Petrie (2009): HL Jacobson and Petrie (2009): EG/B Noussair et al. (2014): CE/MPL[e] Bryan (2019): B/EG Charness et al. (2020): HL Charness et al. (2020): GP Charness et al. (2020): TCN Charness et al. (2020): B/EG			Charness et al. (2020)
		Share significant: 1/30 = 3%			Share significant: 6/9 = 67%		

[a]CE refers to studies that elicit a certain amount of money for a fixed lottery (e.g., Sutter et al., 2013), or alternatively elicit the high outcome of a lottery for which the certainty equivalent stays fixed (e.g., Samek et al., 2021, following Benjamin et al., 2013). Moreover, B/EG: Binswanger (1980)/Eckel and Grossman (2008), GP: Gneezy and Potters (1997), HL: Holt and Laury (2002), HO: Hey and Orme (1994), and TCN: Tanaka et al. (2010). WTR refers to the one-item survey question on general willingness to take risks (Dohmen et al., 2011).

[b]−/−−/− − − and +/++/+ + + indicate significant relations at the 10 / 5 / 1 percent level.

[c]Credit card debt, loan take-up and take-up of informal credit

[d]Portfolio risk, risky investment, and stock ownership

[e]Regression results on this were not reported, but the finding is discussed in the text.

Table C-1: Literature Overview: Risk Aversion and Field Behavior *(continued)*

Domain	Outcome	Experimental Measure of Risk Aversion[a]			Survey Measure (WAR = 10 - WTR)[a]		
		Significant Relation	Sign[b]	**No Relation**	**Significant Relation**	Sign[b]	**No Relation**
Eco-Friendly Behavior	Avoidance and Reuse of Plastic	Fuhrmann-Riebel et al. (2021): CE/Titration	+ + +				
	Sustainable Energy Consumption[c]	Schleich et al. (2019): HL	+	Fuhrmann-Riebel et al. (2021): CE/Titration			Ziegler (2020)
	Energy Consumption			Fuhrmann-Riebel et al. (2021): CE/Titration	Volland (2017)	− −	Fischbacher et al. (2021)
		Share significant: 2/4 = 50%			Share significant: 1/3 = 33.3%		
Planning Behavior	Exam Preparation	Ding et al. (2010): CE/MPL (Prob.)	+	Ding et al. (2010): CE/MPL	Ding et al. (2010)	+ +	
		Share significant: 1/2 = 50%			Share significant: 1/1 = 100%		
Preference for Competitive Income	Self-Employment	Beauchamp et al. (2019): CE/Categorization	−	Hardeweg et al. (2013): HL	Caliendo et al. (2007)	− −	Charness et al. (2020)
				Holm et al. (2013): HL	Dohmen et al. (2011)	− − −	
				Holm et al. (2013): CE/MPL	Hardeweg et al. (2013)	− − −	
				Andersen et al. (2014): HL	Beauchamp et al. (2019)	− − −	
				Charness et al. (2020): HL	Nicolaou and Shane (2019)	− − −	
				Charness et al. (2020): GP			
				Charness et al. (2020): TCN			
				Charness et al. (2020): B/EG			
	Non-stable Income[d]	Falco (2014): HL	− −	van Veldhuizen (2022, Experiment 1): HL[e]	Bonin et al. (2007)	− − −	
		Buser et al. (2014): B/EG	− − −	van Veldhuizen (2022, Experiment 2): HL[f]	Dohmen and Falk (2011)	− − −	
		Gillen et al. (2019): CE/MPL	− − −	van Veldhuizen (2022, Experiment 2): B/EG	Buser et al. (2014)	− − −	
		Gillen et al. (2019): GP	− − −		Fouarge et al. (2014)	− − −	
		van Veldhuizen (2022, Experiment 1): B/EG	−		Gillen et al. (2019)[g]	−	
					van Veldhuizen (2022, Exp. 1)	− − −	
					van Veldhuizen (2022, Exp. 2)	− − −	
		Share significant: 6/17 = 35%			Share significant: 12/13 = 92%		

[a]CE refers to studies that elicit a certain amount of money for a fixed lottery (e.g., Sutter et al., 2013), or alternatively elicit the high outcome of a lottery for which the certainty equivalent stays fixed (e.g., Samek et al., 2021, following Benjamin et al., 2013). Moreover, B/EG: Binswanger (1980)/Eckel and Grossman (2008), GP: Gneezy and Potters (1997), HL: Holt and Laury (2002), HO: Hey and Orme (1994), and TCN: Tanaka et al. (2010). WTR refers to the one-item survey question on general willingness to take risks (Dohmen et al., 2011).

[b]−/− −/− − − and +/++/+ + + indicate significant relations at the 10 / 5 / 1 percent level.

[c]Energy-efficient technology use, energy saving behavior and switching to green electricity contracts

[d]Earnings risk and choice of performance-based payment scheme

[e](Non-reported) coefficient from Column 4 of Table 2 (see replication material). Interestingly, the Eckel-Grossmann measure is not significant in that specification.

[f](Non-reported) coefficient from Column 4 of Table 4 (see replication material).

[g]Replacing the risk measures in the specifications of Columns 4 and 5 (GP and CE/MPL) with the survey measure yields this result (see replication material).

Table C-2: Summary of Empirical Literature on Higher-Order Risk Preferences (HORP) and Field Behavior

Domain	**Outcome**	**Related HORP (Section I.C)**	**Prior Empirical Evidence (Field Behavior)**	**Empirical Relation (Prior Work on HORP)**[a]
Health	Unhealthy Behavior (in General) Addictive Behavior Abusive Smartphone Usage	Prudence	N/A[b]	
Financial Decision-Making	Precautionary Saving Risky Investment Insurance Demand	Prudence Temperance Prudence	Noussair et al. (2014)[c]	+ − ?
Prevention and Environment	Short-Term Prevention (in General) Long-Term Prevention (in General) Eco-Friendly Behavior	Prudence Prudence Prudence	N/A[d,e] N/A[e] N/A	
Planning	Cautious Planning	Prudence	N/A	
Competition	Preference for Competitive Income	Prudence	N/A	

[a]A plus sign (+) indicates a positive relationship, while a minus sign (−) indicates a negative relationship. A question mark (?) indicates that there is no robust evidence in either direction, and a combination of signs indicates that the same model leads to different predictions; also see the corresponding footnotes.

[b]Attema et al. (2019) document individuals in the lab choosing the prudent lottery when hypothetical outcomes are the remaining years to live instead of monetary outcomes.

[c]Lambregts et al. (2021) conduct an insurance experiment with risk of non-reimbursement, and report a lower insurance take-up rate among prudent decision-makers, for reasons similar in spirit to those discussed in our theoretical considerations in Appendix E.4.

[d]Krieger and Mayrhofer (2016) measure prevention through preferences over two lotteries with two outcomes each, where they interpret choosing the "prevention" lottery (lower expected value in return for a lower variance) as prevention. They document a negative relationship with prudence.

[e]Masuda and Lee (2019) study prevention behavior in the laboratory using monetary loss events that can be prevented by monetary investments. They explicitly differentiate between the one- and the two-period setting in their prevention games, and find negative correlations with prudence in both settings (providing support for the predictions by Peter, 2017).

Table C-3: Selection of Additional Related Empirical Literature on Higher-Order Risk Preferences

Study	**Summary**	**Relation to our Work**
Joshi et al. (2021)	Study finds that a sample of Indian farmers is, on average, prudent and temperate	Findings are informative for the relation between prudence and occupational choice, which comes closest to our outcome regarding a preference for a competitive income. Note, however, that both studies rely on arguably self-selected samples while lacking a suitable comparison group.
Schaap (2021)	Study documents that their sample of Chilean fishermen is not prudent	

References

Andersen, S., A. D. Girolamo, G. W. Harrison, and M. I. Lau (2014). Risk and time preferences of entrepreneurs: evidence from a Danish field experiment. *Theory and Decision 77*(3), 341–357.

Anderson, L. R. and J. M. Mellor (2008). Predicting health behaviors with an experimental measure of risk preference. *Journal of Health Economics 27*(5), 1260–1274.

Attema, A. E., O. l'Haridon, and G. van de Kuilen (2019). Measuring multivariate risk preferences in the health domain. *Journal of Health Economics 64*, 15–24.

Barsky, R. B., F. T. Juster, M. S. Kimball, and M. D. Shapiro (1997). Preference parameters and behavioral heterogeneity: An experimental approach in the health and retirement study. *Quarterly Journal of Economics 112*(2), 537–579.

Beauchamp, J. P., D. J. Benjamin, D. I. Laibson, and C. F. Chabris (2019). Measuring and controlling for the compromise effect when estimating risk preference parameters. *Experimental Economics 23*(4), 1069–1099.

Benjamin, D. J., S. A. Brown, and J. M. Shapiro (2013). Who is 'behavioral'? Cognitive ability and anomalous preferences. *Journal of the European Economic Association 11*(6), 1231–1255.

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

Bonin, H., T. Dohmen, A. Falk, D. Huffman, and U. Sunde (2007). Cross-sectional earnings risk and occupational sorting: The role of risk attitudes. *Labour Economics 14*(6), 926–937.

Bryan, G. (2019). Ambiguity aversion decreases the impact of partial insurance: Evidence from African farmers. *Journal of the European Economic Association 17*(5), 1428–1469.

Buser, T., M. Niederle, and H. Oosterbeek (2014). Gender, competitiveness, and career choices. *Quarterly Journal of Economics 129*(3), 1409–1447.

Caliendo, M., F. M. Fossen, and A. S. Kritikos (2007). Risk attitudes of nascent entrepreneurs–new evidence from an experimentally validated survey. *Small Business Economics 32*(2), 153–167.

Charness, G., T. Garcia, T. Offerman, and M. C. Villeval (2020). Do measures of risk attitude in the laboratory predict behavior under risk in and outside of the laboratory? *Journal of Risk and Uncertainty 60*(2), 99–123.

de Oliveira, A. C. M., T. C. M. Leonard, K. Shuval, C. S. Skinner, C. Eckel, and J. C. Murdoch (2016). Economic preferences and obesity among a low-income African American community. *Journal of Economic Behavior & Organization 131*, 196–208.

Ding, X., J. Hartog, and Y. Sun (2010). Can we measure individual risk attitudes in a survey? IZA Discussion Paper No. 4807.

Dohmen, T. and A. Falk (2011). Performance pay and multidimensional sorting: Productivity, preferences, and gender. *American Economic Review 101*(2), 556–590.

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

Eckel, C. C. and P. J. Grossman (2008). Forecasting risk attitudes: An experimental study using actual and forecast gamble choices. *Journal of Economic Behavior & Organization 68*(1), 1–17.

Fairley, K. and U. Weitzel (2017). Ambiguity and risk measures in the lab and students' real-life borrowing behavior. *Journal of Behavioral and Experimental Economics 67*, 85–98.

Falco, P. (2014). Does risk matter for occupational choices? Experimental evidence from an African labour market. *Labour Economics 28*, 96–109.

Falk, A., A. Becker, T. Dohmen, B. Enke, D. Huffman, and U. Sunde (2018). Global evidence on economic preferences. *Quarterly Journal of Economics 133*(4), 1645–1692.

Fischbacher, U., S. Schudy, and S. Teyssier (2021). Heterogeneous preferences and investments in energy saving measures. *Resource and Energy Economics 63*, 101202.

Fouarge, D., B. Kriechel, and T. Dohmen (2014). Occupational sorting of school graduates: The role of economic preferences. *Journal of Economic Behavior & Organization 106*, 335–351.

Fuhrmann-Riebel, H., B. D'Exelle, and A. Verschoor (2021). The role of preferences for pro-environmental behaviour among urban middle class households in Peru. *Ecological Economics 180*, 106850.

Galizzi, M. M., S. R. Machado, and R. Miniaci (2016). Temporal stability, cross-validity, and external validity of risk preferences measures: Experimental evidence from a UK representative sample. Working Paper, LSE, London, UK.

Galizzi, M. M. and M. Miraldo (2017). Are you what you eat? Healthy behaviour and risk preferences. *The B.E. Journal of Economic Analysis & Policy 17*(1), 20160081.

Galizzi, M. M., M. Miraldo, and C. Stavropoulou (2016). In sickness but not in wealth. *Medical Decision Making 36*(4), 503–517.

Gillen, B., E. Snowberg, and L. Yariv (2019). Experimenting with measurement error: Techniques with applications to the Caltech Cohort Study. *Journal of Political Economy 127*(4), 1826–1863.

Gneezy, U. and J. Potters (1997). An experiment on risk taking and evaluation periods. *Quarterly Journal of Economics 112*(2), 631–645.

Hardeweg, B., L. Menkhoff, and H. Waibel (2013). Experimentally validated survey evidence on individual risk attitudes in rural Thailand. *Economic Development and Cultural Change 61*(4), 859–888.

Harrison, G. W., A. Hofmeyr, D. Ross, and J. T. Swarthout (2018). Risk preferences, time preferences, and smoking behavior. *Southern Economic Journal 85*(2), 313–348.

Harrison, G. W., M. I. Lau, and E. E. Rutström (2010). Individual discount rates and smoking: Evidence from a field experiment in Denmark. *Journal of Health Economics 29*(5), 708–717.

Hey, J. D. and C. Orme (1994). Investigating generalizations of expected utility theory using experimental data. *Econometrica 62*(6), 1291–1326.

Holm, H. J., S. Opper, and V. Nee (2013). Entrepreneurs under uncertainty: An economic experiment in China. *Management Science 59*(7), 1671–1687.

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

Jacobson, S. and R. Petrie (2009). Learning from mistakes: What do inconsistent choices over risk tell us? *Journal of Risk and Uncertainty 38*(2), 143–158.

Jaeger, D. A., T. Dohmen, A. Falk, D. Huffman, U. Sunde, and H. Bonin (2010). Direct evidence on risk attitudes and migration. *Review of Economics and Statistics 92*(3), 684–689.

Joshi, K., T. Ranganathan, and R. Ranjan (2021). Exploring higher order risk preferences of farmers in a water-scarce region: Evidence from a field experiment in West Bengal, India. *Journal of Quantitative Economics 19*(2), 317–344.

Krieger, M. and T. Mayrhofer (2016). Prudence and prevention: An economic laboratory experiment. *Applied Economics Letters 24*(1), 19–24.

Lades, L. K., K. Laffan, and T. O. Weber (2021). Do economic preferences predict pro-environmental behaviour? *Ecological Economics 183*, 106977.

Lambregts, T. R., P. van Bruggen, and H. Bleichrodt (2021). Insurance decisions under nonperformance risk and ambiguity. *Journal of Risk and Uncertainty 63*(3), 229–253.

Lejuez, C. W., J. P. Read, C. W. Kahler, J. B. Richards, S. E. Ramsey, G. L. Stuart, D. R. Strong, and R. A. Brown (2002). Evaluation of a behavioral measure of risk taking: The Balloon Analogue Risk Task (BART). *Journal of Experimental Psychology: Applied 8*(2), 75–84.

Leonard, T., K. Shuval, A. de Oliveira, C. S. Skinner, C. Eckel, and J. C. Murdoch (2013). Health behavior and behavioral economics: Economic preferences and physical activity stages of change in a low-income African-American community. *American Journal of Health Promotion 27*(4), 211–221.

Masuda, T. and E. Lee (2019). Higher order risk attitudes and prevention under different timings of loss. *Experimental Economics 22*(1), 197–215.

Nicolaou, N. and S. Shane (2019). Common genetic effects on risk-taking preferences and choices. *Journal of Risk and Uncertainty 59*(3), 261–279.

Noussair, C. N., S. T. Trautmann, and G. van de Kuilen (2014). Higher order risk attitudes, demographics, and financial decisions. *The Review of Economic Studies 81*(1), 325–355.

Oosterbeek, H. and A. van den Broek (2009). An empirical analysis of borrowing behaviour of higher education students in the Netherlands. *Economics of Education Review 28*(2), 170–177.

Pástor, Ľ. and P. Veronesi (2020). Political cycles and stock returns. *Journal of Political Economy 128*(11), 4011–4045.

Peter, R. (2017). Optimal self-protection in two periods: On the role of endogenous saving. *Journal of Economic Behavior & Organization 137*, 19–36.

Samek, A., A. Gray, A. Datar, and N. Nicosia (2021). Adolescent time and risk preferences: Measurement, determinants and field consequences. *Journal of Economic Behavior & Organization 184*, 460–488.

Schaap, R.-J. (2021). The prevalence of prudence in a risky occupation. *Economics Letters 207*, 110019.

Schleich, J., X. Gassmann, T. Meissner, and C. Faure (2019). A large-scale test of the effects of time discounting, risk aversion, loss aversion, and present bias on household adoption of energy-efficient technologies. *Energy Economics 80*, 377–393.

Sutter, M., M. G. Kocher, D. Glätzle-Rützler, and S. T. Trautmann (2013). Impatience and uncertainty: Experimental decisions predict adolescents field behavior. *American Economic Review 103*(1), 510–531.

Szrek, H., L.-W. Chao, S. Ramlagan, and K. Peltzer (2012). Predicting (un)healthy behavior: A comparison of risk-taking propensity measures. *Judgment and Decision Making 7*, 716–727.

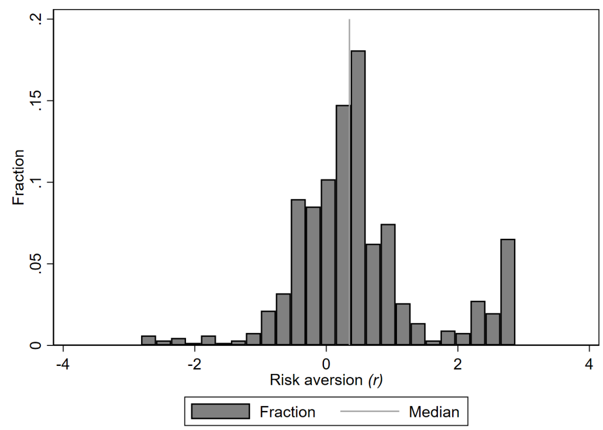
Tanaka, T., C. F. Camerer, and Q. Nguyen (2010). Risk and time preferences: Linking experimental and household survey data from Vietnam. *American Economic Review 100*(1), 557–571.

van Veldhuizen, R. (2022). Gender differences in tournament choices: Risk preferences, overconfidence, or competitiveness? *Journal of the European Economic Association 20*(4), 1595–1618.

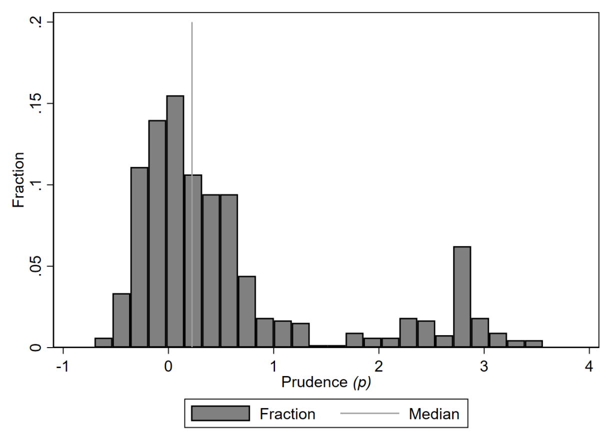
Volland, B. (2017). The role of risk and trust attitudes in explaining residential energy demand: Evidence from the United Kingdom. *Ecological Economics 132*, 14–30.

Ziegler, A. (2020). Heterogeneous preferences and the individual change to alternative electricity contracts. *Energy Economics 91*, 104889.

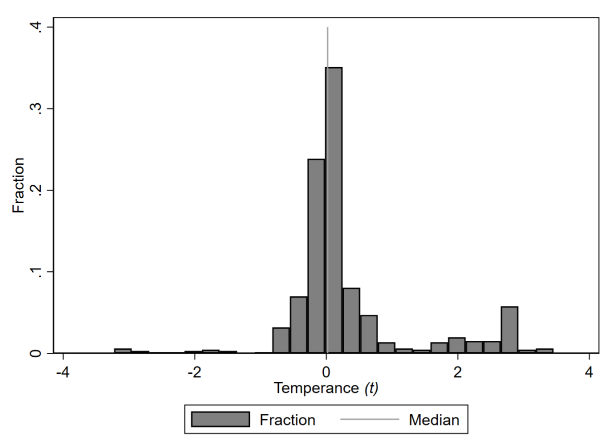
D Online Appendix: Additional Figures and Tables



(a) Histogram of the Arrow-Pratt risk-aversion measure

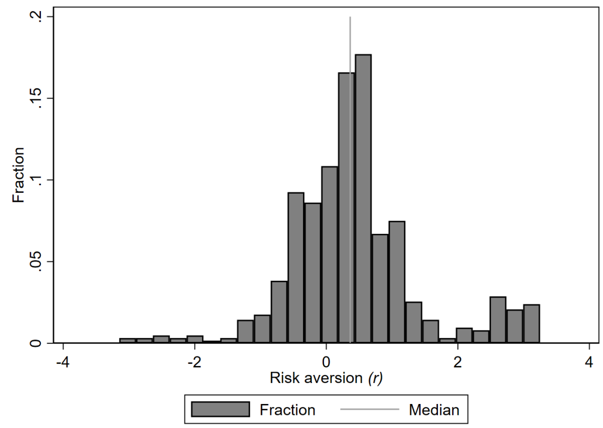


(b) Histogram of the Crainich and Eeckhoudt prudence measure

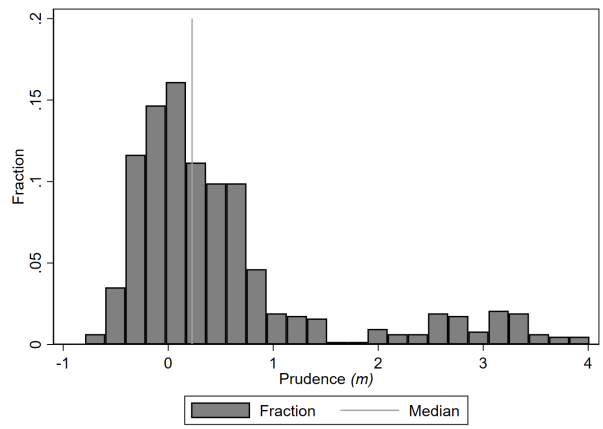


(c) Histogram of the Denuit and Eeckhoudt temperance measure

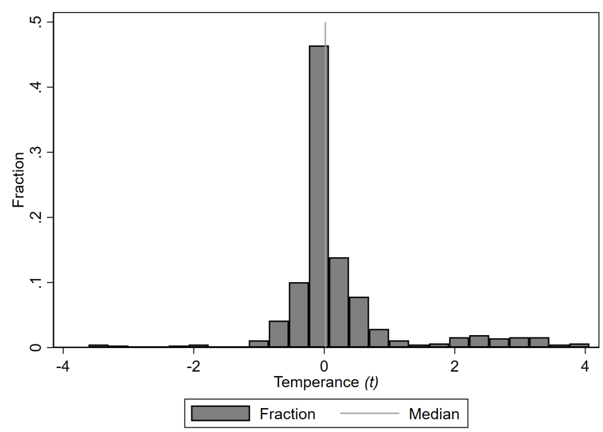
Figure D-1: Prevalences of (Higher-Order) Risk Preferences (in standard deviations)



(a) Histogram of the Arrow-Pratt risk-aversion measure



(b) Histogram of the Crainich and Eeckhoudt prudence measure



(c) Histogram of the Denuit and Eeckhoudt temperance measure

Figure D-2: Prevalences of (Higher-Order) Risk Preferences (in standard deviations) – When Excluding Observations Corresponding to Simple Patterns of Decision Making

Notes: See Figure D-1 above for the distribution of risk-preference measures without excluding observations corresponding to simple patterns of decision-making as discussed in Online Appendix Section B.1.

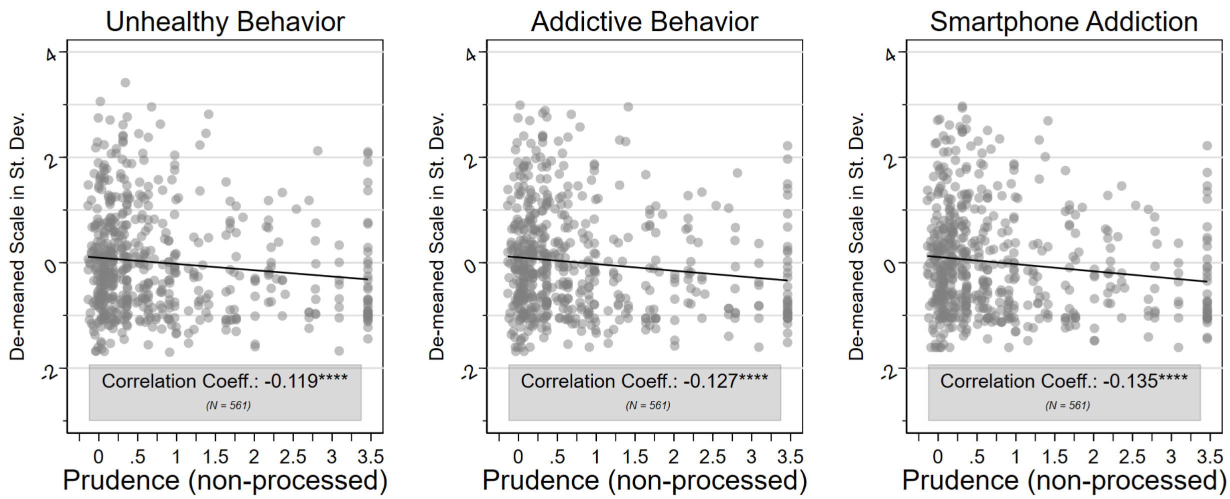


Figure D-3: Health-Related Behavior and Non-Processed, Simple Measure of Prudence p^{CE}: Raw (Pairwise/Pearson) Correlations with Linear Fit

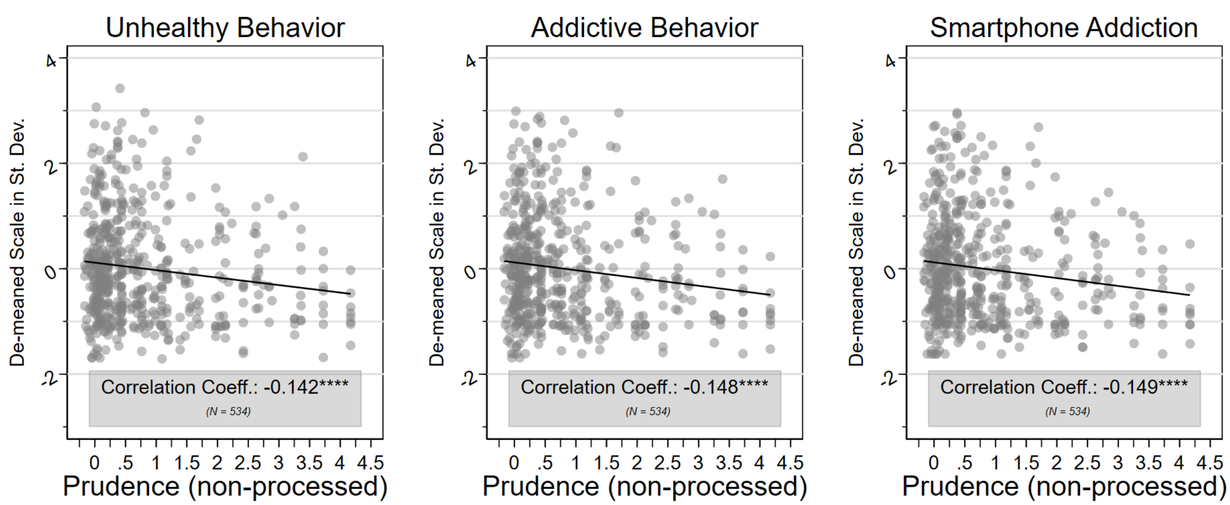


Figure D-4: Health-Related Behavior and Non-Processed, Simple Measure of Prudence p^{CE}: Raw (Pairwise/Pearson) Correlations with Linear Fit – When Excluding Observations Corresponding to Simple Patterns of Decision Making (c.f. Online Appendix Section B.1)

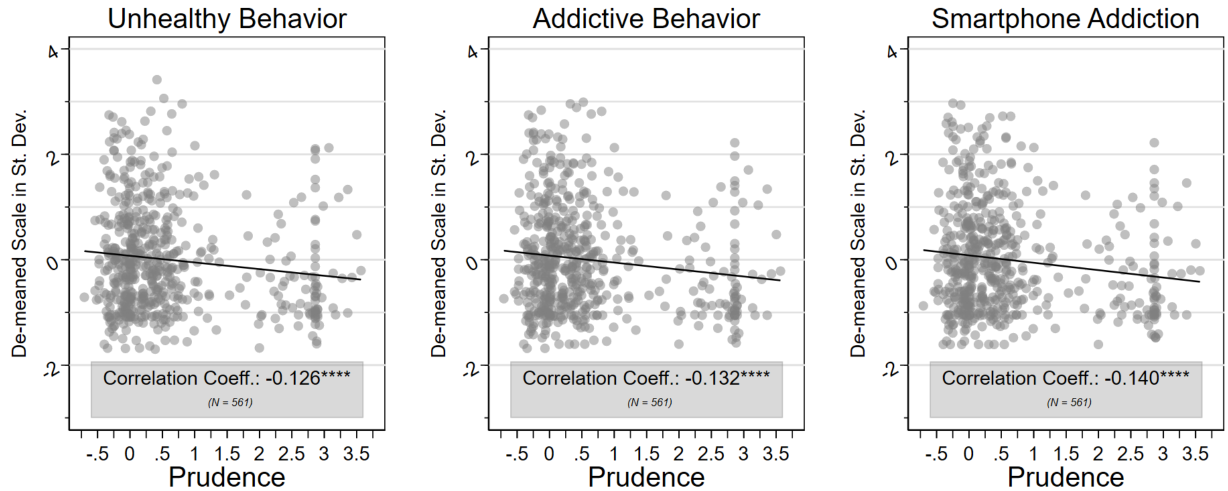


Figure D-5: Health-Related Behavior and (Sophisticated Measure of) Prudence p: Raw (Pairwise/Pearson) Correlations with Linear Fit

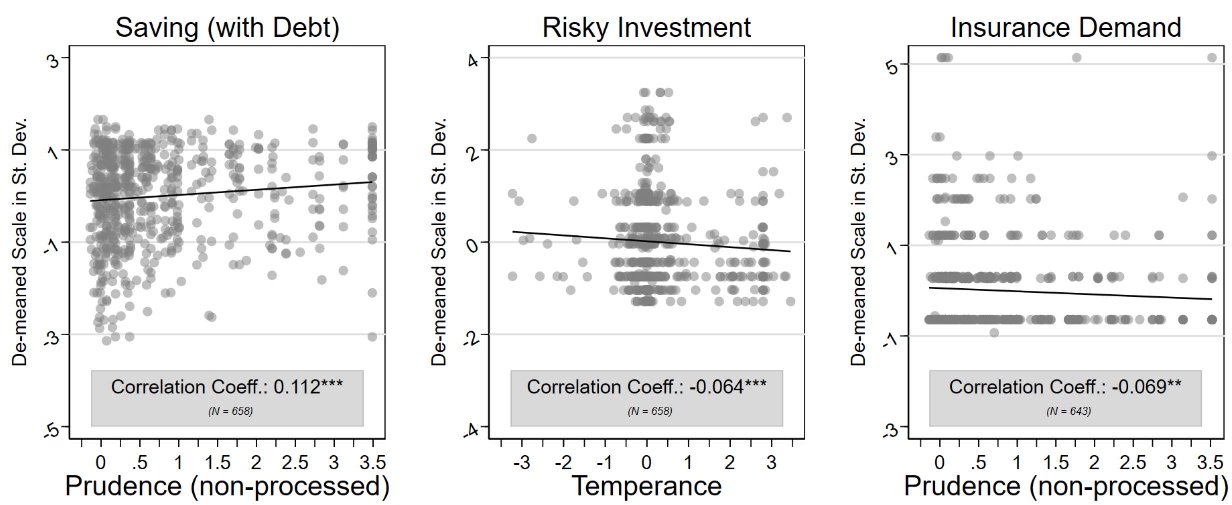


Figure D-6: Financial Decision-Making and Non-Processed, Simple Measure of Prudence p^{CE} (Sophisticated Measure of Temperance t): Raw (Pairwise/Pearson) Correlations with Linear Fit

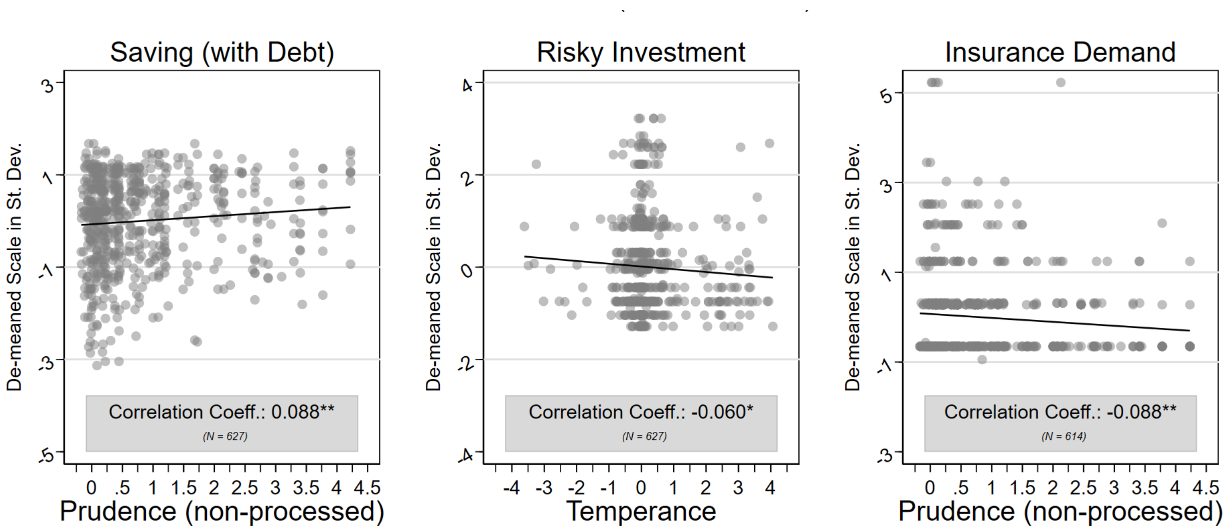


Figure D-7: Financial Decision-Making and Non-Processed, Simple Measure of Prudence p^{CE} (Sophisticated Measure of Temperance t): Raw (Pairwise/Pearson) Correlations with Linear Fit – When Excluding Observations Corresponding to Simple Patterns of Decision Making (c.f. Online Appendix Section B.1)

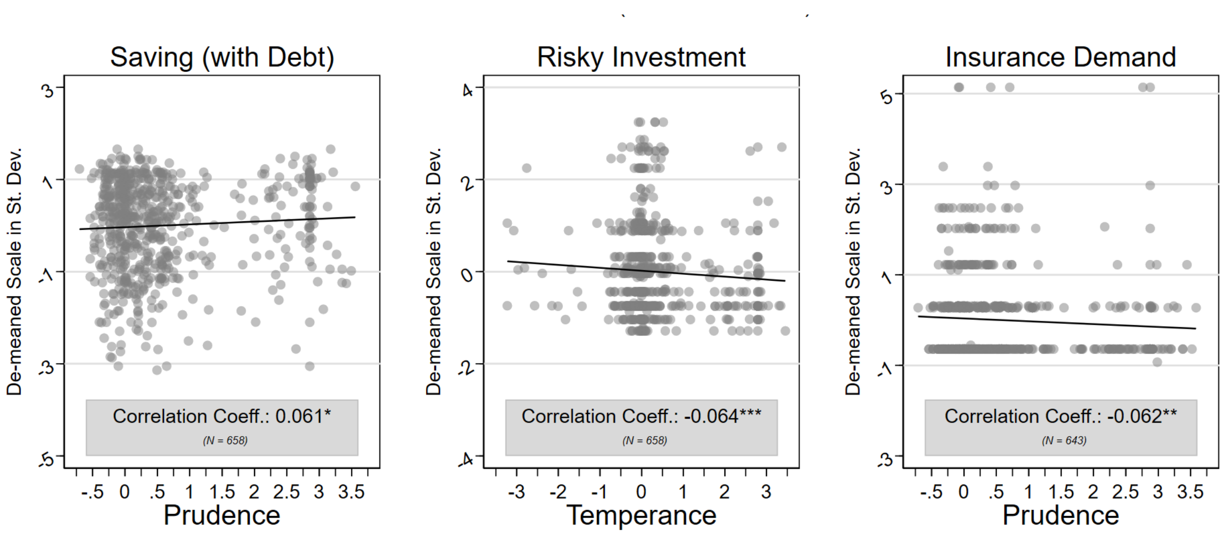


Figure D-8: Financial Decision-Making and (Sophisticated Measure of) Higher-Order Risk Preferences p and t: Raw (Pairwise/Pearson) Correlations with Linear Fit

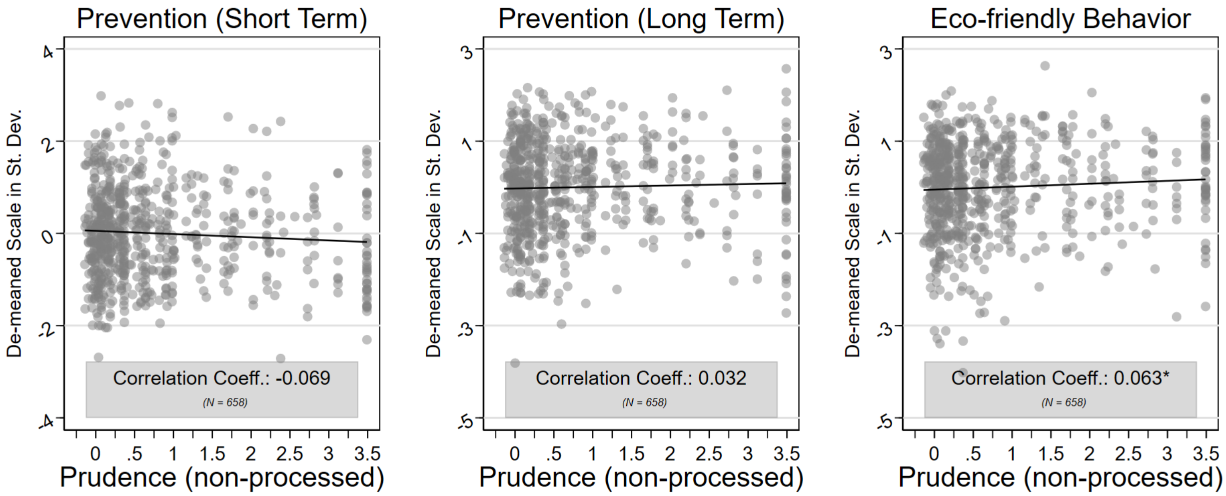


Figure D-9: Prevention, Eco-Friendly Behavior and Non-Processed, Simple Measure of Prudence p^{CE}: Raw (Pairwise/Pearson) Correlations with Linear Fit

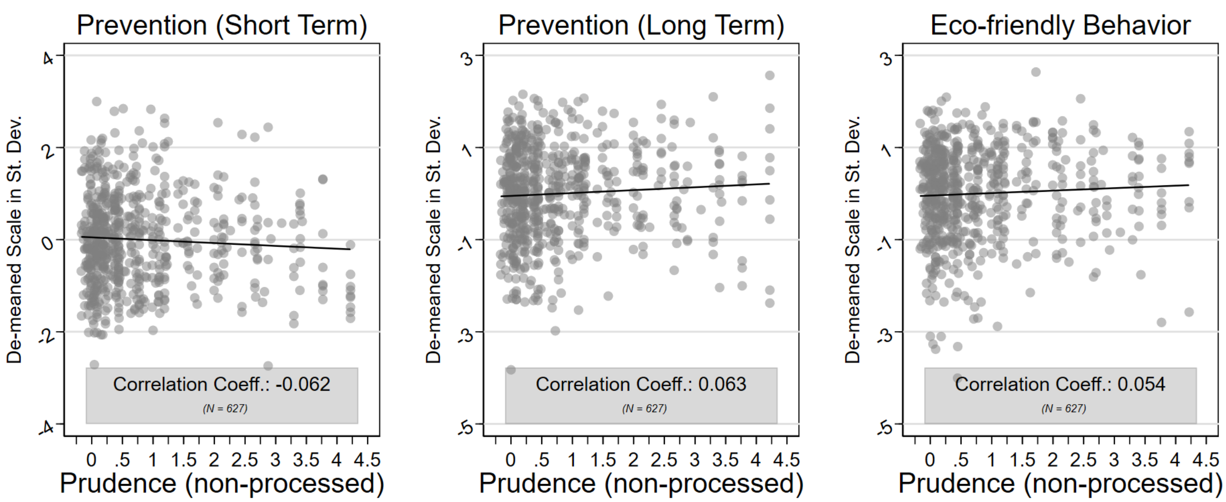


Figure D-10: Prevention, Eco-Friendly Behavior and Non-Processed, Simple Measure of Prudence p^{CE}: Raw (Pairwise/Pearson) Correlations with Linear Fit – When Excluding Observations Corresponding to Simple Patterns of Decision Making (c.f. Online Appendix Section B.1)

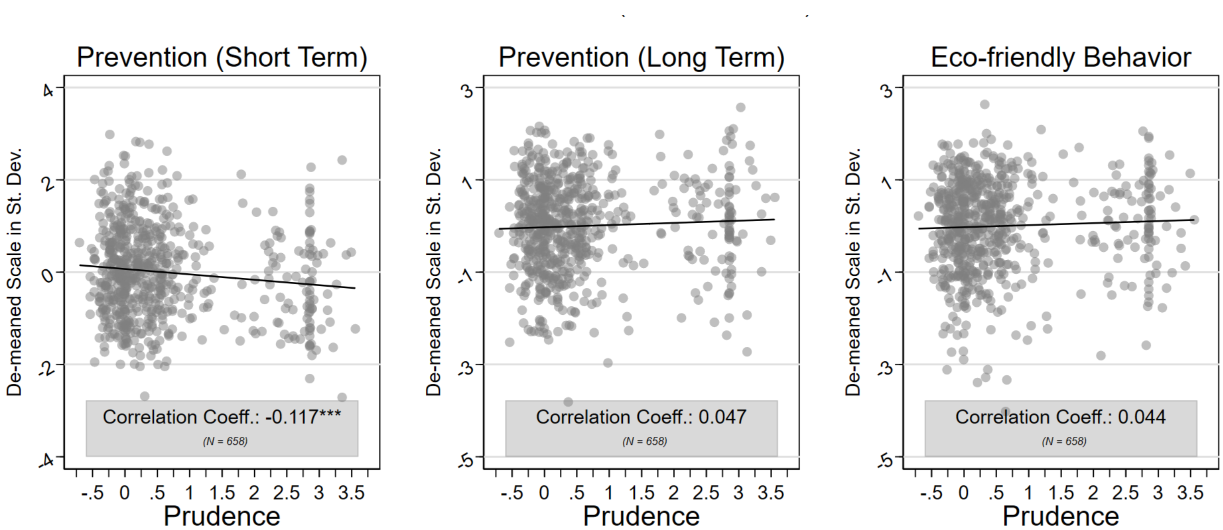


Figure D-11: Prevention, Eco-Friendly Behavior, and (Sophisticated Measure of) Prudence p: Raw (Pairwise/Pearson) Correlations with Linear Fit

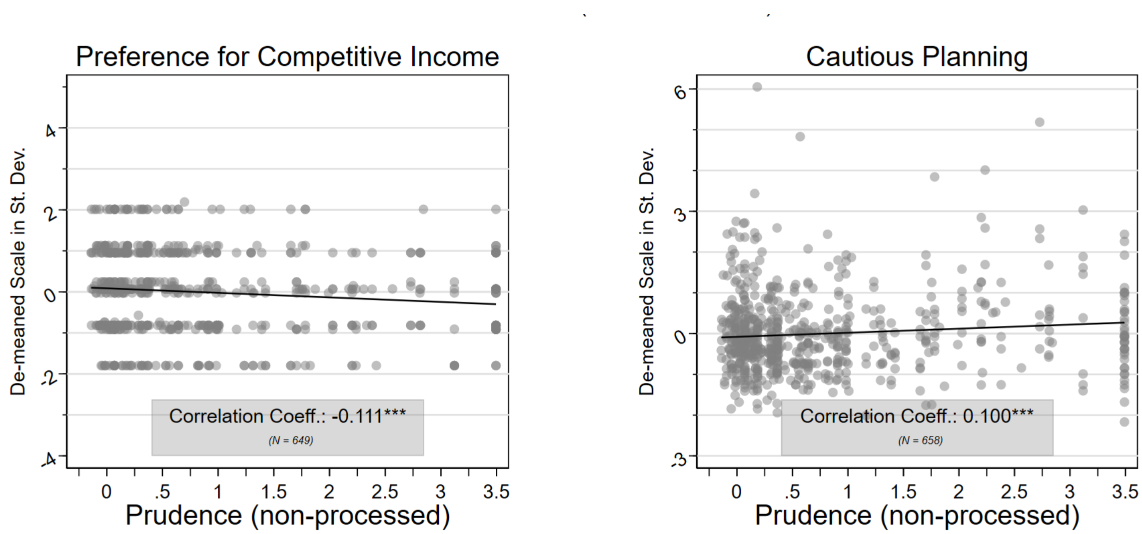


Figure D-12: Planning Behavior, Preference for a Competitive Income, and Non-Processed, Simple Measure of Prudence p^{CE}: Raw (Pairwise/Pearson) Correlations with Linear Fit

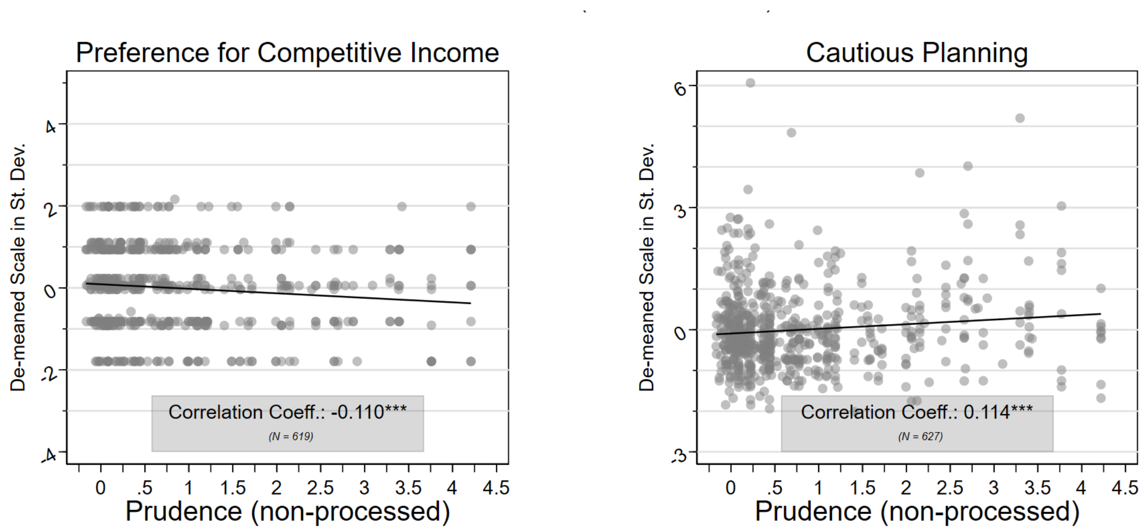


Figure D-13: Planning Behavior, Preference for a Competitive Income, and Non-Processed, Simple Measure of Prudence p^{CE}: Raw (Pairwise/Pearson) Correlations with Linear Fit – When Excluding Observations Corresponding to Simple Patterns of Decision Making (c.f. Online Appendix Section B.1)

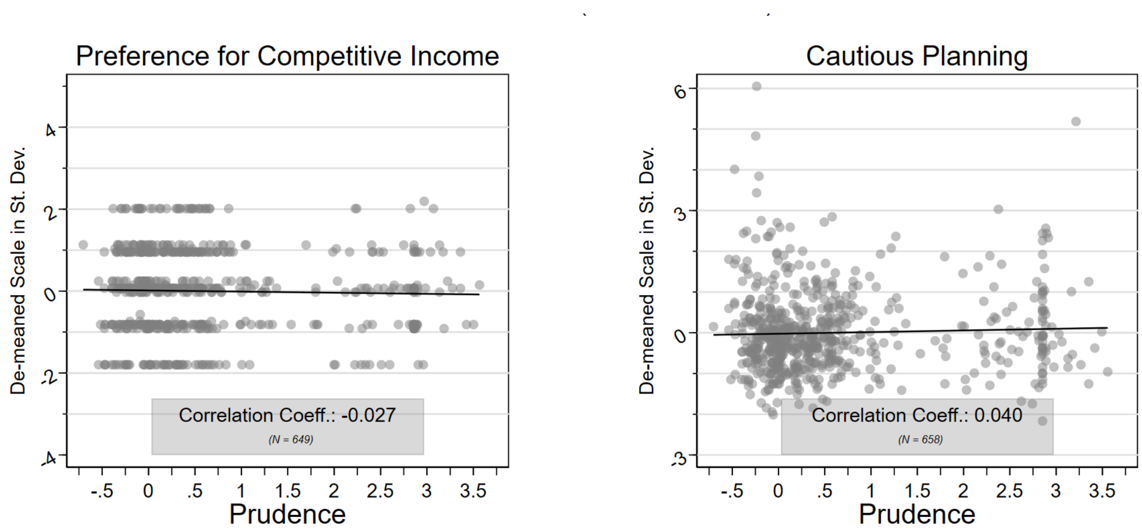


Figure D-14: Planning Behavior, Preference for a Competitive Income, and (Sophisticated Measure of) Prudence p: Raw (Pairwise/Pearson) Correlations with Linear Fit

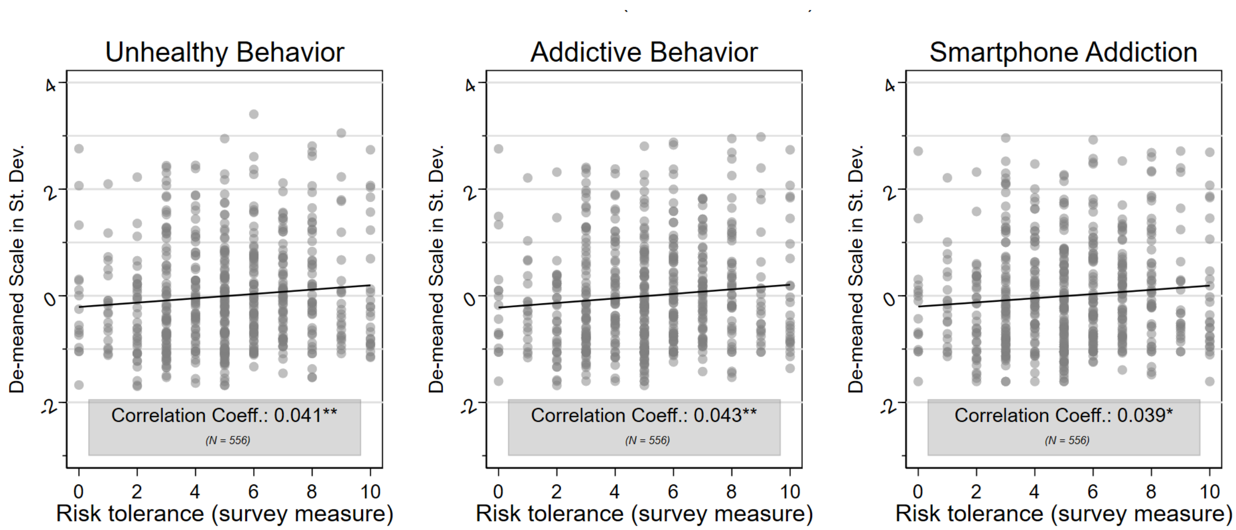


Figure D-15: Health-Related Behavior and Survey Measure on Willingness to Take Risks: Raw (Pairwise) Correlations (Regression Coefficient of Non-standardized Survey Measure) with Linear Fit

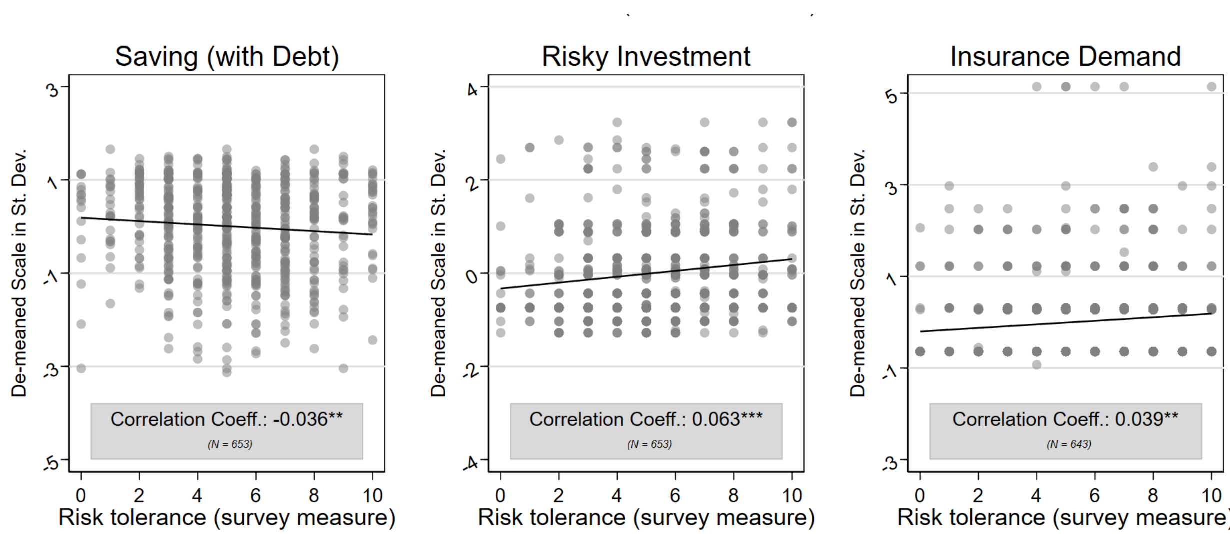


Figure D-16: Financial Decision Making and Survey Measure on Willingness to Take Risks: Raw (Pairwise) Correlations (Regression Coefficient of Non-standardized Survey Measure) with Linear Fit

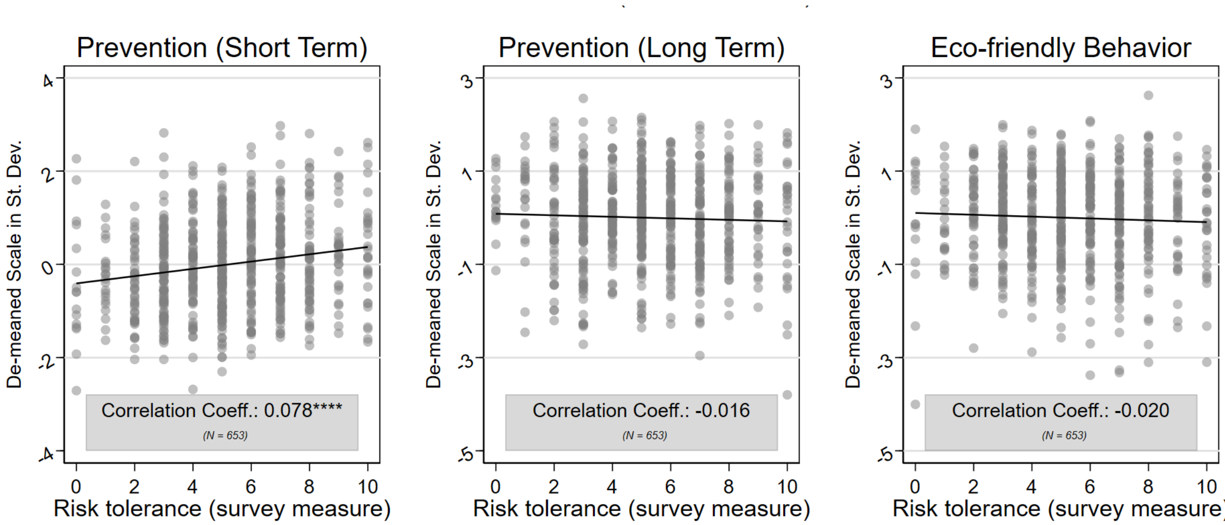


Figure D-17: Prevention, Eco-Friendly Behavior, and Survey Measure on Willingness to Take Risks: Raw (Pairwise) Correlations (Regression Coefficient of Non-standardized Survey Measure) with Linear Fit

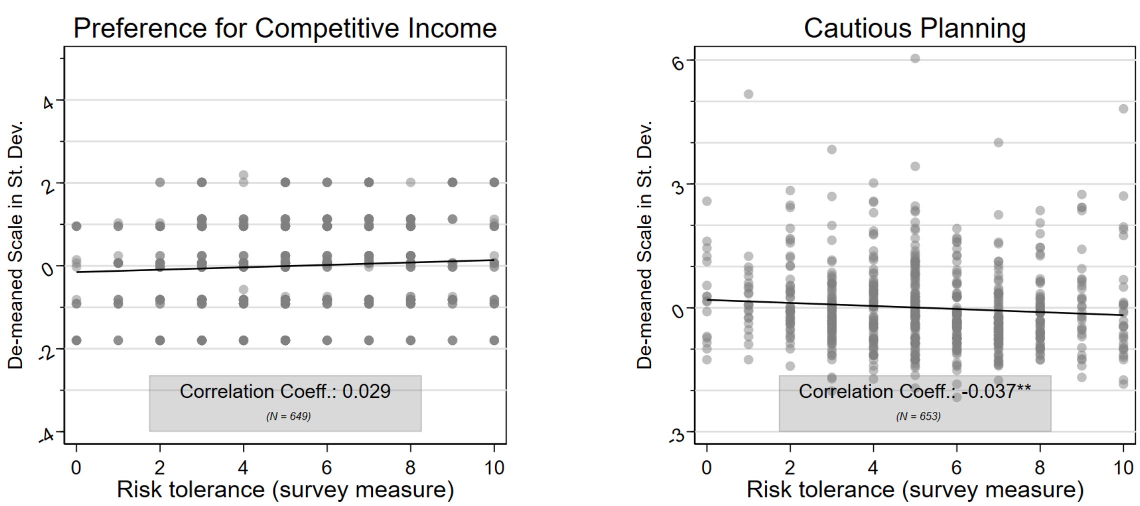


Figure D-18: Planning Behavior, Preference for a Competitive Income, and Survey Measure on Willingness to Take Risks: Raw (Pairwise) Correlations (Regression Coefficient of Non-standardized Survey Measure) with Linear Fit

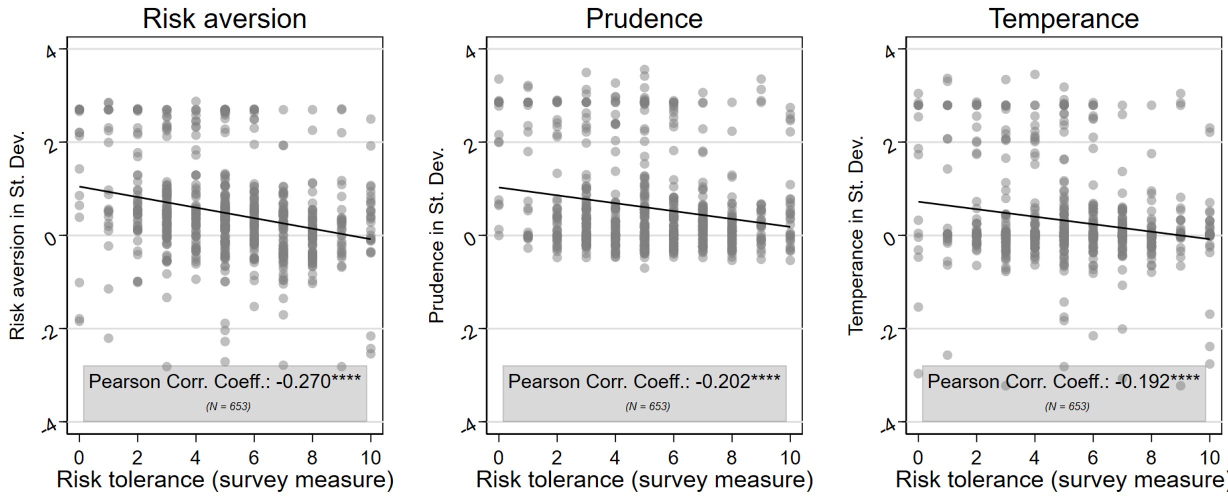


Figure D-19: Sophisticated Measures of Higher-Order Risk Preferences (r, p, and t) and the Survey Measure on Willingness to Take Risks: Raw (Pairwise/Pearson) Correlations with Linear Fit

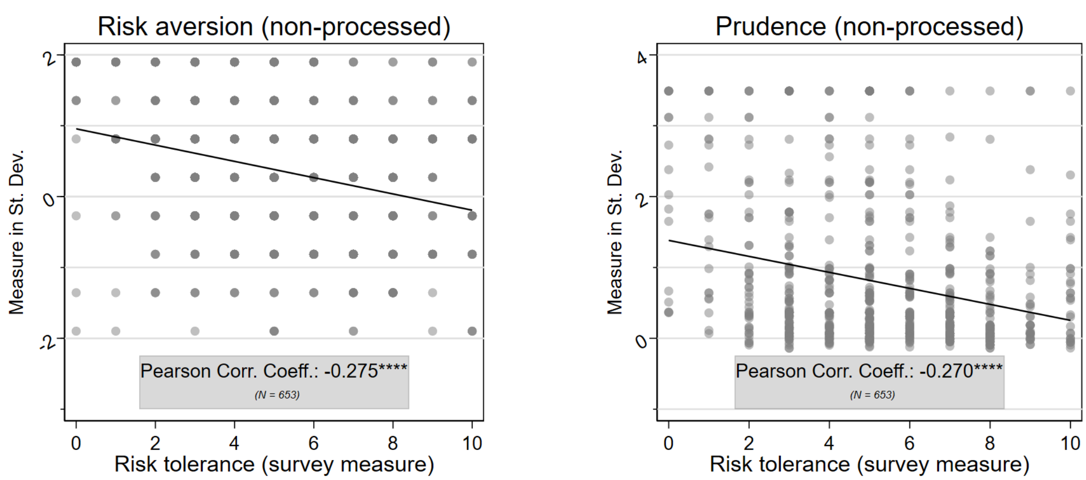


Figure D-20: Raw, Non-Processed Measures of Risk Aversion (r^{CE}) and Prudence (p^{CE}) and the Survey Measure on Willingness to Take Risks: Raw (Pairwise/Pearson) Correlations with Linear Fit

Table D-1: Characteristics of Participants: Age and Gender

Average Age (in years)	Grade	Total	Female	Male
11.6	6th	153	70	83
13.6	8th	169	80	89
15.7	10th	174	91	82
17.6	12th	162	89	73
Total		658	330	327

Table D-2: Validation of Elicitation Method: Identification of Prudence

	Sophisticated Measures			
	Prudence (p)		Risk Aversion (r)	
	Intensity♯	Classification†	Intensity♯	Classification†
# Prudent choices (Heinrich & Shachat)	0.295**** (0.017)	0.248**** (0.034)	−0.000 (0.017)	0.065* (0.033)
	Non-Processed Measures			
	(Simple) Prudence (p^{CE})		(Simple) Risk Aversion (r^{CE})	
	Intensity♯	Classification†	Intensity♯	Classification†
# Prudent choices (Heinrich & Shachat)	0.388**** (0.016)	0.621**** (0.048)	−0.008 (0.024)	0.047 (0.033)

Notes: Positive coefficients imply increasing intensity measures of prudence and risk aversion as resulting from our sophisticated and raw methods or an increased likelihood of being classified as prudent or risk-averse, respectively. Regression coefficients result from OLS regression (without constants) for intensities and probit regression for classifications. Intensity measures expressed in standard deviations. Bootstrapped standard errors (1000 reps.) clustered at the session level in parentheses.

♯For prudence, we use the (utility-based) intensity measure p by Crainich and Eeckhoudt (2008), and its simple approximation, p^{CE}. For risk, we use the Arrow-Pratt measure r and its simple, behavioral approximation, r^{CE}, as used in most of the empirical part of the paper, i.e., after having removed the variation explained by prudence. See Sections II.A and II.B for details and definition of the measures.

†We classify subjects as prudent and risk-averse if their respective intensity measure is positive, i.e., $p > 0$ and $r > 0$, respectively (and equivalently for the simple measures).

****/***/**/* denotes significance at the 0.1 / 1 / 5 / 10 percent level.

Table D-3: Influence Factors of Risk Aversion r (Arrow-Pratt Measure)

	Dependent Variable: Risk Aversion (r)							
	(1)		(2)		(3)		(4)	
Age (in years)	−0.031***	(0.010)	−0.016	(0.012)			−0.014	(0.015)
Cognitive ability			−0.114**	(0.041)	−0.125***	(0.038)	−0.132***	(0.043)
Female (=1)							0.259**	(0.091)
Impatience							−0.889***	(0.275)
Pocket money per week							−0.003	(0.002)
Math grade							−0.022	(0.051)
German grade							−0.018	(0.066)
Number of siblings							−0.010	(0.033)
Migration background (=1)							−0.039	(0.090)
Education mother: A-Levels (=1)							−0.108	(0.097)
Education father: A-Levels (=1)							−0.003	(0.105)
BMI							−0.012	(0.014)
Parents Occupation								
Full-time and part-time							0.066	(0.075)
One full-time							0.143	(0.137)
Don't work/other regularity							0.085	(0.102)
Religion								
Protestant							0.046	(0.118)
Other or no religion							−0.069	(0.097)
R^2	0.01		0.02		0.02		0.08	
Observations	658		658		658		658	

Notes: Positive coefficients imply increasing risk aversion, which is expressed in standard deviations. Cognitive-ability scores, relative German grade and relative math grade are standardized, such that above average scores are positive. Reference categories for parents' occupation is 'Both full-time', and 'Catholic' for religion. For 24 participants, some demographic information has been imputed with either 0, or the variable's mean value, or the 'other' category for binary, or continuous or categorical variables, respectively. We control for imputation with indicator variables. Robust standard errors clustered at the session level in parentheses. P-values for factors added only in the last column of this table except for gender are corrected for multiple testing using the Romano-Wolf procedure with 1,000 iterations (Romano and Wolf, 2005a,b, 2016; Clarke et al., 2020).

*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

Table D-4: Influence Factors of Prudence p (Crainich-Eeckhoudt Measure)

	Dependent Variable: Prudence (p)							
	(1)		(2)		(3)		(4)	
Age (in years)	−0.018	(0.013)	−0.010	(0.015)			−0.007	(0.019)
Cognitive ability			−0.056	(0.054)	−0.063	(0.048)	−0.062	(0.058)
Female (=1)							0.205*	(0.105)
Impatience							−0.584**	(0.234)
Pocket money per week							−0.006	(0.003)
Math grade							−0.029	(0.045)
German grade							−0.026	(0.064)
Number of siblings							0.022	(0.037)
Migration background (=1)							0.031	(0.076)
Education mother: A-Levels (=1)							−0.068	(0.091)
Education father: A-Levels (=1)							0.012	(0.108)
BMI							−0.008	(0.013)
Parents Occupation								
Full-time and part-time							0.071	(0.078)
One full-time							0.271	(0.136)
Don't work/other regularity							0.030	(0.119)
Religion								
Protestant							0.007	(0.101)
Other or no religion							−0.011	(0.113)
R^2	0.00		0.00		0.00		0.06	
Observations	658		658		658		658	

Notes: Positive coefficients imply increasing prudence, which is expressed in standard deviations. Cognitive-ability scores, relative German grade and relative math grade are standardized, such that above average scores are positive. Reference categories for parents' occupation is 'Both full-time', and 'Catholic' for religion. For 24 participants, some demographic information has been imputed with either 0, or the variable's mean value, or the 'other' category for binary, or continuous or categorical variables, respectively. We control for imputation with indicator variables. Robust standard errors clustered at the session level in parentheses. P-values for factors added only in the last column of this table except for gender are corrected for multiple testing using the Romano-Wolf procedure with 1,000 iterations (Romano and Wolf, 2005a,b, 2016; Clarke et al., 2020).

*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

Table D-5: Influence Factors of Temperance t (Denuit-Eeckhoudt Measure)

	Dependent Variable: Temperance (t)							
	(1)		(2)		(3)		(4)	
Age (in years)	−0.015	(0.011)	−0.001	(0.013)			0.002	(0.015)
Cognitive ability			−0.099**	(0.039)	−0.100**	(0.035)	−0.118**	(0.042)
Female (=1)							0.166*	(0.085)
Impatience							−0.703**	(0.283)
Pocket money per week							−0.003	(0.002)
Math grade							−0.029	(0.058)
German grade							0.019	(0.065)
Number of siblings							−0.002	(0.034)
Migration background (=1)							−0.093	(0.084)
Education mother: A-Levels (=1)							−0.113	(0.090)
Education father: A-Levels (=1)							0.032	(0.100)
BMI							−0.004	(0.012)
Parents Occupation								
Full-time and part-time							0.124	(0.074)
One full-time							0.209	(0.125)
Don't work/other regularity							0.166	(0.118)
Religion								
Protestant							−0.048	(0.105)
Other or no religion							−0.079	(0.109)
R^2	0.00		0.01		0.01		0.06	
Observations	658		658		658		658	

Notes: Positive coefficients imply increasing temperance, which is expressed in standard deviations. Cognitive-ability scores, relative German grade and relative math grade are standardized, such that above average scores are positive. Reference categories for parents' occupation is 'Both full-time', and 'Catholic' for religion. For 24 participants, some demographic information has been imputed with either 0, or the variable's mean value, or the 'other' category for binary, or continuous or categorical variables, respectively. We control for imputation with indicator variables. Robust standard errors clustered at the session level in parentheses. P-values for factors added only in the last column of this table except for gender are corrected for multiple testing using the Romano-Wolf procedure with 1,000 iterations (Romano and Wolf, 2005a,b, 2016; Clarke et al., 2020).

*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

Table D-6: Influence Factors of (Simple Measures of) Risk Aversion (r^{CE}) and Prudence (p^{CE})

	[1] Risk Aversion (r^{CE})		[2] Prudence (p^{CE})	
Age (in years)	−0.022	(0.016)	−0.009	(0.012)
Cognitive ability	−0.110**	(0.038)	−0.091	(0.059)
Female (=1)	0.310***	(0.085)	0.208**	(0.089)
Impatience	−0.982***	(0.257)	−0.813***	(0.312)
Other Factors	10		10	
R^2	0.10		0.09	
Observations	658		658	

Notes: OLS regressions of (higher-order) risk preferences using simple, raw measures of risk aversion and prudence (see Section II.A). Positive coefficients imply increasing risk aversion, and prudence, which are expressed in standard deviations. Cognitive-ability scores are standardized, such that above average scores are positive. Other possible influence factors controlled for are relative math grade, relative German grade (where positive variables imply above average performance relative to the grade), the amount of pocket money per week, relative BMI, the number of siblings, religion, migration background, parents' education, as well as their occupation; see Tables D-7, and D-8 for detailed regressions results. Robust standard errors clustered at the session level in parentheses.
****/***/**/* denotes significance at the 0.1 / 1 / 5 / 10 percent level.

Table D-7: Influence Factors of (the Simple Measure of) Risk Aversion (r^{CE})

	Dependent Variable: Risk Aversion (r^{CE})							
	(1)		(2)		(3)		(4)	
Age (in years)	−0.036*	(0.017)	−0.023	(0.019)			−0.022	(0.016)
Cognitive ability			−0.095*	(0.045)	−0.112**	(0.044)	−0.110**	(0.038)
Female (=1)							0.310***	(0.085)
Impatience							−0.982***	(0.257)
Pocket money per week							−0.002	(0.003)
Math grade							−0.023	(0.043)
German grade							−0.046	(0.052)
Number of siblings							−0.018	(0.030)
Migration background (=1)							−0.010	(0.087)
Education mother: A-Levels (=1)							−0.059	(0.092)
Education father: A-Levels (=1)							−0.041	(0.088)
BMI							−0.022	(0.014)
Parents Occupation								
Full-time and part-time							0.004	(0.071)
One full-time							0.063	(0.125)
Don't work/other regularity							0.079	(0.117)
Religion								
Protestant							0.062	(0.120)
Other or no religion							−0.112	(0.073)
R^2	0.01		0.01		0.01		0.10	
Observations	658		658		658		658	

Notes: Positive coefficients imply increasing risk aversion using a simple, raw measure of risk aversion (see Section II.A), which is expressed in standard deviations. Cognitive-ability scores, relative German grade and relative math grade are standardized, such that above average scores are positive. Reference categories for parents' occupation is 'Both full-time', and 'Catholic' for religion. For 24 participants, some demographic information has been imputed with either 0, or the variable's mean value, or the 'other' category for binary, or continuous or categorical variables, respectively. We control for imputation with indicator variables. Robust standard errors clustered at the session level in parentheses. P-values for factors added only in the last column of this table except for gender are corrected for multiple testing using the Romano-Wolf procedure with 1,000 iterations (Romano and Wolf, 2005a,b, 2016; Clarke et al., 2020).

*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

Table D-8: Influence Factors of (the Simple Measure of) Prudence (p^{CE})

	Dependent Variable: Prudence (p^{CE})							
	(1)		(2)		(3)		(4)	
Age (in years)	−0.027**	(0.012)	−0.016	(0.014)			−0.009	(0.012)
Cognitive ability			−0.081	(0.057)	−0.092	(0.053)	−0.091	(0.059)
Female (=1)							0.208**	(0.089)
Impatience							−0.813***	(0.312)
Pocket money per week							−0.005	(0.002)
Math grade							0.021	(0.035)
German grade							−0.020	(0.063)
Number of siblings							−0.026	(0.026)
Migration background (=1)							−0.028	(0.094)
Education mother: A-Levels (=1)							0.039	(0.087)
Education father: A-Levels (=1)							−0.062	(0.115)
BMI							−0.000	(0.015)
Parents Occupation								
Full-time and part-time							0.111	(0.088)
One full-time							0.333**	(0.137)
Don't work/other regularity							0.051	(0.096)
Religion								
Protestant							−0.011	(0.091)
Other or no religion							−0.086	(0.111)
R^2	0.00		0.01		0.01		0.09	
Observations	658		658		658		658	

Notes: Positive coefficients imply increasing prudence using a simple, raw measure of prudence (see Section II.A), which is expressed in standard deviations. Cognitive-ability scores, relative German grade and relative math grade are standardized, such that above average scores are positive. Reference categories for parents' occupation is 'Both full-time', and 'Catholic' for religion. For 24 participants, some demographic information has been imputed with either 0, or the variable's mean value, or the 'other' category for binary, or continuous or categorical variables, respectively. We control for imputation with indicator variables. Robust standard errors clustered at the session level in parentheses. P-values for factors added only in the last column of this table except for gender are corrected for multiple testing using the Romano-Wolf procedure with 1,000 iterations (Romano and Wolf, 2005a,b, 2016; Clarke et al., 2020).

*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

Table D-9: Unhealthy Behavior

	Dependent Variable: Unhealthy Behavior							
	(1)		(2)		(3)		(4)	
Risk aversion (r)			0.030	(0.088)	0.053	(0.087)	0.024	(0.086)
Prudence (p)	−0.136****	(0.035)	−0.136****	(0.034)	−0.123****	(0.034)	−0.140****	(0.025)
Temperance (t)			−0.022	(0.062)	−0.008	(0.061)	−0.011	(0.047)
Impatience					0.161****	(0.038)	0.116***	(0.037)
Cognitive ability							−0.015	(0.039)
Age (in years)							0.011	(0.020)
Female (=1)							0.265**	(0.103)
Pocket money per week							0.006	(0.004)
Math grade							−0.113**	(0.045)
German grade							−0.171****	(0.033)
Number of siblings							−0.009	(0.034)
Migration background (=1)							0.163	(0.107)
Education mother: A-Levels (=1)							0.024	(0.058)
Education father: A-Levels (=1)							−0.010	(0.097)
BMI							0.033*	(0.020)
Parents Occupation								
Full-time and part-time							−0.102	(0.125)
One full-time							−0.038	(0.205)
Don't work/other regularity							0.133	(0.180)
Religion								
Protestant							−0.180*	(0.104)
Other or no religion							−0.121	(0.105)
R^2	0.02		0.02		0.04		0.14	
Observations	561		561		561		561	

Notes: Positive coefficients imply increasing engagement in unhealthy or addictive behavior. Risk and time measures are expressed in standard deviations. Non-orthogonalized measure: Prudence. Cognitive-ability scores, relative German grade and relative math grade are standardized, such that above average scores are positive. Reference categories for parents' occupation is 'Both full-time', and 'Catholic' for religion. Missing demographic information has been imputed and controlled for. Outcome indices are formed using PCA weights and are standard normalized. Questions included in this index are listed in Section E.3 of the questionnaire in Online Appendix E. Bootstrapped standard errors (1000 repetitions) clustered at the session level in parentheses.
****/***/**/* denotes significance at the 0.1 / 1 / 5 / 10 percent level.

Table D-10: Addictive Behavior

	Dependent Variable: Addictive Behavior							
	(1)		(2)		(3)		(4)	
Risk aversion (r)			0.016	(0.087)	0.039	(0.086)	0.007	(0.087)
Prudence (p)	−0.142****	(0.034)	−0.142****	(0.034)	−0.129****	(0.033)	−0.146****	(0.025)
Temperance (t)			−0.016	(0.060)	−0.001	(0.059)	−0.002	(0.047)
Impatience					0.163****	(0.038)	0.122***	(0.038)
Cognitive ability							−0.009	(0.039)
Age (in years)							0.009	(0.020)
Female (=1)							0.264**	(0.103)
Pocket money per week							0.006	(0.004)
Math grade							−0.116***	(0.044)
German grade							−0.156****	(0.032)
Number of siblings							−0.006	(0.034)
Migration background (=1)							0.172*	(0.102)
Education mother: A-Levels (=1)							0.015	(0.060)
Education father: A-Levels (=1)							−0.017	(0.095)
BMI							0.016	(0.018)
Parents Occupation								
Full-time and part-time							−0.112	(0.125)
One full-time							−0.020	(0.201)
Don't work/other regularity							0.126	(0.183)
Religion								
Protestant							−0.178*	(0.102)
Other or no religion							−0.123	(0.101)
R^2	0.02		0.02		0.04		0.13	
Observations	561		561		561		561	

Notes: Positive coefficients imply increasing engagement in addictive behavior. Risk and time measures are expressed in standard deviations. Non-orthogonalized measure: Prudence. Cognitive-ability scores, relative German grade and relative math grade are standardized, such that above average scores are positive. Reference categories for parents' occupation is 'Both full-time', and 'Catholic' for religion. Missing demographic information has been imputed and controlled for. Outcome indices are formed using PCA weights and are standard normalized. Questions included in this index are listed in Section E.3 of the questionnaire in Online Appendix E. Bootstrapped standard errors (1000 repetitions) clustered at the session level in parentheses. ****/***/**/* denotes significance at the 0.1 / 1 / 5 / 10 percent level.

Table D-11: Addictive Usage of Smartphone and Social Media

	Dependent Variable: Smartphone Addiction							
	(1)		(2)		(3)		(4)	
Risk aversion (r)			0.002	(0.089)	0.024	(0.087)	−0.023	(0.086)
Prudence (p)	−0.152****	(0.033)	−0.152****	(0.033)	−0.139****	(0.032)	−0.160****	(0.025)
Temperance (t)			0.001	(0.052)	0.015	(0.052)	0.017	(0.045)
Impatience					0.154****	(0.036)	0.115***	(0.037)
Cognitive ability							−0.001	(0.040)
Age (in years)							0.008	(0.020)
Female (=1)							0.314***	(0.101)
Pocket money per week							0.004	(0.004)
Math grade							−0.093**	(0.043)
German grade							−0.139****	(0.036)
Number of siblings							−0.000	(0.036)
Migration background (=1)							0.224**	(0.100)
Education mother: A-Levels (=1)							−0.017	(0.056)
Education father: A-Levels (=1)							0.004	(0.083)
BMI							0.008	(0.017)
Parents Occupation								
Full-time and part-time							−0.094	(0.122)
One full-time							−0.016	(0.201)
Don't work/other regularity							0.176	(0.192)
Religion								
Protestant							−0.159*	(0.094)
Other or no religion							−0.109	(0.101)
R^2	0.02		0.02		0.04		0.12	
Observations	561		561		561		561	

Notes: Positive coefficients imply increasing engagement in addictive behavior. Risk and time measures are expressed in standard deviations. Non-orthogonalized measure: Prudence. Cognitive-ability scores, relative German grade and relative math grade are standardized, such that above average scores are positive. Reference categories for parents' occupation is 'Both full-time', and 'Catholic' for religion. Missing demographic information has been imputed and controlled for. Outcome indices are formed using PCA weights and are standard normalized. Questions included in this index are listed in Section E.3 of the questionnaire in Online Appendix E. Bootstrapped standard errors (1000 repetitions) clustered at the session level in parentheses. ****/***/**/* denotes significance at the 0.1 / 1 / 5 / 10 percent level.

Table D-12: Saving (with Debt)

	Dependent Variable: Saving (with Debt)							
	(1)		(2)		(3)		(4)	
Risk aversion (r)			0.093	(0.114)	0.060	(0.111)	0.096	(0.105)
Prudence (p)	0.060*	(0.031)	0.060**	(0.030)	0.040	(0.030)	0.058*	(0.034)
Temperance (t)			0.110**	(0.048)	0.090*	(0.047)	0.087**	(0.035)
Impatience					−0.222***	(0.034)	−0.197***	(0.032)
Pocket money risky +							−0.063	(0.044)
Pocket money risky −							−0.037	(0.133)
Earnings side job stable							0.099**	(0.046)
Earnings side job per week							0.001	(0.002)
Earnings side job self-det.							0.049	(0.050)
Cognitive ability							−0.021	(0.036)
Age (in years)							−0.067**	(0.028)
Female (=1)							−0.323***	(0.068)
Pocket money per week							−0.008***	(0.002)
Math grade							0.118**	(0.053)
German grade							0.091*	(0.049)
Number of siblings							−0.003	(0.030)
Migration background (=1)							−0.003	(0.075)
Education mother: A-Levels (=1)							−0.017	(0.057)
Education father: A-Levels (=1)							0.104	(0.085)
BMI							−0.031**	(0.014)
Parents Occupation								
Full-time and part-time							0.092	(0.084)
One full-time							0.112	(0.149)
Don't work/other regularity							−0.043	(0.107)
Religion								
Protestant							0.018	(0.058)
Other or no religion							−0.173**	(0.085)
R^2	0.00		0.01		0.06		0.19	
Observations	658		658		658		658	

Notes: Positive coefficients imply increasing likelihood to save. Risk and time measures are expressed in standard deviations. Non-orthogonalized measure: Prudence. Cognitive-ability scores, relative German grade and relative math grade are standardized, such that above average scores are positive. Reference categories for parents' occupation is 'Both full-time', and 'Catholic' for religion. Missing demographic information has been imputed and controlled for. Additional influence factors compared to those used in Section III control for income risk, as the saving theory that we rely on here is about precautionary saving; see our empirical strategy (Section E.2) for details. Outcome indices are formed using PCA weights and are standard normalized. Questions included in this index are listed in Section E.4 of the questionnaire in Online Appendix E. Bootstrapped standard errors (1000 repetitions) clustered at the session level in parentheses.
****/***/**/* denotes significance at the 0.1 / 1 / 5 / 10 percent level.

Table D-13: Risky Investment

	Dependent Variable: Risky Investment							
	(1)		(2)		(3)		(4)	
Risk aversion (r)			−0.161**	(0.074)	−0.159**	(0.075)	−0.066	(0.070)
Prudence (p)			−0.103**	(0.048)	−0.102**	(0.048)	−0.045	(0.043)
Temperance (t)	−0.064***	(0.023)	−0.064***	(0.024)	−0.062***	(0.023)	−0.032**	(0.016)
Impatience					0.013	(0.025)	0.009	(0.026)
Pocket money risky +							0.083*	(0.050)
Pocket money risky −							−0.031	(0.120)
Earnings side job stable							0.109**	(0.049)
Earnings side job per week							0.003*	(0.002)
Earnings side job self-det.							0.054	(0.086)
Cognitive ability							−0.033	(0.044)
Age (in years)							−0.025	(0.026)
Female (=1)							−0.540***	(0.057)
Pocket money per week							0.007**	(0.004)
Math grade							0.073**	(0.029)
German grade							−0.034	(0.036)
Number of siblings							−0.054**	(0.023)
Migration background (=1)							0.011	(0.064)
Education mother: A-Levels (=1)							0.236***	(0.066)
Education father: A-Levels (=1)							0.196**	(0.085)
BMI							0.007	(0.011)
Parents Occupation								
Full-time and part-time							0.103	(0.079)
One full-time							−0.111	(0.113)
Don't work/other regularity							−0.125	(0.099)
Religion								
Protestant							−0.060	(0.090)
Other or no religion							−0.216***	(0.047)
R^2	0.00		0.02		0.02		0.18	
Observations	658		658		658		658	

Notes: Positive coefficients imply increasing likelihood to invest in risky assets. Risk and time measures are expressed in standard deviations. Non-orthogonalized measure: Temperance. Cognitive-ability scores, relative German grade and relative math grade are standardized, such that above average scores are positive. Reference categories for parents' occupation is 'Both full-time', and 'Catholic' for religion. Missing demographic information has been imputed and controlled for. Additional influence factors compared to those used in Section III control for income risk, as the theory that we rely on here is about investment when facing background risks; see our empirical strategy (Section E.2) for details. Outcome indices are formed using PCA weights and are standard normalized. Questions included in this index are listed in Section E.4 of the questionnaire in Online Appendix E. Bootstrapped standard errors (1000 repetitions) clustered at the session level in parentheses.
****/***/**/* denotes significance at the 0.1 / 1 / 5 / 10 percent level.

Table D-14: Financial Insurance Demand

	Dependent Variable: Insurance Demand							
	(1)		(2)		(3)		(4)	
Risk aversion (r)			−0.017	(0.049)	−0.014	(0.050)	−0.018	(0.044)
Prudence (p)	−0.039**	(0.017)	−0.039**	(0.016)	−0.037**	(0.017)	−0.023	(0.024)
Temperance (t)			−0.010	(0.042)	−0.008	(0.042)	−0.001	(0.046)
Impatience					0.018	(0.026)	−0.003	(0.025)
Cognitive ability							0.017	(0.015)
Age (in years)							−0.036***	(0.010)
Female (=1)							−0.070	(0.047)
Pocket money per week							0.005***	(0.002)
Math grade							−0.025	(0.022)
German grade							0.019	(0.037)
Number of siblings							−0.022	(0.017)
Migration background (=1)							0.101***	(0.036)
Education mother: A-Levels (=1)							−0.012	(0.037)
Education father: A-Levels (=1)							0.008	(0.046)
BMI							0.011	(0.007)
Parents Occupation								
Full-time and part-time							−0.081	(0.060)
One full-time							−0.109	(0.072)
Don't work/other regularity							−0.249***	(0.070)
Religion								
Protestant							−0.021	(0.063)
Other or no religion							−0.065	(0.072)
R^2	0.00		0.00		0.01		0.06	
Observations	643		643		643		643	

Notes: Positive coefficients imply increasing likelihood to possess an insurance. Risk and time measures are expressed in standard deviations. Non-orthogonalized measure: Prudence. Cognitive-ability scores, relative German grade and relative math grade are standardized, such that above average scores are positive. Reference categories for parents' occupation is 'Both full-time', and 'Catholic' for religion. Missing demographic information has been imputed and controlled for. Outcome indices are formed using PCA weights and are standard normalized. Questions included in this index are listed in Section E.4 of the questionnaire in Online Appendix E. Bootstrapped standard errors (1000 repetitions) clustered at the session level in parentheses. ****/***/**/* denotes significance at the 0.1 / 1 / 5 / 10 percent level.

Table D-15: Prevention (Short-Term)

	Dependent Variable: Prevention (Short Term)							
	(1)		(2)		(3)		(4)	
Risk aversion (r)			0.168**	(0.079)	0.178**	(0.078)	0.120*	(0.063)
Prudence (p)	−0.117***	(0.036)	−0.117***	(0.037)	−0.110***	(0.038)	−0.119****	(0.033)
Temperance (t)			0.025	(0.038)	0.032	(0.038)	0.017	(0.039)
Impatience					0.073***	(0.025)	0.055**	(0.025)
Cognitive ability							−0.048	(0.030)
Age (in years)							−0.011	(0.024)
Female (=1)							0.257****	(0.058)
Pocket money per week							0.007****	(0.002)
Math grade							−0.096**	(0.043)
German grade							−0.055	(0.058)
Number of siblings							−0.041*	(0.024)
Migration background (=1)							0.039	(0.095)
Education mother: A-Levels (=1)							−0.077	(0.053)
Education father: A-Levels (=1)							−0.026	(0.075)
BMI							−0.022	(0.016)
Parents Occupation								
Full-time and part-time							−0.252**	(0.128)
One full-time							−0.309****	(0.088)
Don't work/other regularity							−0.176	(0.131)
Religion								
Protestant							−0.098	(0.082)
Other or no religion							−0.280***	(0.105)
R^2	0.01		0.02		0.03		0.10	
Observations	658		658		658		658	

Notes: Positive coefficients imply increasing prevention effort. Risk and time measures are expressed in standard deviations. Non-orthogonalized measure: Prudence. Cognitive-ability scores, relative German grade and relative math grade are standardized, such that above average scores are positive. Reference categories for parents' occupation is 'Both full-time', and 'Catholic' for religion. Missing demographic information has been imputed and controlled for. Outcome indices are formed using PCA weights and are standard normalized. Questions included in this index are listed in Section E.5 in the questionnaire in Online Appendix E. Bootstrapped standard errors (1000 repetitions) clustered at the session level in parentheses.
****/***/**/* denotes significance at the 0.1 / 1 / 5 / 10 percent level.

Table D-16: Prevention (Long-Term)

	Dependent Variable: Prevention (Long Term)							
	(1)		(2)		(3)		(4)	
Risk aversion (r)			0.245***	(0.077)	0.233***	(0.077)	0.167***	(0.060)
Prudence (p)	0.047	(0.056)	0.047	(0.053)	0.040	(0.054)	0.010	(0.042)
Temperance (t)			0.129***	(0.046)	0.121***	(0.047)	0.078**	(0.034)
Impatience					−0.081**	(0.034)	−0.088***	(0.032)
Cognitive ability							−0.170****	(0.026)
Age (in years)							0.008	(0.022)
Female (=1)							0.391****	(0.083)
Pocket money per week							0.001	(0.003)
Math grade							−0.087*	(0.049)
German grade							0.071*	(0.043)
Number of siblings							−0.007	(0.023)
Migration background (=1)							−0.126	(0.080)
Education mother: A-Levels (=1)							−0.132**	(0.065)
Education father: A-Levels (=1)							0.009	(0.052)
BMI							−0.023*	(0.014)
Parents Occupation								
Full-time and part-time							0.013	(0.072)
One full-time							−0.052	(0.089)
Don't work/other regularity							0.017	(0.127)
Religion								
Protestant							−0.058	(0.095)
Other or no religion							−0.209	(0.128)
R^2	0.00		0.03		0.03		0.13	
Observations	658		658		658		658	

Notes: Positive coefficients imply increasing prevention effort. Risk and time measures are expressed in standard deviations. Non-orthogonalized measure: Prudence. Cognitive-ability scores, relative German grade and relative math grade are standardized, such that above average scores are positive. Reference categories for parents' occupation is 'Both full-time', and 'Catholic' for religion. Missing demographic information has been imputed and controlled for. Outcome indices are formed using PCA weights and are standard normalized. Questions included in this index are listed in Section E.5 in the questionnaire in Online Appendix E. Bootstrapped standard errors (1000 repetitions) clustered at the session level in parentheses.
****/***/**/* denotes significance at the 0.1 / 1 / 5 / 10 percent level.

Table D-17: Eco-friendly behavior

	Dependent Variable: Eco-friendly Behavior							
	(1)		(2)		(3)		(4)	
Risk aversion (r)			0.208**	(0.100)	0.188**	(0.095)	0.201**	(0.081)
Prudence (p)	0.044	(0.035)	0.044	(0.036)	0.031	(0.034)	0.022	(0.027)
Temperance (t)			0.061	(0.064)	0.048	(0.063)	0.034	(0.057)
Impatience					−0.140****	(0.027)	−0.086***	(0.029)
Cognitive ability							0.011	(0.040)
Age (in years)							0.004	(0.022)
Female (=1)							0.100	(0.077)
Pocket money per week							−0.011****	(0.002)
Math grade							0.044	(0.028)
German grade							0.188****	(0.055)
Number of siblings							−0.004	(0.023)
Migration background (=1)							−0.172***	(0.063)
Education mother: A-Levels (=1)							−0.060	(0.055)
Education father: A-Levels (=1)							−0.002	(0.061)
BMI							−0.020**	(0.010)
Parents Occupation								
Full-time and part-time							0.045	(0.068)
One full-time							0.022	(0.075)
Don't work/other regularity							−0.056	(0.145)
Religion								
Protestant							−0.027	(0.106)
Other or no religion							−0.082	(0.142)
R^2	0.00		0.01		0.03		0.12	
Observations	658		658		658		658	

Notes: Positive coefficients imply increasing eco-friendly behavior. Risk and time measures are expressed in standard deviations. Non-orthogonalized measure: Prudence. Cognitive-ability scores, relative German grade and relative math grade are standardized, such that above average scores are positive. Reference categories for parents' occupation is 'Both full-time', and 'Catholic' for religion. Missing demographic information has been imputed and controlled for. Outcome indices are formed using PCA weights and are standard normalized. Questions included in this index are listed in Section E.5 in the questionnaire in Online Appendix E. Bootstrapped standard errors (1000 repetitions) clustered at the session level in parentheses. ****/***/**/* denotes significance at the 0.1 / 1 / 5 / 10 percent level.

Table D-18: Planning Behavior

	Dependent Variable: Cautious Planning							
	(1)		(2)		(3)		(4)	
Risk aversion (r)			−0.033	(0.048)	−0.034	(0.048)	−0.063	(0.050)
Prudence (p)	0.040	(0.047)	0.040	(0.045)	0.040	(0.043)	0.028	(0.051)
Temperance (t)			0.166****	(0.050)	0.166****	(0.049)	0.170****	(0.045)
Impatience					−0.009	(0.039)	−0.016	(0.042)
Cognitive ability							−0.029	(0.029)
Age (in years)							0.002	(0.028)
Female (=1)							0.123	(0.090)
Pocket money per week							−0.002	(0.003)
Math grade							−0.021	(0.052)
German grade							0.048	(0.048)
Number of siblings							−0.023	(0.028)
Migration background (=1)							0.139**	(0.069)
Education mother: A-Levels (=1)							−0.091	(0.063)
Education father: A-Levels (=1)							−0.143*	(0.074)
BMI							−0.003	(0.017)
Parents Occupation								
Full-time and part-time							−0.016	(0.082)
One full-time							0.049	(0.113)
Don't work/other regularity							0.070	(0.147)
Religion								
Protestant							0.075	(0.056)
Other or no religion							0.109	(0.095)
R^2	0.00		0.02		0.02		0.05	
Observations	658		658		658		658	

Notes: Positive coefficients imply more cautios planning behavior. Risk and time measures are expressed in standard deviations. Non-orthogonalized measure: Prudence. Cognitive-ability scores, relative German grade and relative math grade are standardized, such that above average scores are positive. Reference categories for parents' occupation is 'Both full-time', and 'Catholic' for religion. Missing demographic information has been imputed and controlled for. Outcome indices are formed using PCA weights and are standard normalized. Questions included in this index are listed in Section E.6 in the questionnaire in Online Appendix E. Bootstrapped standard errors (1000 repetitions) clustered at the session level in parentheses. ****/***/**/* denotes significance at the 0.1 / 1 / 5 / 10 percent level.

Table D-19: Preference for Competitive Income

	Dependent Variable: Preference for Competitive Income							
	(1)		(2)		(3)		(4)	
Risk aversion (r)	−0.094****	(0.023)	−0.093****	(0.023)	−0.096****	(0.023)	−0.070***	(0.023)
Prudence (p)			0.052**	(0.026)	0.052**	(0.026)	0.054**	(0.024)
Temperance (t)			0.053	(0.048)	0.053	(0.048)	0.032	(0.047)
Impatience					−0.017	(0.028)	−0.022	(0.025)
Cognitive ability							0.015	(0.019)
Age (in years)							0.021**	(0.009)
Female (=1)							−0.265****	(0.045)
Pocket money per week							0.002	(0.001)
Math grade							0.006	(0.019)
German grade							0.025	(0.029)
Number of siblings							0.001	(0.018)
Migration background (=1)							0.068	(0.048)
Education mother: A-Levels (=1)							−0.040	(0.041)
Education father: A-Levels (=1)							0.022	(0.078)
BMI							0.002	(0.006)
Parents Occupation								
Full-time and part-time							−0.028	(0.045)
One full-time							0.056	(0.068)
Don't work/other regularity							−0.010	(0.079)
Religion								
Protestant							0.025	(0.049)
Other or no religion							0.068	(0.042)
R^2	0.03		0.03		0.04		0.12	
Observations	649		649		649		649	

Notes: Positive coefficients imply increasing preference for competitive income. Risk and time measures are expressed in standard deviations. Non-orthogonalized measure: Risk aversion. Cognitive-ability scores, relative German grade and relative math grade are standardized, such that above average scores are positive. Reference categories for parents' occupation is 'Both full-time', and 'Catholic' for religion. Missing demographic information has been imputed and controlled for. Outcome indices are formed by adding z-Scores and are standard normalized. Questions included in this index are listed in Section E.7 in the questionnaire in Online Appendix E. Bootstrapped standard errors (1000 repetitions) clustered at the session level in parentheses.
****/***/**/* denotes significance at the 0.1 / 1 / 5 / 10 percent level.

Table D-20: Regression Results Using Correlated Simple, Raw Risk Measures (c.f. Table 2, Col. [4]-[5])

(a) Health-Related Behavior

	Unhealthy Behavior		Addictive Behavior		Smartphone Addiction	
Risk Aversion (r^{CE})	0.045	(0.044)	0.044	(0.045)	0.051	(0.048)
Prudence (p^{CE})	−0.136***	(0.042)	−0.143****	(0.040)	−0.159****	(0.037)
Impatience	0.118***	(0.039)	0.123***	(0.040)	0.117***	(0.039)
Other Factors	13		13		13	
R^2	0.13		0.12		0.12	
Observations	561		561		561	

(b) Financial Decision-Making

	Saving (with Debt)		Risky Investment		Insurance Demand	
Risk Aversion (r^{CE})	0.043	(0.048)	−0.056	(0.044)	−0.014	(0.021)
Prudence (p^{CE})	0.062	(0.041)	0.014	(0.044)	−0.020	(0.027)
Impatience	−0.197***	(0.033)	0.010	(0.027)	−0.004	(0.025)
Other Factors	18		18		13	
R^2	0.19		0.18		0.06	
Observations	658		658		643	

(c) Prevention and Environmentally-Friendly Behavior

	Prevention (Short Term)		Prevention (Long Term)		Eco-friendly Behavior	
Risk Aversion (r^{CE})	0.084*	(0.043)	0.197****	(0.028)	0.119***	(0.042)
Prudence (p^{CE})	−0.114***	(0.037)	−0.130****	(0.033)	−0.045	(0.028)
Impatience	0.058**	(0.024)	−0.086***	(0.033)	−0.086***	(0.029)
Other Factors	13		13		13	
R^2	0.09		0.15		0.12	
Observations	658		658		658	

(d) Planning Behavior and Preference for Competitive Income

	Cautious Planning		Preference for Competitive Income	
Risk Aversion (r^{CE})	0.007	(0.042)	−0.073***	(0.024)
Prudence (p^{CE})	0.096**	(0.037)	−0.003	(0.024)
Impatience	−0.011	(0.043)	−0.023	(0.025)
Other Factors	13		13	
R^2	0.05		0.11	
Observations	658		649	

Notes: Positive coefficients imply increasing preference for the respective behavior. Risk and time measures are expressed in standard deviations. Results obtained from OLS regressions with *correlated* risk measures. Consequently, coefficient estimates might be biased and standard errors incorrect. See Table 2, Columns [4]-[5], as well as Tables D-9 to D-19 in Appendix D for results from orthogonalized risk measures and for additional notes on the respective models, respectively. Bootstrapped standard errors (1000 repetitions) clustered at the session level in parentheses.
****/***/**/* denotes significance at the 0.1 / 1 / 5 / 10 percent level.

Table D-21: Regression Results Using Correlated Sophisticated Risk Measures (c.f. Table 2, Col. [6]-[8])

(a) Health-Related Behavior

	Unhealthy Behavior		Addictive Behavior		Smartphone Addiction	
Risk aversion (r)	0.024	(0.086)	0.007	(0.087)	−0.023	(0.086)
Prudence (p)	−0.133****	(0.034)	−0.145****	(0.034)	−0.171****	(0.032)
Temperance (t)	−0.032	(0.103)	−0.008	(0.102)	0.036	(0.095)
Impatience	0.116***	(0.037)	0.122***	(0.038)	0.115***	(0.037)
Other Factors	13		13		13	
R^2	0.14		0.13		0.12	
Observations	561		561		561	

(b) Financial Decision-Making

	Saving (with Debt)		Risky Investment		Insurance Demand	
Risk aversion (r)	0.096	(0.105)	−0.066	(0.070)	−0.018	(0.044)
Prudence (p)	0.002	(0.040)	−0.046	(0.043)	−0.023	(0.041)
Temperance (t)	0.003	(0.105)	0.055	(0.071)	0.015	(0.068)
Impatience	−0.197***	(0.032)	0.009	(0.026)	−0.003	(0.025)
Other Factors	18		18		13	
R^2	0.19		0.18		0.06	
Observations	658		658		643	

(c) Prevention and Environmentally-Friendly Behavior

	Prevention (Short Term)		Prevention (Long Term)		Eco-friendly Behavior	
Risk aversion (r)	0.120*	(0.063)	0.167***	(0.060)	0.201**	(0.081)
Prudence (p)	−0.130****	(0.034)	−0.040	(0.049)	0.001	(0.047)
Temperance (t)	−0.088*	(0.053)	−0.067	(0.054)	−0.141	(0.089)
Impatience	0.055**	(0.025)	−0.088***	(0.032)	−0.086***	(0.029)
Other Factors	13		13		13	
R^2	0.10		0.13		0.12	
Observations	658		658		658	

(d) Planning Behavior and Preference for Competitive Income

	Cautious Planning		Preference for Competitive Income	
Risk aversion (r)	−0.063	(0.050)	−0.124***	(0.042)
Prudence (p)	−0.084	(0.063)	0.046*	(0.027)
Temperance (t)	0.225****	(0.066)	0.032	(0.047)
Impatience	−0.016	(0.042)	−0.022	(0.025)
Other Factors	13		13	
R^2	0.05		0.12	
Observations	658		649	

Notes: Positive coefficients imply increasing preference for the respective behavior. Risk and time measures are expressed in standard deviations. Results obtained from OLS regressions with *correlated* risk measures. Consequently, coefficient estimates might be biased and standard errors incorrect. See Table 2, Columns [6]-[8], or Tables D-9 to D-19 in Appendix D for results from orthogonalized risk measures and for additional notes on the respective models. Bootstrapped standard errors (1000 repetitions) clustered at the session level in parentheses.
****/***/**/* denotes significance at the 0.1 / 1 / 5 / 10 percent level.

Table D-22: Relevance of Controlling for Higher-Order Risk Preferences in Predicting Field Behavior Using Common Approaches

	Simple Risk Measure		Survey Measures	
Coefficient of risk preferences (columns show estimated coefficients)	[1] Simple Risk (r^{CE})	[2] Simple Risk (r^{CE}); controlling for HORPs	[3] Survey Measure (Risk Tolerance)	[3] Survey Measure (Risk Tolerance); controlling for HORPs
Health-Related Behavior				
Unhealthy Behavior	−0.031	0.034	0.080*	0.057
	(−0.945)	(0.714)	(1.684)	(1.260)
Addictive Behavior	−0.036	0.025	0.084*	0.061
	(−1.038)	(0.526)	(1.739)	(1.327)
Smartphone Addiction	−0.038	0.020	0.080	0.055
	(−1.007)	(0.397)	(1.524)	(1.117)
Financial Decision Making				
Saving (with Debt)	0.067	0.034	−0.057	−0.048
	(1.576)	(0.592)	(−1.440)	(−1.196)
Risky Investment	−0.056	−0.053	0.135***	0.130***
	(−1.529)	(−1.070)	(3.121)	(2.889)
Insurance Demand	−0.025	−0.025	0.044*	0.041
	(−1.281)	(−0.873)	(1.699)	(1.586)
Prevention and Eco-Friendly Behavior				
Prevention (Short Term)	0.020	0.090***	0.173****	0.156****
	(0.495)	(2.618)	(6.218)	(5.427)
Prevention (Long Term)	0.124****	0.142****	−0.032	−0.028
	(4.145)	(3.989)	(−0.668)	(−0.596)
Eco-friendly Behavior	0.094**	0.113***	−0.013	−0.007
	(2.460)	(2.714)	(−0.383)	(−0.226)
Planning Behavior and Preference for Competitive Income				
Cautious Planning	0.062*	−0.014	−0.082**	−0.071*
	(1.745)	(−0.443)	(−2.142)	(−1.933)
Preference for Competitive Income	−0.075****	−0.071***	0.039	0.041
	(−3.447)	(−2.662)	(1.468)	(1.583)

Notes: Positive coefficients imply increasing preference for the respective type of behavior. All measures in standard deviations. Column [1] shows the regression coefficients when only using the simple, raw risk aversion measure (r^{CE}) as measure of risk preferences. Column [2] shows these coefficients from regressions where we account for the higher-order risk preferences prudence (p) and temperance (t). Column [3] shows regression coefficients when risk preferences are measured with the one-item survey question on willingness to take risk in general. Column [4] shows these coefficients from regressions where we account for the higher-order risk preferences prudence (p) and temperance (t). In all the regressions, we control for the influence factors from Section III. Bootstrapped t-statistics (1000 rep.) clustered at the session level in parentheses. ****/***/**/* denotes significance at the 0.1 / 1 / 5 / 10 percent level.

Table D-23: Regression Results Using Instrumental Variables to Account for Errors in Risk Aversion (c.f. Table 2, Col. [6]-[8])

(a) Health-Related Behavior

	Unhealthy Behavior		Addictive Behavior		Smartphone Addiction	
Risk Aversion	0.041	(0.075)	0.026	(0.076)	0.011	(0.078)
Prudence (p)	−0.139****	(0.025)	−0.145****	(0.026)	−0.159****	(0.026)
Temperance (t)	−0.010	(0.047)	−0.002	(0.047)	0.017	(0.045)
Impatience	0.118***	(0.038)	0.123***	(0.039)	0.117***	(0.037)
Other Factors	13		13		13	
Observations	561		561		561	

(b) Financial Decision-Making

	Saving (with Debt)		Risky Investment		Insurance Demand	
Risk Aversion	0.072	(0.092)	−0.068	(0.069)	−0.030	(0.042)
Prudence (p)	0.058*	(0.034)	−0.045	(0.043)	−0.024	(0.024)
Temperance (t)	0.087**	(0.034)	−0.032**	(0.016)	−0.001	(0.046)
Impatience	−0.197****	(0.032)	0.008	(0.026)	−0.004	(0.025)
Other Factors	18		18		13	
Observations	658		658		643	

(c) Prevention and Eco-Friendly Behavior

	Prevention (Short Term)		Prevention (Long Term)		Eco-friendly Behavior	
Risk Aversion	0.127**	(0.053)	0.193****	(0.050)	0.178***	(0.068)
Prudence (p)	−0.118****	(0.033)	0.011	(0.042)	0.023	(0.027)
Temperance (t)	0.018	(0.038)	0.079**	(0.035)	0.035	(0.058)
Impatience	0.058**	(0.025)	−0.083**	(0.032)	−0.084***	(0.029)
Other Factors	13		13		13	
Observations	658		658		658	

(d) Planning Behavior and Preference for Competitive Income

	Cautious Planning		Preference for Competitive Income	
Risk Aversion	−0.035	(0.044)	−0.081***	(0.027)
Prudence (p)	0.028	(0.051)	0.049**	(0.024)
Temperance (t)	0.170****	(0.045)	0.007	(0.050)
Impatience	−0.015	(0.042)	−0.023	(0.024)
Other Factors	13		13	
Observations	658		649	

Notes: Positive coefficients imply increasing preference for the respective behavior. Risk and time measures are expressed in standard deviations. Results obtained from IV estimations to account for measurement error (see, e.g., Durbin, 1954, for an early overview), where we follow Gillen et al. (2019) as described in Section B.1. We use orthogonalized risk measures as used in the main text (see Table S-3 in the supplementary material for results when using this instrumental variables approach on correlated risk measures – which shows the same pattern as without using instrumental variables, see Table D-21). See Table 2, Columns [6]-[8], as well as Tables D-9 to D-19 in Appendix D for results without using this instrumental variables technique and for additional notes on the respective models. Bootstrapped standard errors (1000 repetitions) clustered at the individual level in parentheses.
****/***/**/* denotes significance at the 0.1 / 1 / 5 / 10 percent level.

Table D-24: Regression Results – When Excluding Observations Corresponding to Simple Patterns of Decision Making (c.f. Table 2, Col. [6]-[8])

(a) Health-Related Behavior

	Unhealthy Behavior		Addictive Behavior		Smartphone Addiction	
Risk aversion (r)	0.015	(0.079)	0.003	(0.080)	−0.023	(0.080)
Prudence (p)	−0.162****	(0.027)	−0.167****	(0.027)	−0.177****	(0.027)
Temperance (t)	−0.061	(0.052)	−0.053	(0.051)	−0.023	(0.046)
Impatience	0.136****	(0.034)	0.137****	(0.035)	0.127****	(0.034)
Other Factors	13		13		13	
R^2	0.14		0.13		0.12	
Observations	534		534		534	

(b) Financial Decision-Making

	Saving (with Debt)		Risky Investment		Insurance Demand	
Risk aversion (r)	0.083	(0.094)	−0.062	(0.060)	−0.024	(0.039)
Prudence (p)	0.026	(0.033)	−0.045	(0.046)	−0.032	(0.023)
Temperance (t)	0.085**	(0.037)	−0.023	(0.023)	−0.014	(0.051)
Impatience	−0.216***	(0.032)	0.013	(0.028)	−0.013	(0.018)
Other Factors	18		18		13	
R^2	0.19		0.18		0.07	
Observations	627		627		614	

(c) Prevention and Environmentally-Friendly Behavior

	Prevention (Short Term)		Prevention (Long Term)		Eco-friendly Behavior	
Risk aversion (r)	0.097*	(0.057)	0.146***	(0.054)	0.168**	(0.068)
Prudence (p)	−0.118***	(0.037)	0.038	(0.047)	0.025	(0.033)
Temperance (t)	−0.014	(0.045)	0.048	(0.034)	0.020	(0.051)
Impatience	0.059**	(0.027)	−0.099****	(0.027)	−0.089***	(0.034)
Other Factors	13		13		13	
R^2	0.10		0.14		0.12	
Observations	627		627		627	

(d) Planning Behavior and Preference for Competitive Income

	Cautious Planning		Preference for Competitive Income	
Risk aversion (r)	−0.048	(0.043)	−0.071***	(0.026)
Prudence (p)	0.029	(0.063)	0.054***	(0.019)
Temperance (t)	0.157****	(0.045)	0.021	(0.036)
Impatience	−0.021	(0.041)	−0.022	(0.025)
Other Factors	13		13	
R^2	0.06		0.12	
Observations	627		619	

Notes: Positive coefficients imply increasing preference for the respective behavior. Risk and time measures are expressed in standard deviations. Results obtained from OLS regressions with orthogonalized risk measures as done in the main text, but excluding observations corresponding to simple patterns of decision-making as discussed in Section B.1. See Table 2, Columns [6]-[8], as well as Tables D-9 to D-19 in Appendix D for results with the full sample and for additional notes on the respective models. Bootstrapped standard errors (1000 repetitions) clustered at the session level in parentheses. ****/***/**/* denotes significance at the 0.1 / 1 / 5 / 10 percent level.

Table D-25: Regression Results – After Having Accounted for Probability Weighting in Estimating Utility-Based Measures (c.f. Table 2, Col. [6]-[8])

(a) Health-Related Behavior

	Unhealthy Behavior		Addictive Behavior		Smartphone Addiction	
Risk aversion (r)	0.002	(0.066)	−0.013	(0.067)	−0.037	(0.066)
Prudence (p)	−0.138****	(0.028)	−0.142****	(0.028)	−0.153****	(0.027)
Temperance (t)	−0.029	(0.050)	−0.021	(0.050)	−0.003	(0.048)
Impatience	0.116***	(0.038)	0.122***	(0.039)	0.116***	(0.037)
Other Factors	13		13		13	
R^2	0.14		0.13		0.12	
Observations	561		561		561	

(b) Financial Decision-Making

	Saving (with Debt)		Risky Investment		Insurance Demand	
Risk aversion (r)	0.099	(0.082)	−0.043	(0.061)	−0.014	(0.039)
Prudence (p)	0.050	(0.035)	−0.044	(0.040)	−0.019	(0.022)
Temperance (t)	0.090**	(0.035)	−0.035**	(0.015)	−0.001	(0.044)
Impatience	−0.197***	(0.032)	0.009	(0.026)	−0.002	(0.025)
Other Factors	18		18		13	
R^2	0.19		0.18		0.06	
Observations	658		658		643	

(c) Prevention and Environmentally-Friendly Behavior

	Prevention (Short Term)		Prevention (Long Term)		Eco-friendly Behavior	
Risk aversion (r)	0.078	(0.058)	0.145***	(0.049)	0.183***	(0.064)
Prudence (p)	−0.118****	(0.034)	0.012	(0.046)	0.014	(0.028)
Temperance (t)	−0.001	(0.035)	0.063**	(0.027)	0.026	(0.055)
Impatience	0.056**	(0.025)	−0.087***	(0.032)	−0.084***	(0.029)
Other Factors	13		13		13	
R^2	0.10		0.13		0.12	
Observations	658		658		658	

(d) Planning Behavior and Preference for Competitive Income

	Cautious Planning		Preference for Competitive Income	
Risk aversion (r)	0.005	(0.049)	−0.066***	(0.023)
Prudence (p)	0.004	(0.052)	0.052**	(0.021)
Temperance (t)	0.137****	(0.038)	0.025	(0.033)
Impatience	−0.018	(0.043)	−0.021	(0.025)
Other Factors	13		13	
R^2	0.05		0.12	
Observations	658		649	

Notes: Positive coefficients imply increasing preference for the respective behavior. Risk and time measures are expressed in standard deviations. Results obtained from OLS regressions with orthogonalized risk measures as done in the main text, but after having accounted for probability weighting in the elicitation procedure of utility points as summarized in Section B.1 and discussed in detail in Section B.3. See Table 2, Columns [6]-[8], as well as Tables D-9 to D-19 in Appendix D for results without accounting for probability weighting and for additional notes on the respective models. Bootstrapped standard errors (1000 repetitions) clustered at the session level in parentheses.
****/***/**/* denotes significance at the 0.1 / 1 / 5 / 10 percent level.

Table D-26: Regression Results – After Having Accounted for Probability Weighting in Estimating Utility-Based Measures and Regressions (c.f. Table 2, Col. [6]-[8])

(a) Health-Related Behavior

	Unhealthy Behavior		Addictive Behavior		Smartphone Addiction	
Risk aversion (r)	−0.174	(0.155)	−0.183	(0.160)	−0.192	(0.166)
Prudence (p)	−0.186****	(0.038)	−0.189****	(0.041)	−0.196****	(0.042)
Temperance (t)	−0.109	(0.069)	−0.098	(0.072)	−0.073	(0.072)
Probability Weighting	−5.233	(3.401)	−5.050	(3.543)	−4.596	(3.678)
Impatience	0.114***	(0.039)	0.120***	(0.039)	0.113***	(0.038)
Other Factors	13		13		13	
R^2	0.14		0.13		0.12	
Observations	561		561		561	

(b) Financial Decision-Making

	Saving (with Debt)		Risky Investment		Insurance Demand	
Risk aversion (r)	0.149	(0.107)	−0.005	(0.092)	0.069	(0.059)
Prudence (p)	0.064*	(0.037)	−0.043	(0.039)	0.004	(0.024)
Temperance (t)	0.111**	(0.050)	−0.018	(0.048)	0.035	(0.041)
Probability Weighting	1.479	(1.907)	1.118	(2.488)	2.484	(1.577)
Impatience	−0.196***	(0.032)	0.010	(0.026)	−0.001	(0.025)
Other Factors	18		18		13	
R^2	0.19		0.18		0.07	
Observations	658		658		643	

(c) Prevention and Environmentally-Friendly Behavior

	Prevention (Short Term)		Prevention (Long Term)		Eco-friendly Behavior	
Risk aversion (r)	0.059	(0.124)	0.285****	(0.080)	0.347****	(0.065)
Prudence (p)	−0.123***	(0.046)	0.050	(0.050)	0.058*	(0.033)
Temperance (t)	−0.010	(0.051)	0.124****	(0.036)	0.098	(0.063)
Probability Weighting	−0.593	(2.583)	4.181**	(2.107)	4.893***	(1.623)
Impatience	0.056**	(0.025)	−0.086***	(0.033)	−0.083***	(0.029)
Other Factors	13		13		13	
R^2	0.10		0.14		0.13	
Observations	658		658		658	

(d) Planning Behavior and Preference for Competitive Income

	Cautious Planning		Preference for Competitive Income	
Risk aversion (r)	−0.063	(0.096)	−0.002	(0.041)
Prudence (p)	−0.015	(0.059)	0.044*	(0.023)
Temperance (t)	0.107**	(0.054)	−0.006	(0.034)
Probability Weighting	−2.033	(2.363)	2.642*	(1.358)
Impatience	−0.018	(0.042)	−0.020	(0.024)
Other Factors	13		13	
R^2	0.05		0.12	
Observations	658		649	

Notes: Positive coefficients imply increasing preference for the respective behavior. Risk and time measures are expressed in standard deviations. Results obtained from OLS regressions with orthogonalized risk measures as done in the main text, but accounting for probability weighting in estimating utility-based measures and in regressions as summarized in Section B.1 and discussed in detail in Section B.3. See Table 2, Columns [6]-[8], as well as Tables D-9 to D-19 in Appendix D for results without accounting for probability weighting and for additional notes on the respective models. Bootstrapped standard errors (1000 repetitions) clustered at the session level in parentheses.
****/***/**/* denotes significance at the 0.1 / 1 / 5 / 10 percent level.

References

Clarke, D., J. P. Romano, and M. Wolf (2020). The Romano-Wolf multiple-hypothesis correction in Stata. *Stata Journal 20*(4), 812–843.

Crainich, D. and L. Eeckhoudt (2008). On the intensity of downside risk aversion. *Journal of Risk and Uncertainty 36*(3), 267–276.

Durbin, J. (1954). Errors in variables. *Revue de l'Institut International de Statistique / Review of the International Statistical Institute 22*(1/3), 23.

Gillen, B., E. Snowberg, and L. Yariv (2019). Experimenting with measurement error: Techniques with applications to the Caltech Cohort Study. *Journal of Political Economy 127*(4), 1826–1863.

Romano, J. P. and M. Wolf (2005a). Exact and approximate stepdown methods for multiple hypothesis testing. *Journal of the American Statistical Association 100*(469), 94–108.

Romano, J. P. and M. Wolf (2005b). Stepwise multiple testing as formalized data snooping. *Econometrica 73*(4), 1237–1282.

Romano, J. P. and M. Wolf (2016). Efficient computation of adjusted p-values for resampling-based stepdown multiple testing. *Statistics & Probability Letters 113*, 38–40.

E Online Appendix: Theory and Questionnaire

Here, we present our full questionnaire in verbatim with some theoretical considerations on the relation between the elicited behavior and higher-order risk preferences.

E.1 A Generic Theoretical Framework of Precautionary Behavior

Our generic framework based on Leland (1968) addresses cases of behavior where the decision to be taken involves a trade-off between the present and the (uncertain) future. The model is a two-period optimization problem of consumption. In period 1, the quantity of interest I_1 is known, e.g., quality adjusted life-years (QALYs), natural resources, income, etc. In period 2, however, it is assumed to be uncertain with subjectively known variance σ^2 and expectation I_2^*. What is not consumed in the first period, i.e., kI_1, where $k < 1$ is the control variable that measures the rate of preserving resources for the future, adds up to consumption in the second period, i.e., $C_2 = I_2 + (1+r)kI_1$, where $r > -1$ is the interest rate.

Mathematically, the problem is to

$$\begin{aligned} \max\nolimits_k \mathbb{E}[U(C_1, C_2)] \qquad & \text{s.th.} \\ C_1 = (1-k)I_1, & \\ C_2 = I_2 + (1+r)kI_1, & \\ E[I_2] = I_2^*, & \\ E[(I_2 - I_2^*)^2] = \sigma^2, & \end{aligned} \tag{6}$$

where U is the overall utility function and C_i denotes consumption in period i.

Departing from the point of optimality in case of certainty, the optimization problem can be solved using Taylor expansions (see *Analysis* below). The following condition determines whether k increases if the future becomes more uncertain:

$$U_{122} - (U_1/U_2)U_{222} < 0,$$

where U_1 denotes the derivative of U with respect to C_1 etc. Assuming additively time-separable utility U with $U_{222} = u_{222} = u_{111} = u'''$ results in $U_{122} = 0$, and we can immediately derive that a positive measure of prudence $p = u'''/u'$ implies an increase in k.[38] We note that the solution does not contain the second derivative of the utility function at all – that is, unlike prudence, risk aversion does not imply a demand for precautionary behavior in this setting.

[38] More realistic assumptions regarding U (which as a result also allow $u'' > 0$ corresponding to risk-seeking individuals) are discussed, e.g., in Schneider (2017). An even simpler (but maybe less realistic) framework restricted to risk lovers is presented in Crainich et al. (2013).

Analysis The solution to the maximization problem outlined above follows the canonical recipe for two-period consumption problems with future uncertainty: It is first assumed that an optimal solution $\tilde{k}$ exists in absence of uncertainty. For this $\tilde{k}$, the (deterministic) first-order condition

$$U_1 = (1 + r)U_2$$

needs to be fulfilled. For an uncertain future, we wonder whether for this $\tilde{k}$,

$$\mathbb{E}[U_1] - (1 + r)\mathbb{E}[U_2] \tag{7}$$

still equals 0, implying that the same level of k is optimal also in the uncertain case. To address this, we re-write this expression using (local linear) functions of the degree of uncertainty $(C_2 - \tilde{C}_2)$, where $\tilde{C}_2$ is the consumption level in period 2 that corresponds to $\tilde{k}$. In other words, we use a Taylor expansion around $\tilde{C}_2$. Ignoring the higher-order terms of the Tayler expansion, (7) can be rewritten as follows (with all derivatives of U evaluated at $\tilde{k}$):

$$\begin{aligned} U_1+&U_{12}\mathbb{E}(C_2 - \tilde{C}_2) + 1/2U_{122}\mathbb{E}(C_2 - \tilde{C}_2)^2 \\ &- (1 + r)\left(U_2 + U_{22}\mathbb{E}(C_2 - \tilde{C}_2) + 1/2U_{222}\mathbb{E}(C_2 - \tilde{C}_2)^2\right). \end{aligned}$$

For a fixed k, $(C_2 - \tilde{C}_2) = (I_2 - I_2^*)$, and we obtain, repeatedly using the first equation,

$$\begin{aligned} U_1 + 1/2U_{122}\sigma^2 - (1 + r)\left(U_2 + 1/2U_{222}\sigma^2\right) &= 1/2\left(U_{122}\sigma^2 - (1 + r)U_{222}\sigma^2\right) \\ &= 1/2\left(U_{122}\sigma^2 - (U_1/U_2)U_{222}\sigma^2\right). \end{aligned} \tag{8}$$

Using the second-order condition of the maximization problem (in the uncertain case), a negative (8) implies a higher k compared to $\tilde{k}$. See Leland (1968) for a more detailed derivation of the solution.

E.2 Empirical Strategy and Mapping of Questionnaire and Theory

Our generic framework and the models on risky investment (below) concern an individual's reaction to an increase in future risk. Generally, in these models, the future risk cannot be insured against (e.g., income risk, health risks, etc). However, it can (partly) be coped with by a given behavior (e.g., saving, healthy behavior, etc). Empirically, for behavior described by these models, we test directly whether our measures of higher-order risk preferences predict the relevant behavior instead of the changes in behavior that come with an increase in future risk.[39] When differences regarding the relevant future risk and other factors affecting

[39] Questions on cautious planning explicitly address the change in riskiness. For insurance and short-term prevention, we describe the empirical strategy below in Appendix Sections E.4 and E.5.

the corresponding behavior are either sufficiently controlled for or sufficiently similar, the variation in prudence or temperance should correlate with the variation in the respective behavior. In the case of precautionary saving and risky investment, in addition to the standard set of controls for demographic and socio-economic information, including parental education, we thus control for information capturing income risk (following Noussair et al., 2014). Specifically, these are: possible positive and negative shocks in pocket money ("Is your pocket money cut sometimes?", "Do you get additional pocket money for larger purchases and expenses?"), the average amount that participants earn on their side job (if any), whether or not those earnings are stable, and whether or not participants can control it themselves (e.g., by deciding themselves on the hours they want to work). For all other behavior, the same holds regarding factors affecting the behavior under study, while the relevant future risks are arguably sufficiently similar in our sample. For example, with respect to health, we assume that (subjective) health risks beyond our controls are sufficiently homogeneous within our adolescent sample for our empirical strategy to work.

E.3 Health-Related Behavior

Theory Theoretically, precautionary health behavior can be linked to higher-order risk preferences within our generic framework above (Section E.1): While the quality of life one may enjoy or consume while young (first period, C_1 in the model) is known, the quality of life one may enjoy in the future (C_2, expressed in quality adjusted life-years, QALYs) depends on both previous behavior with its effects on the stock of QALYs ($(1+r)kI_1$ in the model), and on other factors that are uncertain and beyond one's control (like genetic influences, I_2 in the model). Here, the model predicts a demand for precautionary health-related behavior for prudent individuals in the following sense: They are expected to avoid overconsumption of QALYs while young and hence engage more in healthy rather than unhealthy behavior.

In our generic framework, addictive behavior more generally (i.e., beyond smoking and drinking) might be interpreted as overconsumption of QALYs by bringing joy instantaneously, at the cost of risking unwanted side-effects in the longer run that negatively affect the quality of life, such as the inability to pursue goals and duties, or health consequences, such as cardiovascular diseases or depression (e.g., associated with social-media addiction, see Orben and Przybylski, 2019).

Survey Questions We capture health-related behavior with the questions listed in this subsection.

For all questions listed in this section, answer possibilities are given in brackets. For likert scale items, ranges of numbers indicate the scale from which participants could choose. The extreme options of the scales were labeled as e.g., "totally agree/totally disagree", "at every occasion/never", or "very often/never", depending on the item.

- Body height (in cm); body weight (in kilograms)
- How often do you exercise/play sports (e.g., soccer, volleyball, dancing, running, ...)? ["never", "less than 1x per month", "1x per month", "2x per month", "1x per week", "2x per week", "more than 2x per week"]

Sub-index of Health-Related Behavior: Questions Targeting Addictive Behavior

- [Grades 8, 10 and 12 only] Do you smoke cigarettes? ["I do not smoke", "I do not smoke, but I have tried it", "I smoke approx. 1-2 cigarette(s) per day", "I smoke approx. one pack of cigarettes per week", "I smoke more than one pack of cigarettes per week"]
- [Grades 8, 10 and 12 only] Do you drink any alcohol? ["no, never", "yes, rarely (up to 1-2x per month)", "yes, occasionally, one to two drinks (up to 1-2x per week)", "yes, occasionally, more than two drinks (up to 1-2x per week)", "yes, regularly (more often than 2x per week)"]

Sub-index of Addictive Behavior: Questions Targeting Excessive Smartphone Usage

- When I take a photo with my cell phone or experience a special situation, I immediately think about posting it on Facebook, Instagram, Snapchat or the like. [0-5]
- I get into trouble with my parents or friends or with my girlfriend/boyfriend, because I use my smartphone that much. [0-5]
- I feel uncomfortable (e.g., nervous or fretful or disquiet or a bit sad) when I cannot use my smartphone for a considerable time, because of an empty battery, no signal, or because my smartphone was taken away. [0-5]
- When I feel bad or when I face a difficult task, I distract myself with my smartphone. [0-5]
- My smartphone disturbs me while doing my homework or studying. [0-5]
- I often check my phone while eating with my family to see if there are any news. [y, n]

E.4 Financial Decision-Making

Saving w./ Debt

The theory on precautionary saving that we consider here is described in Section E.1 above. We capture saving behavior with the following questions:

- How do you handle your pocket money/income? ["I spend everything quickly", "I save less than the half", "I save approximately the half", "I save more than the half", "I save everything"]
- Assuming that you get 50 euros for Christmas or for your birthday. What will you do with the money? ["I spend everything quickly", "I save less than the half", "I save approximately the half", "I save more than the half", "I save everything"]
- Do you have a bank account? [y, n]
- Do you borrow money from your parents? ["Yes, actually every month", "Yes, several times per year

(more than 4 times per year; but not every month)", "Yes, rarely (less than 4 times per year)", "No, never"]

- Do you have a debit card? [y, n]

Risky Investment

Gollier and Pratt (1996, eq. 10) analyze the effect of a small mean-zero background risk $\tilde{\varepsilon}$ on the risk premium of another independent risk. Building on Pratt (1964), this effect can be approximated for small risks by the relative increase in the absolute risk aversion measure $\Delta r/r$ that is due to the addition of $\tilde{\varepsilon}$. Assuming u thrice continuously differentiable with $u' > 0$ and $u'' > 0$, Gollier and Pratt (1996) show that

$$\frac{\Delta r}{r} \cong 0.5\sigma_{\varepsilon}^2(t - pr),$$

where t, p and r are the measures of temperance, prudence, and risk aversion used here.[40] The larger t, the larger grows the risk premium due to taking background risk into account. Following the examples in Gollier and Pratt (1996), this is typically used to characterize the "tempering" effect of a background risk – such as income risk – on the risk aversion towards another independent risk, that is, the degree of reluctance to hold a risky portfolio. Hence, for the same increase in background risk, a larger t implies less risky investment than a smaller t would do.

We capture risky investment with the following questions:

- Do you know what a stock is? [y, n]
- Do you have any stocks? [y, n]
- Do you think you will buy some stocks in the future? [y, n]
- Have you ever used money that was originally intended for something else at a subsequent date (e.g., for holidays or a present), for a bet or invested it in stocks? [y, n]

Financial Insurance

Fei and Schlesinger (2008) discuss different settings of insurance demand with two states of the world: Either, there is a loss event with a certain probability p (loss states), or no loss occurs (no-loss states). The settings then differ by whether or not background risk is considered, and if so, in which state(s) of the world background risk is added. Here, we consider the case where background risk is added to the no-loss states: The exact value of an insured item is uncertain (e.g., because it is a used item and/or an item of daily use), and the contract is written based upon the appraised value. We do so as this is arguably the only realistic case where adolescents might consider insurance coverage – for their phone or their bike, both of which they use regularly, and which are thus of uncertain value after

[40] Note that for temperance and prudence, the measures in Gollier and Pratt (1996) are defined slightly differently, for which reason their Equation 10 looks differently.

their first usage. Insurance payments in such cases usually depend on an appraised value based solely on the age of the object. In this case, following Fei and Schlesinger (2008, eq. 3 with $\beta = 0$), a consumer (with utility u, four times differentiable, satisfying $u' > 0$ and $u'' < 0$) chooses their insurance demand maximizing their expected utility according to

$$\max_\alpha EU = pu(y_L) + (1-p)(y_N),$$

where y_L and y_N are the wealth levels in the loss and no-loss states, respectively. They are given by $y_L = W - L + \alpha L - P(\alpha)$ and $y_N = W + \tilde{\varepsilon} - P(\alpha)$, where W is the initial wealth level, L is the size of the loss, α is the co-insurance rate, and thus αL is the payment in case of a loss. The background risk $\tilde{\varepsilon}$ is a not directly hedgeable zero-mean random variable with positive variance that is uncorrelated with the loss. $P(\alpha) = (1 + m)\alpha pL$ is the premium depending on the loading factor $m \geq 0$. Fei and Schlesinger (2008) show that in comparison to the case without background risk, for a fair insurance, a prudent consumer purchases less than full coverage: They save expenditure on the premium, thereby shifting wealth to the no-loss states to mitigate the pain from the background risk. An increase in insurance demand compared to the case without background risk happens if and only if a consumer is imprudent.

We measure the insurance demand in this settings with the following questions:

- Do you have a cell phone insurance? [yes, no, I do not know]
- Did you take it out yourself? [yes, no, I do not have a cell phone insurance/I do not know if I have one]
- Do you have a bike insurance? [yes, no, I do not know]
- Did you take it out yourself? [yes, no, I do not have a bike insurance/I do not know if I have one]

From this set of questions, we can, for the setting considered, either infer $\alpha > 0$ or $\alpha = 0$. Aggregation over this set of question and over all individuals then yields an aggregate insurance demand α for items that fall into the category considered here. Under the assumption that insurance contracts are evaluated as fair,[41] prudent consumers will have, on average, a lower α than prudent-neutral consumers (who, assuming risk aversion, opt for full insurance coverage just as in the case without background risk). Thus, as we observe risk aversion in the aggregate, a negative correlation between our index of financial insurance demand (aggregate α) and our measures of prudence would support this theory.

[41] For the proof of the general case for arbitrary loading factor m, the combination of less prudence, but more risk aversion, is used to reach this conclusion, see Proposition 4 in Fei and Schlesinger (2008).

E.5 General Prevention and Eco-Friendly Behavior

General Prevention (Short Term)

Eeckhoudt and Gollier (2005) model prevention effort where an agent is facing the risk of a loss $L > 0$. The probability of its occurrence $p(e)$ is a decreasing function of the effort e expressed in monetary terms that is invested to prevent the loss. Utility u is assumed to be increasing and differentiable. The agent wants to maximize their expected utility given by

$$U(e) = p(e)u(w - L - e) + (1 - p(e))u(w - e),$$

where U is assumed to be concave in e and w denotes wealth. To analyze the optimal level of prevention of an agent with utility function u, Eeckhoudt and Gollier (2005) depart from the optimal prevention level e_n of a risk-neutral agent, which is such that an additional dollar in prevention reduces the expected loss by one dollar. They then investigate whether the agent with utility function u optimally invests more or less effort in prevention. For the case $p(e_n) = 1/2$, that is, when the loss happens with probability 1/2, their Proposition 2 states that a prudent agent, either risk-averse or risk-lover, selects a level of effort smaller than e_n (and imprudent agents analogously select a larger level of effort).

To capture prevention effort, we use the following survey questions:

- I mutually interchange secrets with my friends to make sure they do not disclose mine. [0-5]
- To make sure that I can always use my mobile phone and that I can be reached, I have a powerbank with me. [0-5]
- Because the others do the same, I prefer to go to the bakery or to the kiosk instead of taking food from home. [0-5]
- Because I think of packing something to eat and drink during longer journeys by bus, train or car I am not hungry or thirsty in such situations. [0-5]
- When the class is divided up into groups, I make sure that I have at least one student in my group who is good at the subject in question. [0-5]
- [Grades 8, 10 and 12 only] Because (romantic) relationships sometimes go better and sometimes worse, I invest time in relationships with good friends and my family – they are always there for me. [0-5]

General Prevention (Long Term) and Eco-Friendly Behavior: Theory Long-term prevention and eco-friendly behavior might be interpreted within our illustrative framework above (Appendix E.1): The quantities of interest could be the quality of life (long-term prevention) or natural resources (eco-friendly behavior) that one may enjoy or consume, which is known in the present (C_1 in the model), but uncertain in the future (C_2). Uncertainty (σ^2 in the model) could arise from unwanted events that might occur and affect the quality of life (long-term prevention) or from advances in technology or newly discovered tipping points that affect the level of available resources in addition to what has not been consumed (eco-friendly behavior).

Prudent individuals are then predicted to engage in precautionary behavior, for example by shifting (the consumption of) quality of life from the present to the future by engaging in (long-term) prevention effort, or by reducing consumption of natural resources and avoiding overconsumption.

See Menegatti (2009) for a more sophisticated model of (two-period) prevention effort, and Bramoullé and Treich (2009) for a more sophisticated model of eco-friendly behavior.

General Prevention (Long Term): Survey Questions We capture (two-period) prevention effort with the following questions:

- When packing, do you use a packing list to make sure you do not forget anything important? [y, n]
- I brush my teeth as often and as long as I should. [0-5]
- I pay attention to my diet: that it is healthy and balanced, not too much and not too little. [0-5]
- For some subjects, I study more in order to compensate for a worse grade in another subject, for example because I do not like the other subject, or because the tests/exams are often very difficult. [0-5]
- Because the risk of being caught copying, for example from a cheat sheet, is much too high for me, I prefer to learn more and refrain from copying. [0-5]
- On average: How long do you prepare for a test or an exam? ["more than one week", "approximately one week", "a few days", "one day"]
- Because I do not know yet what I would like to become later, I try to get good grades to keep all possibilities open to me. [y, n]
- [Grades 10 and 12 only] If I have to give a presentation at school using PowerPoint, I will always have two options to access the file (e.g. via my e-mail address and a USB key) or I have the presentation as a PDF file with me. [0-5]
- [Grades 8, 10 and 12 only] When looking for a (side) job, an internship or even a university place, it makes sense to send further applications until you have received a written confirmation of the desired option, even if it has already been confirmed orally. [y, n]
- [Grades 8, 10 and 12 only] Every now and then, I check whether the vaccinations according to my vaccination card are up-to-date. [y, n]

Eco-Friendly Behavior: Survey Questions We capture eco-friendly behavior with the following questions:

- I buy second-hand products, for example second-hand clothes, mobile phones, laptops, or the like. [0-5]
- If I leave my room for several hours, I will turn down the heating. [0-5]
- If I am the last to leave the room, I will turn off the light. [0-5]
- If I do not need the water while showering, I will turn it off. [0-5]
- If currently no one is watching, the TV will be turned off. [0-5]
- If I do not use the computer/laptop for a considerable time, I will turn it off resp. put it into the power-saving mode. [0-5]
- When I do the shopping, I use my own bag or backpack. [0-5]
- At school or on the way, I use my own beverage bottle (made of glass or metal). [0-5]

- I use my own cup for coffee or hot chocolate. [0-5]
- I try using the bike, wherever it is possible. [0-5]
- I separate my waste to the best of my knowledge and belief. [0-5]
- If you go to the bathroom, wash your hands and there are only paper towels to dry your hands: How many paper towels do you take? [0-10]
- When you are in the canteen, how many napkins do you take on your tray? [0-10]
- [Grades 8, 10 and 12 only] If you smoke (otherwise leave the question unanswered): I throw the cigarettes on the ground after smoking. [0-5]

E.6 Planning Behavior

Planning behavior may be interpreted within our illustrative model above (Appendix E.1) analogously to the case of (long-term) prevention (see Appendix E.5), where quality of life is shifted from one period to another.

We capture precautionary planning behavior with the following questions:

- Imagine in the next vocabulary test 10 words from the last lesson of the last school year are asked in addition to the current lesson. How much longer are you going to study? ["0 minutes", "10 minutes", "20 minutes", "30 minutes", "45 minutes", "1 hour", "1 hour, 30 minutes", "2 hours", "2 hours, 30 minutes", "3 hours", "4 hours", "5 hours", "6 hours", "7 hours"]
- Imagine you would like to visit us at the Max Planck Institute and have an appointment with us. According to Google Maps you need 20 minutes by bike from the main station in Bonn, where you start either with your own bike or with a borrowed one. However, there are three traffic lights on the route, all of which can be either red or green – or any combination of the two. How many minutes/hours before the meeting should you start at the main station? ["1 hour", "55 minutes", "50 minutes", "45 minutes", "40 minutes", "35 minutes", "30 minutes", "25 minutes", "20 minutes", "15 minutes"]
- [Grades 10 and 12 only] Imagine you have to hand in an important document of several pages printed and bound at a certain time (say, 12 noon), e.g., a seminar paper or a longer presentation with classmates. You decide to have this done in a copy shop right next to the place where you have to hand in the document. Also, imagine you could go there from home and that would take 10 minutes. It is always possible that the USB stick is not readable, the format is wrong, the file is not readable, or there are five customers ahead of you in line. The printing itself and the binding do not last longer than 15 minutes. How many minutes/hours before handing in do you start going to the copy shop from home? ["20 minutes", "25 minutes", "30 minutes", "35 minutes", "40 minutes", "45 minutes", "50 minutes", "55 minutes", "1 hour", "1 hour, 15 minutes", "1 hour, 30 minutes", "1 hour, 45 minutes", "2 hours", "2 hours, 30 minutes", "3 hours"]

E.7 Preference for Competitive Income

Fuchs-Schündeln and Schündeln (2005) document lower precautionary wealth in the presence of income risk when selection into low-income risk occupations is possible. Assuming risk aversion equals prudence, they argue this observation is due to risk-averse/prudent individuals choosing low income risk occupations, who thus have no need to hold precautionary wealth.

To proxy for preferences reflecting acceptance of high income risk occupations, we use the following questions:

- Later, I would like to be self-employed, e.g., as a craftsman, an architect, a cafe owner, etc. [y, n]
- Later, I would like to be a civil servant, e.g., as a teacher, a policeman, in a city's administration or at the tax office, etc. [y, n]
- For the riddles, we will pay a few "Taler" for each correct solution. Although we will not change that: Would you prefer a fixed amount of "Taler" for your payment, regardless of the number of riddles that you have solved correctly? [y, n]
- Or alternatively, would you like to make a small competition out of it? We would assign you a classmate from the room, and the one of you who would have solved more riddles correctly, would get the partner's fixed payment and additionally his own fixed payment. However, the other one would get nothing. [y, n]

E.8 General Risk-Taking Behavior

One-Item Survey Questions

- How do you assess yourself: Are you generally a person who is ready to take risks or do you try to avoid risks? Please tick one of the boxes on the scale, whereby the value 0 means: "not at all ready to take risks" and the value 10 means: "very ready to take risks". With the values in between you can graduate your assessment. [0-10]

(Adapted) Domain-Specific Risk-Taking Scale (DOSPERT)

- [Grades 8, 10 and 12 only] How many times did you drink five or more alcoholic beverages on a single evening in 2018? [0 (never) - 5 (at every occasion)]
- [Grades 10 and 12 only] How often did you copy parts of somebody else's work in 2018 (e.g., copied a longer text from Wikipedia for a presentation or copied some homework)? [0-5]
- Have you ever skied on a slope that has exceeded your abilities or have you skied off-piste? [yes, no, I do not ski]
- [Grades 10 and 12 only] Have you ever gotten involved in unprotected sex? [y, n]
- How many times did you tell a friend's secret to someone else in 2018? [0-5]
- How many times did you not fasten your seat belt while driving in 2018? [0-5]
- How often did you not wear a helmet when riding a scooter or a motorbike (or similar) in 2018? [0-5]
- How often did you not use sun protection even though you were in the sun for a long time in 2018? [0-5]
- How often did you copy (from your neighbor, a cheat sheet, ...) in a class test/exam in 2018? [0-5]
- How often did you fake the signature of another person (e.g., your parents) in 2018? [0-5]
- Have you ever stolen a small item in a shop (e.g., a pencil or a lipstick)? [y, n]
- How often did you wear clothes (including private occasions) that your parents or someone else disapproved of in 2018? [0-5]
- How many times did you steal a small amount of money from someone you know in 2018? [0-5]
- How many times were you involved in a brawl in 2018? [0-5]
- How many times did you cross a red light in 2018? [0-5]
- Instead of using illegal streaming sites, I prefer using Netflix, Amazon Prime Video or similar services that I pay for. [y, n]
- Have you ever gambled away an entire week's pocket

money (or more) in a bet? [y, n]

- How often did you not wear a helmet when you rode a bike in 2018? [0-5]
- Have you ever met a person you got to know online/on social networks/apps? [y, n]
- If I have forgotten my homework, I will not announce it and simply hope that it will not be my turn during the discussion. [yes, a bit of both – it depends, no]
- Do you use your mobile phone in traffic other than for navigation (e.g., when you are driving a car, scooter or bicycle, when you are crossing the road, ...)? [y, n]

E.9 Demographic Information and Income Risk

- I am [female, male]
- Your postcode/I am from [Choice list with possible living areas]
- What grade are you in? [6,8,10,12]
- Your month of birth [1 - 12] and your year of birth [Choice list with birth years]
- Last year, I got the following grades in my report: In mathematics [1, 2, 3, 4, 5, 6]; in German [1, 2, 3, 4, 5, 6]
- I am [Choice list with the most frequent religions]
- Please mark the appropriate statement: ["My parents and I were born in Germany", "I was born in Germany. One parent was not", "I was born in Germany. My mother and my father were not", "I was not born in Germany"]
- My mother has A levels [y, n]; my father has A levels [y, n]
- My parents ["both work full-time (e.g. both father and mother work from Monday to Friday the whole day)", "one works full-time, one works part-time", "both work part-time (e.g. both father and mother only work in the midmorning or only on 2-3 days per week)", "one works full-time", "one works part-time", "work in another regularity", "currently, both do not work."]
- Number of younger sisters [0, 1, 2, 3, 4], older sister [0, 1, 2, 3, 4], younger brothers [0, 1, 2, 3, 4] and older brothers [0, 1, 2, 3, 4]
- Approximate amount of pocket money (from my parents, my grandparents, ... altogether) per week [0-50; steps of 0.5]
- I have a side job, through which I earn the following amount per week (on average; 0 if no side job) [0-150; steps of 1]
- Is your pocket money cut sometimes? [y, n]
- Do you get additional pocket money for larger purchases and expenses? [yes, sometimes/it depends, no]
- [Grades 10 and 12 only] Do you regularly get the same amount of money in your side job? [yes, no, I do not have a side job]
- [Grades 10 and 12 only] Do you have any influence on it (e.g., because you can decide yourself how often you work)? [yes, no, I do not have a side job]

References

Bramoullé, Y. and N. Treich (2009). Can uncertainty alleviate the commons problem? *Journal of the European Economic Association 7*(5), 1042–1067.

Crainich, D., L. Eeckhoudt, and A. Trannoy (2013). Even (mixed) risk lovers are prudent. *American Economic Review 103*(4), 1529–1535.

Eeckhoudt, L. and C. Gollier (2005). The impact of prudence on optimal prevention. *Economic Theory 26*(4), 989–994.

Fei, W. and H. Schlesinger (2008). Precautionary Insurance Demand With State-Dependent Background Risk. *The Journal of Risk and Insurance 75*, 1–16.

Fuchs-Schündeln, N. and M. Schündeln (2005). Precautionary savings and self-selection: Evidence from the German reunification experiment. *Quarterly Journal of Economics 120*(3), 1085–1120.

Gollier, C. and J. W. Pratt (1996). Risk vulnerability and the tempering effect of background risk. *Econometrica 64*, 1109–1123.

Leland, H. E. (1968). Saving and uncertainty: The precautionary demand for saving. *Quarterly Journal of Economics 82*(3), 465–473.

Menegatti, M. (2009). Optimal prevention and prudence in a two-period model. *Mathematical Social Sciences 58*(3), 393–397.

Noussair, C. N., S. T. Trautmann, and G. van de Kuilen (2014). Higher order risk attitudes, demographics, and financial decisions. *The Review of Economic Studies 81*(1), 325–355.

Orben, A. and A. K. Przybylski (2019). The association between adolescent well-being and digital technology use. *Nature Human Behaviour 3*(2), 173–182.

Pratt, J. W. (1964). Risk aversion in the small and in the large. *Econometrica 32*(1/2), 122–136.

Schneider, S. O. (2017). *Advances and applications of experimental measures to test behavioral saving theories and a method to increase efficiency in binary and multiple treatment assignment.* Ph. D. thesis, Göttingen, Georg-August Universität.

F Online Appendix: Instructions (for Online Publication)

F.1 Translated Instructions

Hello and welcome to our study. Glad, that you are here and willing to participate. Within the next 45 minutes, we are going to play some ‘decision games’ with you, you will work on some riddles and then, you are will be asked to complete a questionnaire. This you will do almost exclusively on a tablet computer and we will explain everything explicitly step by step. First we will explain, then you can take action, and then we will explain the next step. We start with the games.

(In the session with the older students): Just a quick comment on the explanations. Since we are doing a scientific study, it is important that we always give the same explanations. As we also conduct the study with younger students, the explanations are at times more detailled than it would be necessary for you. Thus, in case it seems a little elongated to you, rest assured, this has nothing to do with you, but we just have to do it this way and it also ensures that you really understand everything very well.

From now on, please do not talk to each other anymore, leave your cell phone where it is resp. put it away in case you are holding it in your hands and please listen carefully. You can earn money in the games. We will pay you out in cash at the end of the experiment or you will receive the money in an envelope – more on this later. The amount of money you can earn depends on your answers and decisions. That is why it is important for you to understand the rules. So please listen carefully! We will stop our instructions repeatedly, so that you can ask questions. Just raise your hand, then one of us will come to you to answer your question.

Everything OK so far? *(Leave some time for questions; answer questions individually and in private)*

In the first game, you are to decide four times whether you would prefer a specific amount of money today, or a slightly larger amount of money in 3 weeks time. Here you can see such a decision situation. *(Show the slide of the presentation that displays the time preference decision situation.)* This is how the decision screen will look like. On the left, you can see the amount of money you would get immediately, in this example 100 thalers. On the right, you can see the amount you would get in three weeks, in this example 120 thalers.

So if you say, given the ‘basic amount’ of 100 thalers, I would wait three more weeks in order to get 20 additional thalers – which option do you have to choose? *(Assuming that the answer is ‘right’)* Exactly, then you have to choose the right option. If you prefer to have 100 thalers today, you have to choose the left option, accordingly.

We convert thalers to euros and 100 thalers are approximately (*mention the relevant amount only*)

- grade 6: 2 euros.
- grade 8: 2.50 euros.
- grade 10: 3.50 euros.
- grade 12: 5.50 euros.

So think carefully about your preferred option.

You can simply enter your decesion by tapping the 'L' or 'R' button.

Everything ok so far? *(Leave some time for questions; answer questions individually and in private)*

Concerning the payout: In addition to the decision games, there are a few riddles. For each riddle you have solved correctly you will get some additional money.

Besides this game, we are going to play another two types of games with you. Overall, you will make about 25 decisions, and one of those decisions will be paid out for real.

Your tablet randomly chooses one of the three types of decision games and it also randomly picks the number of the decision. It is important that you take every decision seriously, because until the end, you will not know which decision will be payed out.

If this game is randomly chosen for payout by the tablet, you will receive the money either today or in three weeks – depending on your decisions.

If you decided for a payout in three weeks and this decision was randomly chosen for payout, you could collect the money in the secretary's office in three weeks *(adapt to procedure in the corresponding school)*.

Everything ok so far? *(Leave some time for questions; answer questions individually and in private)*

If anyone of you does not want to participate, please let us know now. You will also be able to stop later at any time. Just raise your hand – then one of us will come to you and discuss the next steps.

Does anyone like to stop now or do you have any questions? *(Leave some time for the students to raise their hands resp. for questions; answer questions individually and in private; if someone drops out, write down the tablet's ID-number and the session number in order to be able to delete the observation.)*

Everything ok so far? *(Leave some time for questions; answer questions individually and in private)*

(Black out slide show by pressing the 'B' key)

Okay, then we will play the decision games now.

(Start session)

(As soon as everyone has made their decisions) Now you will decide 18 times whether you would rather have a specific certain amount or you would like to throw a coin with us and end up having either a higher or a lower amount than the certain amount. We will change the amounts from decision situation to decision situation.

Such a decision situation looks like this, for example. *(Show the slide of the presentation that displays the coin tossing decision situation).* On the left, we have a coin and you will get 70 thalers, regardless of whether the coin lands with the white or the black side at the top. So you will get this amount with certainty; we show that by the fact that both for the white side *(point at the upper arrow)* and the black side *(point at the lower arrow)* there are 70 thalers in the end. On the right *(point at the right option)* this is different. Here you will get 140 thalers, if the coin lands with the white side at the top *(point at the upper arrow)*, thus laying on the black side. If the coin lands with the black side at the top *(point at the lower arrow)*, you will get 0 thalers — that is: nothing.

Thus, you have to decide, whether you would rather take home 70 thalers with certainty or whether you would like to have the chance to get 140 thalers, where you can also end up empty-handed. So if you say: "I would rather like to have the chance to get 140 thalers and take the risk of ending up empty-handed with this coin toss", which option do you have to choose? *Assuming that the answer is 'right')* Exactly, you have to choose the right option. Otherwise, if you say you would prefer to play it safe, you have to chose the left option.

To enter your decision, simply tap on the button below the option you prefer. Since the decision situations look rather similar at first sight, you also have to press 'Next' *(point at the 'next' button)*, to make sure you do not accidentally choose the same answer again for a different situation.

Everything ok so far? *(Leave some time for questions; answer questions individually and in private)*

Now turning to the payout: Let's assume, the computer selected decision number one of the coin toss.

Let us additionally assume that you had chosen the left option. Then you would simply get 70 thalers. However, if you had chosen the right option, your tablet would toss a coin. If the coin showed white, you would get 140 thalers in this example.

As I said, we will convert thalers to euros later. 140 thalers are the highest payout you can earn in this game. That is about (*mention the relevant amount only*)

- grade 6: 2,75 euros.
- grade 8: 3,50 euros.
- grade 10: 5,00 euros.

- grade 12: 7,75 euros.

So think carefully about how you decide.

Everything ok so far? *(Leave some time for questions; answer questions individually and in private)*

(Black out slide show by pressing the 'B' key)

(Start subsession 'Certainty Equivalents')

(As soon as everyone has made their decisions) In the last of the three games, in different situations, you can choose whether you prefer to draw a ball from a bag, we call it bag L for left, or a ball from another bag, we call it bag R for right. Of course, this will happen without you being able to look into the bag, so you will not be able to pick out the ball you want. You will draw a ball randomly. Here you can see how such a bag looks like *(Show the slide of the presentation that displays the urn decision situation)*. As you can see, there are four balls in each bag. The number written on a balls indicates how many thalers you will get if you randomly draw the corresponding ball. For example *(point at the ball with a 50 written on it, marked with R)*, on the red ball – R means red, G means green, B means blue — you can read 50. So, if you happen to draw this ball, you get 50 thalers. As I said, you will not be able to look into the bag, so you could draw each of the four balls, and the chances of drawing each of these balls are the same. That is, if you draw out of the right bag, the chances that you will draw a ball with a 50 and get 50 thalers are twice as big as the chance to draw a ball with an 80 or 120, simply because there are two balls with a 50 on it.

You may only draw one ball and only choose once per decision situation from which bag you want to draw. In the next decision situation you will be allowed to draw from another bag. In total, there are three such decision situations.

Everything ok so far? *(Leave some time for questions; answer questions individually and in private)*

Okay. If in this situation you think: "I would prefer to have a higher chance of a rather high payment, even if I could end up going home with the smallest amount"; from which bag would you like to draw here, which option do you prefer? Drawing out of the left or the right bag? *(Assuming that the answer is 'left')* Exactly, then you have to choose the left option. But if you think: "Even if the chances to draw the small amount are higher — it is not that small in comparison – and, besides, I could also draw the highest amount." – then you have to choose the option on the right.

To enter your decision, simply tap on the button below the bag from which you prefer to draw.

Everything clear so far? *(Leave some time for questions; answer questions individually and in private)*

(Black out slide show by pressing the 'B' key)

(Start subsession 'Urn')

(As soon as everyone has made their decisions) In the next part of our study we would like to ask you a few riddles. There are two types of riddles: The first type of riddle is to assign a number to a symbol. As fast as possible. The goal is to assign the correct number to as many symbols as possible in one and a half minutes. *(Show the slide of the presentation that displays the Symbol-Digit-Test)* Up here *(point at the allocation table)* you can see which number belongs to which symbol. You will always see this table. Here on the middle, *(point at the symbol in the center panel)* a symbol is randomly selected. Your task is to press *(point at the buttons)* the correct number as fast as possible. What is the correct number in this case? *(Assuming that the answer is 'eight')* Exactly, eight is the correct answer, and you have to choose eight here *(point at the button labelled with eight)*. Take good care of what you are pressing because there is no going back. If you pressed a number, the next task with the next symbol will come and you shall choose the corresponding number again. It takes a total of one and a half minutes and up here *(point at the time)* you can see how much time you have left.

For the riddles that you will play now and for those that you will play afterwards, those of you who solved the most riddles correctly get approximately (*mention the relevant amount only*)

- grade 6: 1,40 euros in total.
- grade 8: 1,75 euros in total.
- grade 10: 2,50 euros in total.
- grade 12: 3,90 euros in total.

If you solved fewer riddles correctly, you will get proportionally less; so make an effort!

Everything ok so far? *(Leave some time for questions; answer questions individually and in private)*

Okay, then you can play these riddles for a minute and a half now. As soon as you are ready, you can press 'next', but after 15 seconds at most you will be forwarded automatically, and then the time will run. A minute and a half, as many and as correct as possible.

Everything ok so far? *(Leave some time for questions; answer questions individually and in private)*

(Black out slide show by pressing the 'B' key)

(Start subsession 'Cognitive Ability 1')

(As soon as the time for the Symbol-Digit-Test is up) Now we turn to the second kind of riddles. Here, you will be shown some different patterns, and one pattern is always missing. There are several possibilities to fill in the gap, and the possibilities will be shown to you. You have to choose the number of the fitting possibility. We will show you a total of 10 such patterns and you will have 5 minutes to solve the riddles. *(Show the slide of the presentation that displays the matrix test.)* For example, such a riddle could like this: Up here *(point at the time)* you can see how much time you have left. Here *(point at pattern)* you can see the pattern. Here *(point at gap)* something is missing. Down here *(point at possible options)* you can see different possibilities to fill in the gap. In this example, which option is the correct one? *(Assuming that the answer is 'five')* Exactly, number five is the correct solution. So we choose five down here. Take good care of what you are pressing, because there is no going back. When you have pressed a number, the next riddle starts.

Everything ok so far? *(Leave some time for questions; answer questions individually and in private)*

Okay, then you can play these riddles for five minutes now. Again, you can press 'next' as soon as you are ready, just like before. After 15 seconds at most, however, you will be forwarded automatically. Then, your time will run.

Everything ok so far? *(Leave some time for questions; answer questions individually and in private)*

Alright, you may now start solving the riddles.

(Black out slide show by pressing the 'B' key)

(Start subsession 'Cognitive Ability 2')

(As soon as the time for the pattern riddles is up) Now, you can complete a questionnaire. If you have any questions, just raise your hand and one of us will come and help you. Most questions to tick can be answered relatively quickly. Just read the question and tick what you think. To give you a feeling of how long this should take: That is less than 2 minutes per page.

In the questionnaire, we will not ask for your name. That means we have no way of finding out who completed which questionnaire. So, it is completely anonymous. We only know that a questionnaire belongs to a person in this room, but we have no way of finding out to which person, once you have left the room.

There will be a number of questions where you can enter single letters from your name and your parents' names; e.g., the last letter of your fist name. We did so to be able to match your data, in case we will come cack in two years. You have this information and so you will be able to enter the same data again in two years. For us, however, it is impossible to do anything with it, because we do not know your parents' names. As I said, we do not even save your name, so we cannot figure out whose questionnaire it was. Therefore, this remains anonymous.

Everything ok so far? *(Leave some time for questions; answer questions individually and in private)*

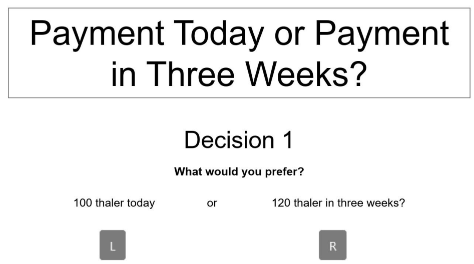
After you have completed the questionnaire, we will go through the rows and pay you. So please just remain seated.

(Start subsession 'Questionnaire')

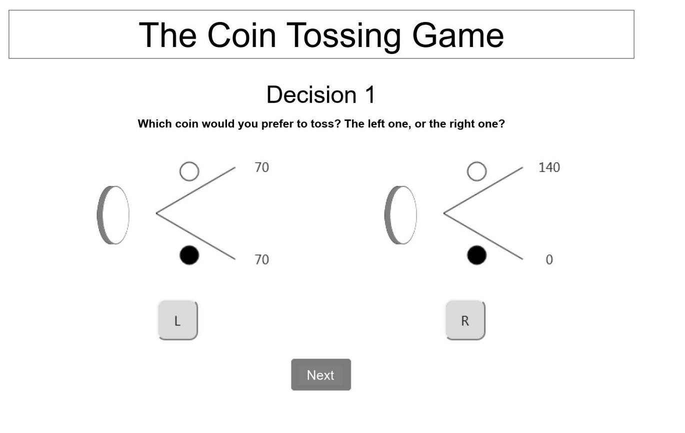
(As soon as all have completed the questionnaires)

F.2 Translated Illustration During Instructions

(a) Now or Later Task



(b) Coin Toss Task (Coins Were Animated And Rotating)



(c) Urn Task

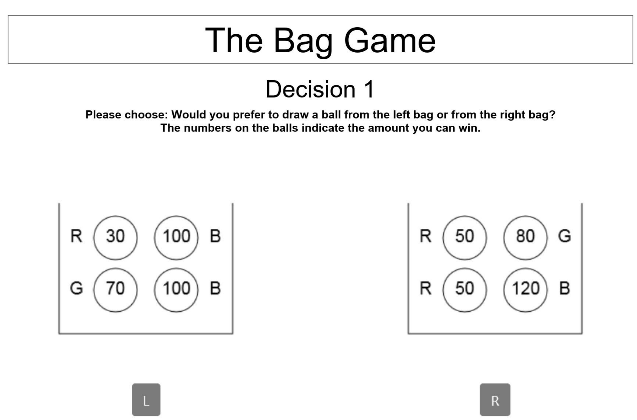


Figure F-1: Illustration of Decision Tasks Shown During Instructions

Notes: Instead of the amounts 30, 70, 100, 100 (Option A) and 50, 50, 80, 120 (Option B) used in the first task of the bag game, in the second task, the amounts 10, 90, 100, 100 (Option A) vs. 50, 50, 60, 140 (Option B) were used, and 60, 100, 100, 100 (Option A) vs. 80, 80, 80, 120 (Option B) were the corresponding amounts in the third task.

F.3 Original Instructions (German)

Hallo und herzlich willkommen zu unserer Studie. Schön, dass ihr hier seid und mitmachen wollt. In den nächsten 45 Minuten werden wir ein paar "Entscheidungsspiele" mit euch spielen, ihr dürft ein paar Rätsel bearbeiten und anschließend einen Fragebogen ausfüllen. Ihr dürft fast alles auf einem Tablet machen und wir erklären alles ausführlich der Reihe nach. Wir erklären, dann dürft ihr aktiv werden, und danach erklären wir den nächsten Schritt. Wir fangen mit den Spielen an.

(In der Session mit den älteren Schülern): Noch ein Kommentar zu den Erklärungen. Da wir eine wissenschaftliche Studie machen, ist es wichtig, dass wir immer die gleichen Erklärungen machen. Da wir die Studie auch mit jüngeren Schülern durchführen, sind die Erklärungen teilweise ausführlicher, als das sonst nötig wäre. Das hat also nichts mit euch zu tun, wenn euch das etwas länglich erscheint, aber wir müssen das so machen und es stellt auch sicher, dass ihr wirklich alles ganz genau versteht.

Redet ab jetzt bitte nicht mehr miteinander, lasst euer Handy wo es ist bzw. legt es weg, wenn ihr es gerade in der Hand haltet und hört gut zu. Ihr könnt in den Spielen Geld verdienen. Das Geld werden wir euch in bar am Ende des Experiments auszahlen oder aber ihr bekommt es in einem Briefumschlag – mehr dazu später. Wieviel Geld ihr verdienen könnt, hängt von euren Antworten und Entscheidungen ab. Daher ist es wichtig, dass ihr die Regeln versteht. Hört also bitte gut zu! Wir werden öfter eine Pause machen, sodass ihr Fragen stellen könnt. Hebt dazu einfach die Hand, einer von uns wird dann zu euch kommen um eure Frage zu beantworten.

Alles klar soweit? *(Zeit lassen für Fragen; Fragen persönlich und vertraulich beantworten)*

Im ersten Spiel sollt ihr viermal entscheiden, ob ihr einen bestimmten Geldbetrag lieber heute haben wollt, oder einen etwas größeren Geldbetrag in 3 Wochen. Hier seht ihr so eine Entscheidungssituation. *(Slide der Präsentation, die die Zeitpräferenz-Entscheidungssituation abbildet, zeigen.)* So sieht das dann aus. Hier links seht ihr den Geldbetrag, den ihr sofort bekommen würdet, in diesem Beispiel sind das 100 Taler. Rechts steht der Betrag, den ihr in drei Wochen bekommen würdet, im Beispiel 120 Taler.

Wenn ihr also sagt, für 20 Taler mehr, da würde ich bei einer Höhe von 100 Taler schon auch drei Wochen warten – welche Option müsst ihr dann wählen? *(Angenommen, die Antwort ist rechts)* Genau, dann müsst ihr die rechte Option wählen. Wenn ihr die 100 Taler allerdings lieber heute hättet, müsst ihr entsprechend die linke Option wählen.

Wir rechnen die Taler in Euro um, und 100 Taler sind ungefähr *(nur den relevanten Betrag nennen)*

- Klasse 6: 2 Euro.

- Klasse 8: 2,50 Euro.
- Klasse 10: 3,50 Euro.
- Klasse 12: 5,50 Euro.

Überlegt also gut, was euch lieber ist.

Eure Entscheidung könnt ihr einfach durch Tippen auf den "L" oder "R" Button eingeben.

Alles klar soweit? *(Zeit lassen für Fragen; Fragen persönlich und vertraulich beantworten)*

Zur Auszahlung: Zusätzlich zu den Entscheidungsspielen haben wir noch ein paar Rätselfragen. Pro richtig gelöstem Rätsel bekommt ihr zusätzlich Geld.

Wir werden außer diesem Spiel noch weitere zwei Arten von Spielen mit euch spielen. Ihr werdet dabei insgesamt rund 25 Entscheidungen treffen, und eine dieser Entscheidungen wird in echt ausgezahlt.

Aus den drei Arten von Entscheidungsspielen wählt euer Tablet zufällig eines aus und wählt außerdem zufällig die Nummer der Entscheidung aus. Da ihr bis zum Schluss nicht wissen werdet, welche Entscheidung ausbezahlt wird, ist es wichtig, dass ihr jede Entscheidung ernst nehmt.

Wenn dieses Spiel vom Tablet zufällig zur Auszahlung ausgewählt wird, dann bekommt ihr euer Geld entweder heute oder in drei Wochen – je nachdem, wie ihr entschieden habt.

Solltet ihr euch für eine Zahlung in drei Wochen entschieden haben und diese Entscheidung zufällig zur Auszahlung ausgewählt werden, könnt ihr das Geld in drei Wochen im Sekretariat abholen. *(Entsprechend der Abmachung mit der Schule anpassen.)*

Alles klar soweit? *(Zeit lassen für Fragen; Fragen persönlich und vertraulich beantworten)*

Falls eine oder einer von euch nicht teilnehmen möchte, lasst es uns bitte jetzt wissen. Ihr werdet auch später zu jedem Zeitpunkt aufhören können. Hebt dafür einfach die Hand – einer von uns kommt dann zu euch und bespricht das weitere Vorgehen.

Möchte jemand jetzt aufhören oder habt ihr Fragen? *(Zeit lassen für Meldungen bzw. Fragen; Fragen persönlich und vertraulich beantworten; bei Abbruch ID-Nummer des Tablets zusammen mit der jeweiligen Session notieren, um Datensatz löschen zu können.)*

Alles klar soweit? *(Zeit lassen für Fragen; Fragen persönlich und vertraulich beantworten)*

(Bildschirmpräsentation mit Druck auf Taste "B" auf schwarz stellen) Okay, dann werden wir jetzt die Entscheidungsspiele spielen.

(Session starten)

(Wenn alle soweit ihre Entscheidungen getätigt haben) Jetzt dürft ihr 18 mal entscheiden, ob ihr lieber einen bestimmten Betrag sicher haben oder aber mit uns eine Münze werfen wollt, und am Ende entweder einen höheren oder aber einen niedrigeren Betrag als den

sicheren Betragt haben wollt. Wir werden die Beträge in den 18 Entscheidungssituationen verändern.

Eine solche Situation sieht zum Beispiel so aus *(Slide der Präsentation, die die Münzwurf-Entscheidungssituation abbildet, zeigen)*. Wir haben hier links eine Münze, und egal, ob die Münze auf der weißen oder auf der schwarzen Seite zum Liegen kommt, bekommt ihr 70 Taler. Diesen Betrag bekommt ihr also sicher; das zeigen wir damit, dass sowohl für weiß *(auf oberen Pfeil zeigen)* als auch für schwarz *(auf unteren Pfeil zeigen)* am Ende 70 Taler stehen. Hier rechts *(auf rechte Option zeigen)* sieht das anders aus. Hier bekommt ihr 140 Taler, wenn die Münze weiß zeigt *(auf oberen Pfeil zeigen)*, also auf der schwarzen Seite liegt. Wenn die Münze nun aber schwarz zeigt *(auf unteren Pfeil zeigen)*, bekommt ihr 0 Taler – also nichts.

Ihr müsst also entscheiden, ob ihr lieber 70 Taler sicher nach Hause nehmen oder lieber die Chance haben wollt, 140 Taler zu bekommen, wobei ihr eben auch leer ausgehen könnt. Wenn ihr also sagt: "Ich möchte lieber die Chance haben, 140 Taler zu bekommen, und nehme das Risiko in Kauf, bei diesem Münzwurf auch leer auszugehen", welche Option müsst ihr dann wählen? *(Angenommen, die Antwort ist "rechts")* Genau, ihr müsst die rechte Option wählen. Andererseits, wenn ihr sagt, ihr wollt lieber auf Nummer Sicher gehen, dann müsst ihr die linke Option wählen.

Um eure Entscheidung einzugeben, tippt bitte einfach auf den Button unter der Option, die ihr lieber hättet. Weil die Entscheidungssituationen auf den ersten Blick sehr ähnlich aussehen, müsst ihr zusätzlich auf "Weiter" drücken *(auf "Weiter"-Button zeigen)*, um sicherzustellen, dass ihr nicht versehentlich noch einmal die gleiche Antwort für eine andere Situation wählt.

Alles klar soweit? *(Zeit lassen für Fragen; Fragen persönlich und vertraulich beantworten)*

Zur Auszahlung hier: Nehmen wir jetzt mal an, der Computer hätte die Entscheidung 1 des Münzwurfs ausgewählt.

Nehmen wir jetzt zusätzlich an, dass ihr euch für die linke Option entschieden hättet. Dann bekommt ihr einfach 70 Taler. Hättet ihr euch hingegen für die rechte Option entschieden, wirft euer Tablet eine Münze. Zeigt die Münze weiß, hättet ihr in diesem Beispiel also 140 Taler bekommen. Wie gesagt rechnen wir die Taler später in Euro um. 140 Taler sind das meiste, was ihr hier mit diesem Spiel verdienen könnt. Das sind ungefähr *(nur den relevanten Betrag nennen)*

- Klasse 6: 2,75 Euro.
- Klasse 8: 3,50 Euro.
- Klasse 10: 5,00 Euro.

- Klasse 12: 7,75 Euro.

Überlegt euch also gut, wie ihr entscheidet.

Alles klar soweit? *(Zeit lassen für Fragen; Fragen persönlich und vertraulich beantworten)*

(Bildschirmpräsentation mit Druck auf Taste „B“ auf schwarz stellen)

(Subsession Certainty Equivalents starten)

(Wenn alle soweit ihre Entscheidungen getätigt haben) Im letzten der drei Spiele müsst ihr in verschiedenen Situationen wählen, ob ihr lieber einen Ball aus einem Beutel, nennen wir ihn Beutel L für links, oder einen Ball aus einem anderen Beutel, den nennen wir Beutel R für rechts, ziehen wollt. Das passiert natürlich, ohne dass ihr in den Beutel schauen könnt, also ihr könnt euch nicht den Ball raussuchen, den ihr gerne hättet. Ihr zieht einen Ball zufällig. Die Beutel sehen jeweils so aus wie auf diesem Bild hier *(Slide der Präsentation, die die Urnen-Entscheidungssituation abbildet, zeigen)*. Ihr seht, in jedem Beutel sind vier Bälle. Die Zahl auf den Bällen gibt an, wieviel Taler ihr bekommt, wenn ihr den entsprechenden Ball zufällig zieht. Zum Beispiel hier *(auf Ball mit der 50, markiert mit R, zeigen)*, auf dem roten Ball – R steht für rot, G steht für grün, B steht für blau – da steht 50 drauf. Wenn ihr also diesen Ball zufällig zieht, bekommt ihr 50 Taler. Wie gesagt, ihr dürft nicht in den Beutel schauen, ihr könntet also jeden der vier Bälle ziehen, und die Chancen, jeden dieser Bälle zu ziehen, sind gleich. Das heißt wenn ihr hier aus dem rechten Beutel zieht, sind die Chancen, dass ihr einen Ball mit einer 50 zieht und 50 Taler bekommt, doppelt so groß, wie die Chance einen Ball mit einer 80 oder 120 zu ziehen, ganz einfach, weil hier zwei Bälle mit einer 50 drin sind.

Ihr dürft nur einen Ball ziehen und nur einmal pro Entscheidungssituation wählen, aus welchem Beutel ihr ziehen wollt. In der nächsten Entscheidungssituation dürft ihr dann wieder aus einem anderen Beutel ziehen. Insgesamt gibt es drei solcher Entscheidungssituationen.

Alles klar soweit? *(Zeit lassen für Fragen; Fragen persönlich und vertraulich beantworten)*

Okay. Wenn ihr in dieser Situation jetzt denkt: “Ich möchte lieber größere Chancen auf eine recht hohe Zahlung, auch wenn ich dabei am Ende mit dem kleinsten Betrag heim gehen könnte” aus welchem Beutel möchtet ihr dann hier ziehen, welche Option bevorzugt ihr? Ziehen aus dem Beutel links oder aus dem Beutel rechts? *(Angenommen, die Antwort ist links)* Genau, dann müsst ihr links wählen. Wenn ihr aber denkt: “Auch wenn die Chancen, den kleinen Betrag zu ziehen höher sind – so klein ist er im Vergleich auch nicht – und außerdem könnte ich ja auch den höchsten Betrag ziehen” — dann müsst ihr die rechte Option wählen.

Um eure Entscheidung einzugeben, tippt bitte einfach auf den Button unter dem Beutel, aus dem ihr lieber ziehen wollt.

Alles klar soweit? *(Zeit lassen für Fragen; Fragen persönlich und vertraulich beantworten)*

(Bildschirmpräsentation mit Druck auf Taste "B" auf schwarz stellen)

(Subsession Urn starten)

(Wenn alle soweit ihre Entscheidungen getätigt haben) Im nächsten Teil unserer Studie wollen wir euch ein paar Rätselfragen stellen. Wir haben zwei Arten von Rätselfragen mitgebracht: Die erste Art von Rätsel besteht darin, einem Symbol eine Zahl zuzuordnen. Und zwar möglichst schnell. Das Ziel ist, in eineinhalb Minuten so vielen Symbolen wie möglich die korrekte Zahl zuzuordnen. *(Slide der Präsentation, die den Symbol-Digit-Test abbildet, zeigen.)* Hier oben *(auf Zuordnungstabelle zeigen)* seht ihr, welche Zahl zu welchem Symbol gehört. Diese Tabelle werdet ihr immer sehen. Hier in der Mitte *(auf Symbol in der Mitte zeigen)* wird dann zufällig ein Symbol ausgewählt. Eure Aufgabe ist es nun, so schnell wie möglich hier unten *(auf die Buttons zeigen)* die richtige Zahl zu drücken. Was ist jetzt hier die richtige Zahl? *(Angenommen, die Antwort ist Acht)* Genau, hier ist Acht richtig, und ihr müsst die Acht hier wählen *(auf Acht zeigen)*. Passt gut auf, was ihr drückt, weil es hier kein Zurück gibt. Wenn ihr eine Zahl gedrückt habt, kommt die nächste Aufgabe mit dem nächsten Symbol und ihr sollt wieder die zugehörige Zahl wählen. Insgesamt dauert das eineinhalb Minuten und hier oben *(auf Zeit zeigen)* seht ihr, wie viel Zeit ihr insgesamt noch habt.

Zusammen für die Rätselspiele, die ihr gleich spielt, und die, die danach kommen, bekommen diejenigen, die am meisten richtig haben, ungefähr *(nur den relevanten Betrag nennen)*

- Klasse 6: 1,40 Euro.
- Klasse 8: 1,75 Euro.
- Klasse 10: 2,50 Euro.
- Klasse 12: 3,90 Euro.

Wer weniger Rätsel richtig gelöst hat, bekommt entsprechend weniger; gebt euch also Mühe!

Alles klar soweit? *(Zeit lassen für Fragen; Fragen persönlich und vertraulich beantworten)*

Okay, dann dürft ihr diese Rätsel jetzt für eineinhalb Minuten spielen. Wenn ihr bereit seid, könnt ihr "Weiter" drücken, aber spätestens nach 15 Sekunden geht es auch automatisch weiter, und ab dann läuft die Zeit. Eineinhalb Minuten, so viel und so richtig wie möglich.

Alles klar soweit? *(Zeit lassen für Fragen; Fragen persönlich und vertraulich beantworten)*

(Bildschirmpräsentation mit Druck auf Taste "B" auf schwarz stellen)

(Subsession "Cognitive Ability 1" starten)

(Wenn die Zeit zur Beantwortung des Symbol-Digit-Tests abgelaufen ist) Nun kommen wir zur zweiten Art von Rätselfragen. Hier bekommt ihr verschiedene Muster gezeigt, und

jeweils ein Muster fehlt. Es gibt verschiedene Möglichkeiten, die fehlende Stelle auszufüllen, und diese Möglichkeiten werden euch angezeigt. Ihr sollt dann die Nummer der Möglichkeit wählen, die passt. Wir zeigen euch insgesamt 10 solcher Muster und ihr habt 5 Minuten Zeit. *(Slide der Präsentation, die den Matrizentest abbildet, zeigen.)* Das sieht zum Beispiel so aus: Hier oben *(auf Zeit zeigen)* seht ihr, wieviel Zeit ihr noch habt. Hier *(auf Muster zeigen)* seht ihr das Muster. Hier *(auf Lücke zeigen)* fehlt etwas. Hier unten *(auf mögliche Optionen zeigen)* seht ihr verschiedene Möglichkeiten, um die fehlende Stelle auszufüllen. Welche ist in diesem Beispiel die richtige Möglichkeit? *(Angenommen, die Antwort ist "fünf")* Genau, die fünfte ist die richtige Lösung. Wir wählen hier unten also Fünf aus. Passt gut auf, was ihr drückt, weil es hier kein Zurück gibt. Wenn ihr eine Zahl gedrückt habt, kommt das nächste Musterrätsel für euch.

Alles klar soweit? *(Zeit lassen für Fragen; Fragen persönlich und vertraulich beantworten)*

Okay, dann dürft ihr diese Rätsel jetzt für fünf Minuten spielen. Wenn ihr bereit seid, könnt ihr wieder "Weiter" drücken, wie vorhin auch schon. Nach spätestens 15 Sekunden geht es aber auch automatisch weiter. Dann läuft eure Zeit.

Alles klar soweit? *(Zeit lassen für Fragen; Fragen persönlich und vertraulich beantworten)*

Dann könnt ihr gleich mit den letzten Rätselfragen starten.

(Bildschirmpräsentation mit Druck auf Taste "B" auf schwarz stellen)

(Subsession „Cognitive Ability 2“ starten)

(Wenn die Zeit zur Beantwortung der Muster-Rätsel abgelaufen ist) Nun dürft ihr noch einen Fragebogen beantworten. Wenn ihr dabei Fragen habt, hebt bitte einfach die Hand, einer von uns kommt dann zu euch und hilft euch. Die meisten Fragen zum Ankreuzen sind recht schnell zu beantworten. Lest einfach die Frage, und kreuzt an, was ihr denkt. Um euch ein Gefühl zu geben, wie lange das dauern sollte: Das sind pro Seite, die gezeigt wird, unter 2 Minuten.

Wir fragen auf dem Fragebogen nicht nach eurem Namen. Das heißt, wir haben keine Möglichkeit, herauszufinden, wer welchen Fragebogen ausgefüllt hat. Das ist also komplett anonym. Wir wissen lediglich, dass ein Fragebogen zu einer Person hier im Raum gehört, haben aber keine Möglichkeit herauszufinden, zu welcher Person, sobald ihr den Raum verlassen habt.

Es wird eine Reihe von Fragen geben, bei denen ihr einzelne Buchstaben aus euren Namen und den Namen eurer Eltern angeben dürft; z.B. den letzten Buchstaben eures Vornamens. Das haben wir gemacht, falls wir in zwei Jahren wieder kommen, um eure Daten zusammen bringen zu können. Diese Informationen habt ihr und könnt damit in zwei Jahren dieselben Daten wieder angeben. Für uns ist es allerdings unmöglich, damit etwas anzufangen, weil wir ja nicht wissen, wie eure Eltern heißen. Wir speichern ja wie gesagt nicht einmal euren

Namen, können also nicht darauf kommen, wessen Fragebogen das war. Das bleibt also dadurch anonym.

Alles klar soweit? *(Zeit lassen für Fragen; Fragen persönlich und vertraulich beantworten)*

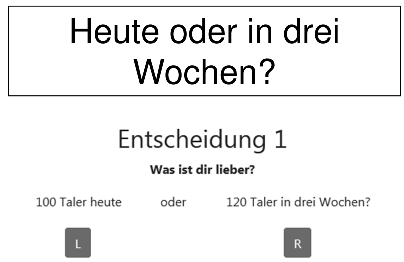
Im Anschluss daran werden wir durch die Reihen gehen und euch bezahlen. Bleibt bitte also einfach sitzen.

(Subsession "Questionnaire" starten)

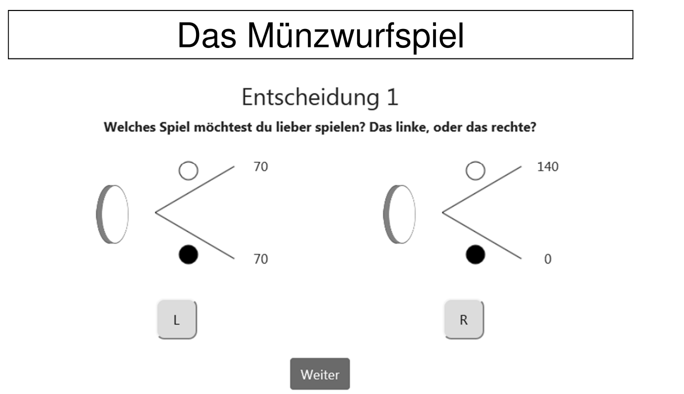
(Wenn alle Fragebögen ausgefüllt wurden)

F.4 Illustration During Instructions (German)

(a) Now or Later Task



(b) Coin Toss Task (Coins Were Animated And Rotating)



(c) Urn Task

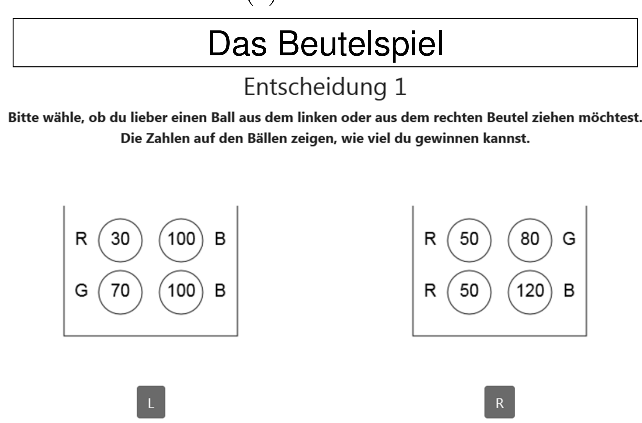


Figure F-2: Illustration of Decision Tasks Shown During Instructions