Higher Order Risk Preferences: New Measures, Determinants and Field Behavior*

Sebastian O. Schneider

Max Planck Institute for Research on Collective Goods, Bonn
sschneider@coll.mpg.de

Matthias Sutter

Max Planck Institute for Research on Collective Goods, Bonn; University of Cologne, IZA Bonn, CESifo Munich & University of Innsbruck
sutter@coll.mpg.de

August 12, 2021

Abstract

We present and use a novel method to elicit and measure the intensities of higher order risk preferences (prudence and temperance) in an experiment with 658 adolescents. In line with theory, we find that higher order risk preferences are strongly related to adolescents' field behavior, including prevention, health-related, and addictive behavior, involving amongst others excessive smartphone usage, or financial decision making. Most importantly, we illustrate how ignoring prudence and temperance might yield largely misleading conclusions about the relation of risk preferences to field behavior. Thus we can put previous work that ignored higher order risk preferences into an encompassing perspective.

Keywords: Higher order risk preferences, prudence, temperance, risk aversion, method, intensity, field behavior, adolescents, health, lab-in-the-field experiment
JEL classification: C91, D01, D81, D91, I12

*We thank Zvonimir Bašić, Gary Charness, Sebastian Ebert, Helga Fehr-Duda, Matteo Galizzi, Lukas Kiessling, Charles Noussair, Hannah Schildberg-Hörisch, Stefan Trautmann, Daniel Wiesen, and participants at the CREED/MPI workshop in Bonn, IMEBESS in Utrecht, 5th M-BEPS, 18th TIBER Symposium, European and North American ESA in Dijon and Los Angeles, Post-ESA Workshop on Experiments with Children and Non-Standard Subjects, NCBEE in Kiel, BEH-net Workshop in Innsbruck, World Congress of the Econometric Society in Milan, Annual Meeting of the German Economic Association in Cologne, CESifo Area Conference on Behavioral Economics, iHEA World Congress, as well as seminar participants at George Mason University, Mainz University, UC Santa Barbara and Boston University for helpful comments and suggestions, as well as Christa Kuon-Rehm, Karl Müller, Christian Neubauer, Eva Schneider and Klaus Schneiderhan for their support and coordination of the study in their respective schools. This study was approved by the IRB of the University of Innsbruck and it was preregistered with the open science foundation (osf.io/n7v2y). Financial support from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2126/1– 390838866 is gratefully acknowledged.

1 Introduction

Risk is an inherent part of life: Decisions about occupation, education, finances or health behavior, to name just a few, regularly involve at least some degree of risk. Consequently, measuring risk preferences is important for a better understanding of behavior in various domains of life. For many years, empirical research has therefore investigated links between risk measures – both incentivized experimental ones as well as hypothetical survey questions – and life outcomes. While a plethora of research has documented a connection between commonly used measures of risk aversion and health outcomes, educational and financial success, or occupational choices (see, e.g., Barsky et al., 1997; Bonin et al., 2007; Anderson and Mellor, 2008; Caliendo et al., 2007; Dohmen et al., 2011; Becker et al., 2012), there is also plenty of work that fails to predict field behavior with subjects' risk aversion (e.g., Sutter et al., 2013; Galizzi et al., 2016; Charness et al., 2020; Samek et al., 2021), or reports rather weak relations between experimental measures and standard survey questions on risk tolerance (see, e.g., Crosetto, 2021, for a meta study).

The somewhat mixed results on the relationship between risk aversion and field behavior may be related to measurement error, causing imperfect correlations (Gillen et al., 2019). However, another potential explanation beyond measurement error is that also risk comes in different forms, analogously to timing decisions (e.g., Laibson, 1997). In fact, measurement error left aside for the moment, the papers on risk preferences cited above have focused on measuring risk aversion and then relating it to field behavior. Contrary to such an approach (that still can be considered as the standard approach in the literature), a growing body of theoretical contributions going back already to Leland (1968) suggests that for explaining a broad range of behaviors, including financial decision making and health-related behavior, higher order risk preferences might be more relevant than standard measures of risk aversion (Kimball, 1990, 1992; Gollier and Pratt, 1996; Courbage and Rey, 2006). A neglect of higher order risk preferences might therefore explain at least partially the often weak and mixed relation between risk preferences and field behavior.

In this paper, we present and use a novel and in-sample validated method to study and quantify higher order risk preferences and to relate them to field behavior. The method accounts for measurement error and is the first to yield non-parametric utility-based intensity measures in the spirit of the Arrow-Pratt measure for risk aversion. As these are the measures applied in related theoretical work, the method provides the ideal basis for theory testing, allowing us to relate our intensity measures of higher order risk preferences to several domains of field behavior, including health and addictive behavior, financial decision making, and eco-friendly behavior.

Besides risk aversion, we focus on prudence and temperance as higher order risk preferences.

Prudence, the third order risk preference, is often defined as the preference to allocate a mean-zero risk to the state of higher wealth instead of to the state of lower wealth (Eeckhoudt and Schlesinger, 2006). An intensity measure of prudence has also been interpreted as a measure of left-skewness aversion or equivalently a preference for right-skewness (Modica and Scarsini, 2005). Under expected utility theory, prudence is equivalent to downside risk aversion (Menezes et al., 1980). Temperance, the fourth order risk preference, can be defined as the preference to disallocate two independent mean-zero risks across two states of the world opposed to accepting both of them in the same state of the world (Eeckhoudt and Schlesinger, 2006). An intensity measure of temperance has been interpreted as a measure of kurtosis aversion (Denuit and Eeckhoudt, 2010).

Via these skewness preference and the kurtosis aversion measures[1], higher order risk preferences thus capture important aspects of the distribution of a risk – beyond its mean and variance –, which suggests that they deserve attention when studying risky behavior in the field. Moreover, in several theoretical models, the intensities of higher order risk preferences have been linked to various types of behaviors, including health-related behavior (e.g., Courbage and Rey, 2006), eco-friendly behavior (e.g., Bramoullé and Treich, 2009), and prevention effort to lower the probability of an undesired event (Eeckhoudt and Gollier, 2005; Menegatti, 2009). Yet, except for their relation to financial decision making (Noussair et al., 2014), there are no empirical studies on the relationship between higher order risk preferences and field behavior in these different domains (Trautmann and van de Kuilen, 2018).

Here, we measure the intensities of the higher order risk preferences prudence and temperance as well as risk aversion with a novel method, based on machine learning techniques, developed for and first used in this and its companion paper (Schneider et al., 2021). The companion paper discusses the statistical details of the method, and cross-validates and compares the available elicitation methods of higher order risk preferences in a sample of the general population. In this paper, we focus on the applicability of the method and the resulting predictive quality for field behavior, as it yields intensity measures, and corrects for measurement error, thus achieving a precise measurement, especially suited for and relevant among adolescents. The method, as implemented here, elicits higher order risk preferences in an intuitive way that is cognitively very accessible for participants. Additionally, it is the only method that yields (aforementioned) utility-based intensity measures

[1]See also Ebert (2012) on the moment characterizations of higher order risk preferences.

of higher order risk preferences non-parametrically. This is important, as parametric approaches are not flexible enough for the study of higher order risk preferences at the individual level (see, e.g., the discussion on this topic in light of their results in Noussair et al., 2014). With our method, participants are repeatedly asked to decide between a safe outcome and a lottery in order to elicit utility points. These are subsequently connected to a utility function with a spline regression approach accounting for measurement error via cross-validation. We have designed this approach specifically for the use with utility functions; it is the (statistical) core of the method. Based on estimated utility functions, utility-based intensity measures of higher order risk preferences can be computed. We validate the method in our sample with a reduced-lottery version of the standard method by Eeckhoudt and Schlesinger (2006).

In the first part of the paper, we examine the distribution and the determinants of higher order risk preferences (and of risk aversion) in a sample of 658 children and adolescents, aged 10 to 21 years. Importantly, all experimental sessions were conducted during regular school-hours with school classes randomly selected within schools, thereby avoiding self-selection into the experiment.[2] Besides measuring higher order risk preferences, we control for cognitive abilities, family characteristics, and time preferences (see, e.g., Epper and Fehr-Duda, 2018, on the importance of accounting for time discounting when studying risk-taking behavior). In the second part of the paper, we then relate these individual experimental measures to self-reported behavior in the field, including general risk taking, health-related behavior, with a focus on addictive behavior, general prevention, eco-friendly behavior, and financial decision making. In the third part of the paper we show that ignoring higher order risk preferences would yield largely misleading results about the relationship of risk aversion to field behavior.

The age group covered in our sample spans the formative years for many habits that shape these adolescents' future prospects. For example, smoking, drinking, or addictive gambling in the teenage years has high predictive power for also showing such behavior in adulthood (DeWit, 2000; Paul et al., 2008; Buchmann et al., 2011; Black et al., 2015). For this reason, it is important for potential interventions to study this age group and learn which factors are predictive of such behaviors in order to help identifying youths at risk.

Our paper contributes to the literature in several ways. The first contribution is to provide a unified experimental framework to measure higher order risk preferences (and risk aversion) for a large sample of children and adolescents, using a

[2]Only 4.2% of eligible pupils did not obtain consent from their parents to participate in our study, and no pupil opted out in any of our sessions.

novel method, developed for this study, to elicit non-parametric utility-based intensity measures. While risk aversion and corresponding intensity measures, such as the Arrow-Pratt measure, have been studied extensively, including this age group (see Sutter et al., 2019, for a review), for prudence and temperance the evidence is scarce or non-existent. There is only one paper that investigates prudence in children and adolescents, yet missing a proper intensity measure: Heinrich and Shachat (2020) let 362 Chinese children and adolescents, aged 8 to 17 years, decide in three binary choice tasks and take the number of prudent decisions as their measure of prudence. That is, like in most of the empirical work on higher order risk preferences, Heinrich and Shachat (2020) do not measure intensities of prudence but approximate intensity with consistency in choice. In contrast to our study, they do not examine theoretically predicted relations of the degree of prudence to field behavior, but investigate determinants of risk aversion and of their measure of prudence and examine transmission of choices from parents to children. Unlike us, Heinrich and Shachat (2020) ignore temperance, which has, in fact, never been studied with children or adolescents. Moreover, with respect to any proper intensity measure of higher order risk preferences, the evidence remains scarce even for adults, with the notable exception of Ebert and Wiesen (2014).

Our second contribution is that we are the first to connect higher order risk preferences with field behavior of adolescents. While in theory, higher order risk preferences have been predicted to relate to various domains of field behavior, there have been no studies so far to test these relationships for adolescents. Even for adults, empirical tests are scarce and confined to the seminal study of Noussair et al. (2014) about how higher order risk preferences of a sample of Dutch adults relate to their financial decisions, albeit they lack proper intensity measures like those resulting from our method. While we also consider the financial behavior of our adolescents, our focus lies on the domains of field behavior that might be considered most formative in this age group. Thus, a major domain under consideration is health-related behavior. Besides smoking and drinking behavior, we focus on a relatively new phenomenon, namely excessive smartphone usage, as this has increasingly been linked to mental health issues, like depression, and poor well-being of adolescents (Przybylski and Weinstein, 2017; Twenge et al., 2018; Orben and Przybylski, 2019). In addition to health-related behavior and financial decision making, we also examine adolescents' behavior towards the environment, their preventive behavior to avoid undesired events, and their general risk-taking behavior.

Our third, and arguably most important, contribution relates back to the often weak or inconclusive results on the relation of experimental risk aversion measures to field behavior (e.g., Sutter et al., 2013; Charness et al., 2020). We argue that

indecisive empirical results may be due to omitting prudence and temperance – even above and beyond measurement error issues – at least in the domains of behavior that are related to higher order risk preferences. Actually, we can show that the omission of prudence and temperance masks the true relation of risk preferences in several cases when we relate them to field behavior. For example, we show how neglecting higher order risk preferences even results in wrongly concluding that health-related behavior, among others, is unrelated with risk – whereas, in reality, only a measure of prudence is missing. Moreover, we find that the commonly used one-item survey measure on risk tolerance (e.g., Falk et al., 2018) successfully relates to field behavior. Yet, we can show that this is the case because this survey measure captures prudence and temperance, but not risk aversion. These insights demonstrate the importance of including higher order risk preferences in empirical analyses and put previous work into a more encompassing perspective that may help to avoid misleading conclusions about the relation of risk preferences and other characteristics to related field behavior.

Our experimental results with respect to the classification of higher order risk preferences are in line with findings on adult populations (see, e.g., the review by Trautmann and van de Kuilen, 2018). In the aggregate, children and adolescents are risk averse, prudent, and temperant. We find no age effect on the intensity of any (higher order) risk preference, which replicates earlier findings on risk aversion of adolescents (see, e.g., the review by Sutter et al., 2019) and of adults (Noussair et al., 2014). Moreover, we replicate the standard finding with respect to gender (e.g., Croson and Gneezy, 2009; Sutter et al., 2019): females are more risk averse than males. Our findings indicate that this pattern extends to higher orders of risk, as females are also more prudent and more temperant than males, which has also been reported with adult populations (Ebert and Wiesen, 2014). High-ability students are less risk averse and less temperant, replicating a common finding for adults with respect to risk aversion (see the review by Dohmen et al., 2018).

Turning to the relationship of our experimental measures to behavior in the field, we find that prudence and, to a lesser extent, temperance are very important for predicting field behavior. In addition to extending earlier results with respect to financial decision making from adults (Noussair et al., 2014) to adolescents, we uncover many novel insights. Most importantly, prudence is strongly related to health-related behavior, but risk aversion is not. For example, our index capturing obsessive use of smartphones is predicted significantly by prudence, but not by risk aversion or temperance. Applying the ORIV technique (Gillen et al., 2019) provides further evidence that this result is not driven by measurement error. We make the same observation when looking at a broader set of health-related behaviors, in-

cluding smoking and drinking, or when forming a general health index that also considers, e.g., the body mass index or the regularity at which participants practice sports. Prudence also matters for other domains of field behavior, such as prevention behavior and general risk taking. However, the relation of risk preferences to field behavior depends strongly on whether or not higher order risk preferences are accounted for, and ignoring higher order risk preferences would lead to largely misleading results. We consider this a key insight of our paper.

In the next section, we describe our subject pool, the general features of the experiment and the method with which we measure higher order risk preferences. Section 3 presents the results on risk aversion, prudence and temperance, and how they depend on socio-demographic variables and cognitive abilities. In Section 4, we introduce the different domains of field behavior that we elicit in our experimental questionnaire, and present what theoretical models predict about their relationship to higher order risk preferences. Section 5 studies the relation between our measures of higher order risk preferences and field behavior. In Section 6 we illustrate the importance of considering prudence and temperance to identify the true relationship of risk preferences to various domains of field behavior. Section 7 discusses our main results and concludes the paper.

2 Methods and Experimental Design

2.1 General Setup

Our study was approved by the IRB of the University of Innsbruck and preregistered with the open science foundation (osf.io/n7v2y), including a pre-analysis plan.

Subject Pool We ran the experiment in four German schools in the federal states Baden-Württemberg, North Rhine-Westphalia and Rhineland-Palatinate in September and October 2018. In every school, classes were selected randomly such that at least one class per grade from grades six, eight, ten, and twelve participated in the study. In total, 658 children and adolescents, aged 10 to 21 years, took part in our experiment. The distribution of adolescents across grades and gender and their average age per grade is summarized in Table 1. Principals and teachers of the participating schools supported our study by allowing us to conduct the experiment in class during regular school hours. Schools made sure that all participating children obtained their parents' informed consent to participate, and more than 95% of parents gave their permission. Adolescents were also asked whether they would be willing to participate in the experiment and no adolescent opted out.

Table 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

General Experimental Setup The whole study was run on tablet computers. First, we elicited adolescents' risk and time preferences in an incentivized experiment. Afterwards, they performed some tasks to measure cognitive abilities (see below) and filled in an extensive survey on field behavior (see Section 4 for a description and Online Appendix C for the entire questionnaire). In the experiment, adolescents could earn "Taler" as our experimental currency. We explained the conversion rate from Taler to Euro carefully and varied it depending on the grade, such that the maximal earnings corresponded to 120% of the weekly amount of pocket money recommended by the German Federal Ministry of Family Affairs for the different age cohorts (see Familien-wegweiser.de, 2018). This was done in order to hold the relative value of a Taler constant across the different age cohorts. For example, the highest possible payment of 280 Taler corresponded to €5.50, €7, €10 and €15.50 for grades 6, 8, 10 and 12, respectively. This includes a show-up fee of 70 Taler and up to 70 Taler for the cognitive ability tasks that were conducted after the experimental elicitation of risk and time preferences.[3]

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 eight 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 with the time given for a task, i.e. (number of matrices solved * 300 + number of correct symbol-digit pairs * 90)/(300 + 90). Finally, for comparison reasons, we center and standardize this measure.

Instructions were the same in every session (see Section S-2 of the supplementary material on our website[4]) and were orally delivered by the first author. We paid

[3]For the cognitive ability task, we always paid 70 Taler to the best student in the classroom. All other students were paid proportionally to the best student.

[4]Supplementary material is available at `http://sebastianoschneider.com/media/supplementary_material/SchneiderSutter2021.pdf`

all participants in cash before they left the classroom, with the exception of future payments in the time preference experiment (described below).

2.2 Elicitation of Certainty and Future Equivalents

The elicitation of risk and time preferences in our experiment is based on the elicitation of indifference values. For risk preferences, we elicit the certain amount of money that makes participants indifferent between playing a lottery and receiving a certain amount of money. This is the foundation of our elicitation of higher order risk preferences, as we will explain in detail in the next subsection. For time preferences, we elicit the amount of money to be paid at the later date that makes them indifferent between receiving a smaller amount of money at the day of the experiment or a larger amount with a three weeks delay. We elicit indifference values using a bisection approach, also called staircase method. This approach is widely used in the economics literature (e.g., Falk et al., 2018) and very easy to understand for participants. Participants are faced with one decision between two options at a time. For the risk elicitation, subjects are presented a choice between a sure payoff and a lottery with two equally likely outcomes. The first choice in the first decision task was between 70 Taler for sure, and an equal chance of getting zero or 140 Taler (see Figure A-1 in Appendix A for a screenshot of this first decision task). If a subject chose the sure payoff, the amount of the sure payoff would be decreased in the next iteration, whereas if she chose the lottery, the sure payoff would be increased for the next decision to be taken. From three such iterations, we deduce indifference values for a specific lottery, the so-called certainty equivalents.[5] In total, we elicit six certainty equivalents for six lotteries (see the next subsection for more details).

For time preferences, one option consists of a certain amount at the day of the experiment, and the other option consists of a larger amount with a three weeks delay. Depending on the choice, the amount paid with a three weeks delay is either increased or decreased, and the decision is repeated. For time preferences, we iterate this step four times to calculate the future equivalent.[6]

We have devoted considerable care to optimize the understanding of our experimental tasks for our sample of adolescents since noise in elicited preferences obviously impedes precise predictions of field behavior (Gillen et al., 2019) and because complexity of an elicitation task can affect measured preferences, even to the extent that it masks existing patterns in the sample. For example, Charness et al. (2018) show that multiple price lists produce enough noise through confusion and inconsistencies to mask a gender difference in risk taking that is found when only a

[5] See Appendix A for an example illustrating in detail the computation of a certainty equivalent for a lottery depending on participants' decisions.

[6] See Appendix A for details on the computation of a future equivalent.

single decision of the choice list is used.[7] Therefore, we do not apply choice lists but ask for one decision at a time.[8] Additionally, we illustrate the risky decision task using a rotating coin to resemble well-known every day decision situations even for our youngest age cohort.

In total, adolescents made 18 decisions between a sure amount and a lottery with two equally likely outcomes, and four decisions between an earlier payoff and a later payoff.[9] Among all decision tasks, one was randomly selected for payment by the computer. If one of the certainty equivalence tasks was selected, and the participant chose the lottery, a coin flip was simulated by the computer to determine the realization of the lottery. The payoffs in these tasks ranged from 0 to 140 Taler. If one of the time preference tasks was selected for payment and the later payment was chosen, it was handed over to the student at the prespecified date. The delivery of delayed payments was guaranteed by headmasters in front of the children, and either secretaries, teachers or headmasters themselves administered the payment in an anonymized way. The payoffs in the time preference tasks ranged from 100 to 140 Taler.[10] As our measure of time preferences (i.e., impatience) we compute the ratio of the future equivalent of the earlier payoff to the early payoff. In all time preference questions, we used 100 Taler as the early payoff option. The measures applied for higher order risk preferences are explained in the following.

2.3 Experimental Measurement of Higher Order Risk Preferences

Higher order risk preferences are now often defined by preferences over the allocation of zero-mean lotteries (Eeckhoudt and Schlesinger, 2006), and recent experimental work to elicit prudence and temperance has built predominantly on these definitions, starting with Deck and Schlesinger (2010), Ebert and Wiesen (2011) and Maier and Rüger (2011). While preferences over these lotteries are model free, they have no specific meaning outside expected utility theory (Eeckhoudt et al., 2020). Under expected utility theory, these definitions are equivalent to definitions based on derivatives of the utility function. For example, just as risk aversion can be

[7] Moreover, estimated preferences from a multiple price list might be biased due to the compromise effect. Suggested solutions rely on parametric approaches (e.g., Beauchamp et al., 2019) that are not flexible enough in our context (as we explain below).

[8] For its adaptive nature, the staircase method is to a certain degree prone to error-propagation. This fact has been considered when developing the method for elicitation of (higher order) risk preferences that we use by building on a spline regression approach, see below. The fact that we find considerable correlations between these measures and field behavior, survey questions, and alternatively elicited measures suggests that error-propagation is of minor concern in our study.

[9] In addition, we included three additional tasks to validate our measure of prudence, see below.

[10] As our focus lies on (higher order) risk preferences here, we omit comparison tables with respect to time preferences in the results section. However, our results are close to identical to those reported in Sutter et al. (2013): For a three-weeks delay, Sutter et al. (2013) report a median annual discount rate of 179%, where we find 176% in this study.

defined based on a negative second, *prudence* is defined via a positive third, and *temperance* is defined as having a negative fourth derivative of the utility function.[11] In fact, showing these equivalences – departing from the utility-based definitions – was the big contribution of Eeckhoudt and Schlesinger (2006), thereby presenting the first accessible way to elicit higher order risk preferences experimentally.

Choice of Method As interpretation of choice behavior with respect to higher order risk preferences is restricted to expected utility theory anyways (Eeckhoudt et al., 2020), we present and use a novel method here that builds on the utility-based definitions. It is the only method that yields utility-based intensity measures of higher order risk preferences (as used in theoretical work) in a non-parametric way. Theoretically, for classification of subjects as risk averse, prudent, or temperant, this method is equivalent to the method by Eeckhoudt and Schlesinger (2006). For an empirical validation, see the subsection on *Identification and Validation* below. In addition to classifying subjects with respect to higher order risk preferences, and contrarily to the method by Eeckhoudt and Schlesinger (2006), our method accounts for measurement errors by virtue of a custom-made (spline) regression approach (see the subsection on *Measurement Error* below), and measures individuals' *strength* of higher order risk preferences. Moreover, our method – as implemented here – is very intuitive and cognitively accessible, even to a sample of adolescents. Participants only have to choose repeatedly between a safe outcome and a two-outcome lottery with equiprobable outcomes.

By this choice, we also avoid the high and increasing complexity inherent in the so-called risk apportionment method (Eeckhoudt and Schlesinger, 2006), where with every order of risk attitude, an additional compound lottery is introduced. For example, to measure temperance, each option consists of the combination of three lotteries. Additionally, choice behavior in these compound tasks might reflect a mixture of higher order risk preferences and ambiguity aversion, as the latter seems to be almost identical to aversion to compound lotteries (Chapman et al., 2018; Gillen et al., 2019). Reduction of the compound lotteries is possible (e.g., Maier and Rüger, 2011; Heinrich and Shachat, 2020), but results in comparing much more demanding options with four (prudence) and eight (temperance) outcomes. The results from Heinrich and Shachat (2020), who study prudence among adolescents, calls for caution in using this method. They find that for their youngest subjects (third graders), choices are not different from random choices across all the three

[11] Higher orders also exist, but we are not aware of any behavioral consequence that has been attributed to, for example, *edginess* (positive fifth derivative), nor do previous results on their prevalence encourage further investigation (Deck and Schlesinger, 2014).

tasks they use to elicit prudence. Moreover, for one of the three tasks, behavior cannot be distinguished from random choices even when pooling all age cohorts.

The continuous, utility-based intensity measures that we use here are obtained in a non-parametric way, and have a clear relation to economic theory. Since theoretical work has relied almost exclusively on utility-based measures, our method is particularly suited for theory testing, as it yields, unlike the much more complex approach by Ebert and Wiesen (2014), exactly these measures – in a *direct way* (see *Attitude Measures* below). Commonly used alternative measures of *intensities* of higher order risk preferences are either based on the argument that consistency in choice equals intensity of a certain trait, where the relation to a utility-based coefficient is still unclear (e.g., Noussair et al., 2014; Heinrich and Shachat, 2020) or rely on parametric assumptions that imply a dependence of the different derivatives.[12]

Method Our method builds on the elicitation of utility points, for which we use the certainty equivalent method here with equally likely outcomes for its high suitability for our sample. We now describe how we can elicit utility points from eliciting certainty equivalents. First we normalize the utility function, such that for the highest possible outcome of $x_{\max} = 140$ Taler we assume $u(x_{\max}) = 1$ and for $x_{\min} = 0$ Taler, the lowest possible outcome, we have $u(x_{\min}) = 0$. Then the expected value of a lottery with these two equally likely outcomes is $0.5u(x_{\max}) + 0.5u(x_{\min}) = 0.5$. As a subject should be indifferent between receiving the elicited certainty equivalent $CE_{.5}$ and the lottery, the utility to her must be the same, thus we have $u(CE_{.5}) = 0.5$. Iterating this procedure, and taking $CE_{.5}$ as either the high outcome of the lottery (where the low outcome remains $x_{\min} = 0$ Taler), or as the low outcome (where the high outcome remains $x_{\max} = 140$ Taler), we also elicit individual utility points $u(CE_x) = x$ for $x = .25$ and $x = .75$. Additionally, depending on the differences between certainty equivalents, we elicit either $CE_{.125}$ or $CE_{.375}$, and either $CE_{.625}$ or $CE_{.875}$. Figure 1 provides an illustration: the distance between $CE_{.5}$ and $CE_{.75}$ is smaller than the distance between $x_{\max}$ and $CE_{.75}$. Therefore, $CE_{.875}$ is elicited; otherwise, we would have elicited $CE_{.625}$. We do so in order to decrease the differences in elicited utility points on the x-axis, or, put differently, to decrease the maximal difference in subsequent elicited certainty equivalents, to get decisions over a wide range of monetary amounts for every participant. Finally, a last point is elicited to decrease the then largest difference in certainty equivalents.[13]

[12] For example, using a power utility function implies that risk averse individuals are always prudent, which is not flexible enough to describe empirically observed choice behavior (e.g., Noussair et al., 2014).

[13] If, for example, up to this last step, the certainty equivalents $CE_{.125}$, $CE_{.25}$, $CE_{.5}$, $CE_{.625}$, $CE_{.75}$ have been elicited, this could be one of the following certainty equivalents: $CE_{.0625}$, $CE_{.1875}$,

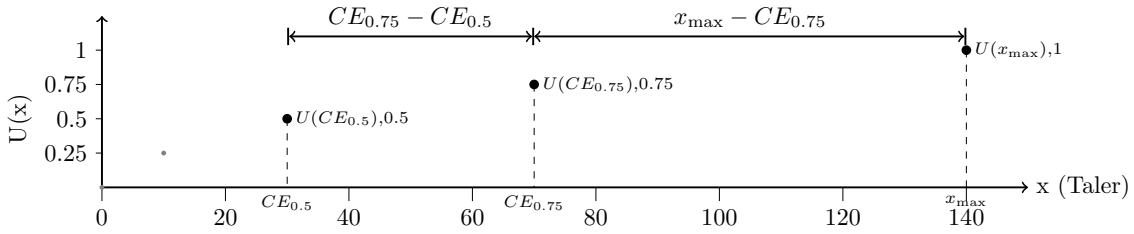


Figure 1: Adaptive Elicitation of Utility Points

Connecting utility points to a non-parametric utility function in a way that accounts for measurement error in the utility points is the statistical core of the method. This we achieve by adapting penalized spline regression (Eilers and Marx, 1996) to the case of individual utility functions such that intensity measures of higher order risk preferences can be computed. Here, we only highlight the key advantages of P-spline regression for our purpose and add more detail in Appendix A.2. The adaption of P-spline regression to our case is non-trivial and described in detail in our companion paper (Schneider et al., 2021). For an implementation of this approach, see our R package *utilityFunctionTools* (Schneider and Baldini, 2021b).

The basic idea of spline regression is to overcome over- and underfitting (that, e.g., polynomial regression is prone to) by regressing on so-called basis functions (e.g., $x, x^2, x^3, \ldots$) that are each defined only on a sub-interval of the function's support. To avoid subjectively deciding about the number of sub-intervals as well as the position of their boundaries, P-spline regression relies on an abundant number of piece-wise defined basis functions. Overfitting is then tackled with a machine-learning approach. A penalization term is introduced to balance the trade-off between smoothness and fit to the data, i.e., to ensure that just "the right amount of flexibility" is used: This amount, i.e., how much coefficients of neighboring piece-wise defined basis functions are allowed to differ, is determined by optimizing the estimated utility function's predictive quality using cross-validation (on the individual level), which is the core of most machine-learning techniques. The vast advantage of this approach over a simple linear interpolation is its ability to (holistically) address measurement error: The regression approach considers the overall shape of the utility function instead of just the change from one point to another while neglecting most of the available data, as linear interpolation does. Another crucial advantage of the spline approach is that from utility functions established in that way, derivatives can be calculated analytically with a closed form solution

$CE_{.375}$, $CE_{.5625}$, $CE_{.6875}$, $CE_{.875}$. See `http://horp-otree-apps.sebastianoschneider.com/` for an illustration of the (publicly available oTree) implementation and Schneider and Baldini (2021a) for the corresponding paper.

without the need for additional numerical computation. Based on the derivatives of the utility function, attitude measures can then be calculated as follows.

Attitude Measures As a measure of individual risk attitude, we use the Arrow-Pratt measure (Pratt, 1964), $r = -u''/u'$, where positive (negative) values indicate risk aversion (risk loving) and risk neutrality corresponds to $r = 0$.[14,15] The theoretical importance of the measure is due to Pratt (1964), who shows that r is proportional to the risk premium and establishes that the measure is suitable to compare individuals regarding their risk attitude.

We measure an individual's prudence level with the measure popularized by Crainich and Eeckhoudt (2008), $p = u'''/u'$, where positive (negative) values indicate (im)prudence.[16,17] Theoretically, our measure is proportional to the prudence utility premium, i.e., the difference of utility between a prudent and an imprudent option, after conversion into monetary terms (Crainich and Eeckhoudt, 2008). Moreover, u'''/u' is also a measure of left-skewness aversion: Modica and Scarsini (2005) show that the increase (decrease) in premium that is due to an increase (decrease) in left-skewness (right-skewness) is proportional to this measure.

As an individual measure of temperance, we use the measure due to Denuit and Eeckhoudt (2010), $t = -u^{iv}/u'$, where again positive (negative) values indicate (in)temperance. Denuit and Eeckhoudt (2010) show that this measure is proportional to the increase in premium due to an increase in fourth-order risk and thus it is a measure for temperance and dislike of kurtosis alike.

Measurement Error It is well-known that measurement error in explanatory variables biases coefficients towards zero. Less was known, however, about the extent of the consequences for applied work relying on experimental measures until recently, when Gillen et al. (2019) have replicated influential experimental work with and without applying techniques to address measurement error. They show

[14] Before building the measures described here, we predict for each derivative separately its value at 1000 evenly spaced points in the support. Note that the division by u' is merely a normalization of the utility function such as to not change the measure if utility is multiplied by a constant; it has, however, no further meaning (Pratt, 1964). Building the measures only after averaging over the different wealth levels has two advantages: First, the dependence of the measures on wealth is "averaged" out. This is desired, as we are connecting these measures to field behavior in general, and not at a specific level of wealth. Second, no further noise is introduced into the measures (resulting, e.g., from division by values virtually indistinguishable from zero).

[15] A positive Arrow-Pratt measure also indicates aversion to mean-preserving spreads (Rothschild and Stiglitz, 1970).

[16] Positive values also correspond to downside risk aversion as defined by Menezes et al. (1980).

[17] Note that our measure is different from the well-known measure introduced by Kimball (1990) in order to be able to compare risk averse and risk seeking subjects: For a prudent individual, the Kimball measure $-u'''/u''$ might be positive or negative depending on her risk attitudes.

that once measurement error is properly accounted for, coefficients of the (experimental) controls increase substantially. In a case study, the so improved controls let previously "found" relations between other variables disappear. The proposed solution to address measurement error is a "mix of statistical tools and design recommendations" (Gillen et al., 2019, p. 1827): They suggest to combine repeated measurement of a characteristic (possibly resulting from another method) with an instrumental variable approach by instrumenting the repeated measures with each other, an approach they call *Obviously Related Instrumental Variables (ORIV)*.

Our method conceptionally reflects the core characteristics of the approach by Gillen et al. (2019): First, our measures rely on repeated measurement (while two certainty equivalents would suffice to determine the sign of the third derivative, i.e., to determine prudence, we rely on six certainty equivalents). Second, we use statistical tools to account for measurement errors, using the common variation in the repeated measures. However, as we rely on (non-parametric) utility functions and the mild assumption of monotonicity, we can go one step further and take a holistic approach, considering the *whole set of information* instead of only taking into account pairs of measurement, as suggested by ORIV: This is the virtue of a (spline) regression approach, where *all* measurements are taken into account to predict a utility function.

These points not withstanding, we show the robustness of the relations between higher order risk preferences and the domains of field behavior under study by additionally using the ORIV estimation proposed by Gillen et al. (2019). Based on certainty equivalents for a two-outcome lottery with equally likely outcomes, a continuous measure of *the intensity* of risk aversion can be defined (see, e.g., Sutter et al., 2013): $r^{\text{CE}} = 0.5 - (\text{CE} - x_{\text{low}})/(x_{\text{high}} - x_{\text{low}})$, where CE denotes the certainty equivalent, and x_{high} and x_{low} the high and low outcomes of the corresponding lottery, respectively. Positive values imply risk aversion, and negative values imply risk-seeking behavior. For the first certainty equivalent that we elicit, we can compute this measure and compare it meaningfully accross all individuals, as $x_{\text{high}} = 140$ and $x_{\text{low}} = 0$ for all participants in the first lottery. We can thus use this simple and widely used measure as additional (duplicate) measure of risk aversion to the measure of risk aversion resulting from our method to apply the ORIV method, providing further evidence 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*.

Identification and Validation From a theoretical point of view, as pointed out above, identification of (higher order) risk preferences with our method is evident,

as we derive our measures from the derivatives of the utility function, in the way higher order risk preferences were initially *defined* in the expected utility framework (Leland, 1968; Kimball, 1990).[18] Clearly, as we use two-outcome lotteries only, identification of, e.g., prudence, comes *with the change of decisions as the amounts of the lotteries change*; it cannot be derived from an isolated decision in these tasks.

We validate our method empirically to show that the theoretically apparent identification of (higher order) risk preferences is not masked by the approximations inherent in our method (such as predicting the utility function with a spline regression approach). We thus have *additionally* elicited prudence using a version of the widely-used lottery task introduced by Eeckhoudt and Schlesinger (2006). To account for our sample, we have adapted the tasks from Heinrich and Shachat (2020) – where compound lotteries are reduced following Maier and Rüger (2011).[19] On average, we find that there is a significantly positive relation ($p < 0.001$) between the number of prudent choices à la Eeckhoudt and Schlesinger (2006) and our intensity measure of prudence; see Table 2 and Table 5 for the corresponding results. Importantly, the measure of risk aversion resulting from our method as used in the regressions in the main text *is completely unrelated* to the number of prudent choices (see Table 2). Yet, our measure of risk aversion is positively and significantly related ($p < 0.001$) to r^{CE}, the simple continuous measure of risk aversion as defined above (and as used, e.g., in Sutter et al., 2013; see Table A-2 in Appendix A for this result) and to the standard survey question on risk tolerance (e.g., Falk et al., 2018; see Table 5 for this result).

3 Results on Higher Order Risk Preferences and Their Determinants

3.1 Results

In our pre-analysis plan, we have specified directional hypotheses for the relation between the risk preferences risk aversion, prudence and temperance and the influence factors age, cognitive ability and gender. For impatience and all other possible influence factors investigated (see the table notes for the full list, and the tables in Appendix B for full regression results), we correct p-values for multiple hypothesis

[18] See Eeckhoudt and Schlesinger (2006) for a proof of the equivalence of deriving (classifications of) higher order risk preferences from the derivatives of the utility function and binary choice tasks. From this proof, the now well-known elicitation method for classification of higher order risk preferences has emerged.

[19] See Figure S-1 in the supplementary material for an illustration of such a task. Instead of the amounts 30, 70, 100, 100 (Option A) and 50, 50, 80, 120 (Option B) used in the first task, 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.

Table 2: Validation of Elicitation Method: Identification of Prudence

	Prudence		Risk Aversion	
	Intensity$^\sharp$	Classification†	Intensity$^\sharp$	Classification†
# Prudent choices (Heinrich & Shachat)	0.295**** (0.017)	0.248**** (0.034)	−0.000 (0.017)	0.065* (0.033)

Notes: Positive coefficients imply increased intensity measures of prudence and risk aversion as resulting from our method or an increased likelihood of being classified as prudent or risk averse, respectively. Regression coefficients resulting 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. $^\sharp$For prudence, we use the intensity measure p by Crainich and Eeckhoudt (2008). For risk, we use the Arrow-Pratt measure r as used in most of the empirical part of the paper, i.e., after having removed the variation explained by prudence. See *Attitude Measures* above for definitions. †We classify subjects as prudent and risk averse if their respective intensitiy measure is positive, i.e., $p > 0$ and $r > 0$, respectively.
****/***/**/* denotes significance at the 0.1 / 1 / 5 / 10 percent level.

testing using the Romano-Wolf procedure with 1,000 iterations (Romano and Wolf, 2005a,b, 2016).

Risk Aversion We find significant risk aversion in our sample. We estimate a mean (median) Arrow-Pratt coefficient of risk aversion, expressed in standard deviations, of $r = .46$ (.35), with 0 indicating risk neutrality (p-value < 0.0001, Wilcoxon signed-rank test,[20] testing whether r is different from 0). For 71% of our sample, we estimate a positive Arrow-Pratt coefficient, implying risk aversion (see Figure B-1 in Appendix B for the full distribution). A regression including a measure for cognitive abilities and demographic background variables is shown in column [1] of Table 3. The regression shows a gender and a cognitive ability effect, as expected: Females are significantly more risk averse than males. Individuals with higher cognitive abilities are significantly less risk averse. Age is unrelated with risk aversion, once we control for cognitive abilities, which is in line with our hypotheses. This suggests that risk aversion is rather affected by the increase in cognitive abilities due to an increase in age than just by growing older and becoming more experienced. One additional influence factor, impatience, is significantly related with a lower degree of risk aversion. All other independent variables mentioned in the legend to Table 3 are not significant.

Prudence On the aggregate level, we find prudence in our sample. The mean (median) estimate of the Crainich-Eeckhoudt measure expressed in standard deviations is $p = .56$ (.22), where positive (negative) values indicate (im)prudence (p-value < 0.0001, Wilcoxon signed-rank test, testing whether the prudence measure p is different from 0). For 68% of our sample, we estimate a positive Crainich-

[20]For ease of exposition, all tests reported in this paper are two-sided.

Table 3: Influence Factors of (Higher Order) Risk Preferences

	[1] Risk aversion		[2] Prudence		[3] Temperance	
Age (in years)	−0.014	(0.015)	−0.007	(0.019)	0.002	(0.015)
Cognitive ability	−0.132***	(0.043)	−0.062	(0.058)	−0.118**	(0.042)
Female (=1)	0.259**	(0.091)	0.205*	(0.105)	0.166*	(0.085)
Impatience	−0.889***	(0.275)	−0.584**	(0.234)	−0.703**	(0.283)
Other Factors	10		10		10	
R^2	0.08		0.06		0.06	
Observations	658		658		658	

Notes: OLS regressions of higher order risk preferences. Positive coefficients imply increasing risk aversion, prudence and temperance, 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, the religion, migration background, parents' education as well as their occupation; see Tables B-1, B-2 and B-3 for detailed regressions results and Tables S-1, S-2, S-3, and S-4 (supplementary material) for regression results excluding participants that reported problems with handling their tablets. For 24 participants, some demographic information has been imputed with 0, the variable's mean value, and the 'other' category for binary, continuous and categorical variables, respectively. We controlled for imputation with indicator variables. Robust standard errors clustered at the session level in parentheses. P-values for factors omitted in this table and for impatience are corrected for multiple hypothesis testing using the Romano-Wolf procedure with 1,000 iterations (Romano and Wolf, 2005a,b, 2016).
****/***/**/* denotes significance at the 0.1 / 1 / 5 / 10 percent level.

Eeckhoudt measure, implying prudence (see Figure B-1 in Appendix B for the full distribution). For the relations with prudence that we are most interested in (age, cognitive abilities, and gender), the regression shown in column [2] of Table 3 only reveals a gender effect: Females are more prudent than males, as expected. However, neither cognitive abilities nor age are significantly related with prudence. All other independent variables (except for impatience) are not significant once p-values are corrected for multiple hypothesis testing.

Temperance In the aggregate, our sample exhibits temperance. For the Denuit-Eeckhoudt measure of temperance, our mean (median) estimate is $t = .30$ (.02), again expressed in standard deviations (p-value < 0.0001, Wilcoxon signed-rank test, testing whether t is different from 0). Given a positive Denuit-Eeckhoudt measure, 58% of our sample can be classified as temperant (see Figure B-1 in Appendix B for the full distribution). The regression of temperance on demographic background variables reported in column [3] of Table 3 shows the same pattern as the regression of risk aversion: Females are more temperant than males (as expected), and students with higher cognitive abilities and larger impatience are less temperant. There is neither an age effect, nor is any other influence factor significant once we correct p-values for multiple hypothesis testing.

Relation between risk aversion, prudence and temperance The measures of risk aversion, prudence and temperance are significantly correlated in our sample (p-value < 0.0001 for all pairwise correlations). The correlation between risk aversion and temperance is the highest ($\rho = .87$). The correlations between prudence and risk aversion ($\rho = .56$) and prudence and temperance ($\rho = .65$) are still large, but a magnitude weaker.[21] Also our measure of impatience is correlated with the (higher order) risk measures (p-value < 0.05 for all pairwise correlations), where the correlation between risk aversion and impatience is the highest ($\rho = -.13$).

3.2 Discussion

In the aggregate, we find significantly risk averse, prudent, and temperant behavior in our sample of children and adolescents. This is in line with earlier studies on risk aversion or prudence with adolescents (e.g., Sutter et al., 2019; Heinrich and Shachat, 2020).[22] Among adults, prudence is wide-spread and has been documented in a number of studies (see, e.g., the review by Trautmann and van de Kuilen, 2018). Our finding of 68% of subjects exhibiting prudent behavior is comparable to results by Tarazona-Gomez (2004) and Deck and Schlesinger (2010), for example, who report 63% of prudent subjects and 61% prudent decisions, respectively. For temperance, however, no study has investigated the prevalence among adolescents before, and results among adults have been mixed. Most studies document temperance, although less prevalent than prudence (Trautmann and van de Kuilen, 2018), which is also what we observe. The correlations between (higher order) risk preferences that we find are higher than the ones reported by Noussair et al. (2014) and Ebert and Wiesen (2014). However, our measures are continuous and account for measurement error, implying a higher precision (Gillen et al., 2019).

With respect to demographic correlates, we find a gender effect for all our risk preferences, but no age effect, neither for risk aversion (once we control for cognitive abilities), nor for prudence or temperance. While also previous studies among adolescents report a gender effect, but no age effect on risk aversion (Sutter et al., 2019), the finding with respect to age and prudence as well as prudence and gender differ from the only other study on prudence with adolescents by Heinrich and Shachat (2020). Yet, they proxy age with grade, have no *absolute* measure of cognitive abilities to control for, and they investigate classifications instead of intensitites

[21] In Appendix A.3, we show the correlations separately for risk averters and risk-loving subjects and explain why this is evidence for the existence of both mixed risk averters (Caballé and Pomansky, 1996) and mixed risk lovers (Crainich et al., 2013).

[22] For example, Sutter et al. (2013) report a mean (median) measure of risk aversion (defined as $r^{\mathrm{CE}} + .5$) of .57 (.53). We obtain values of .58 (.56) for $r^{\mathrm{CE}} + .5$, with Pearson and rank correlation coefficients of .87 between this measure and the Arrow-Pratt measure resulting from our method used throughout this paper.

with their sample of 362 Chinese students aged 8 to 17 years. Moreover, as Sutter et al. (2019) note in their review, notable changes in risk preferences might occur before the age of 10. As Heinrich and Shachat (2020) include one grade with children below that age, they might observe age effects due to inclusion of this grade. Among adults, Noussair et al. (2014) find no age effect either. Regarding the observation that females exhibit more risk averse, prudent and temperant behavior, our findings are in line with Ebert and Wiesen (2014). Similarly, Noussair et al. (2014) document females exhibiting more risk aversion and temperance, but not prudence.

Previous literature among adults has documented a positive relation between prudence and cognitive abilities (Noussair et al., 2014; Breaban et al., 2016), but no relation between temperance and cognitive abilities (Noussair et al., 2014). Among adolescents, only the relationship between risk aversion and cognitive abilities has been investigated, and either no correlation has been observed, or a tendency towards risk-neutrality with increasing cognitive abilities (Sutter et al., 2019). In our sample, we also observe a tendency towards risk-neutrality with increasing cognitive abilities. Yet, and in contrast to previous results among adults, prudence and cognitive abilities are unrelated in our sample, while for temperance we observe the same pattern as for risk aversion, namely a tendency towards temperance-neutrality.[23] This differing finding might be due to sample characteristics, the fact that we focus on intensity measures instead of classifications, the different measure of cognitive abilities that we use, or the reduced complexity in our elicitation method. Given these results, more research in the future will be useful to settle the relationship between higher order risk preferences and cognitive abilities.

4 Eliciting Different Types of Field Behavior

After having presented our results on risk aversion, prudence and temperance and the determinants of these (higher order) risk preferences, we turn to the second main contribution of our paper, namely their relation to adolescents' field behavior. Because the introduction only briefly touched upon the different types of field behavior, we start in this section with a more detailed description of the various types and present in particular also the theoretical predictions about the relationship between higher order risk preferences and field behavior. The predictions are summarized in Table 4, which also indicates whether previous empirical evidence is available for a specific domain.

[23] In this context, note that Noussair et al. (2014) also fail in finding a significant relation between risk aversion and cognitive ability in their (student) sample (N=109), although a negative relationship is well documented in larger studies and when cognition measures target numeracy (see the review on experimental measures and cognitive abilities in Dohmen et al., 2018).

4.1 General Risk Taking and Patience Behavior

Consistency of risk elicitation methods across tasks and survey questions has gained considerable attention, partly due to recurrently less encouraging results (see, e.g., Crosetto and Filippin, 2015, Pedroni et al., 2017 and the discussion in Schildberg-Hörisch, 2018). Recently, such results have been attributed to measurement error (Gillen et al., 2019). As our novel elicitation method for risk preferences addresses measurement error by design, we are first interested in whether it improves upon alternative methods in predicting standard survey questions. Specifically, we use the standard question on willingness to take risk in general, first included in the German socio-economic panel (SOEP; Wagner et al., 2007; Dohmen et al., 2011; Falk et al., 2018). Then we use an adapted subsample of the DOSPERT questionnaire (Weber et al., 2002) that was built to assess risk in the domains of health/safety as well as recreational, ethical, social and financial decisions.[24] Questions used to address general risk taking behavior are printed in Appendix C.1.

To compare our measure of impatience against validated survey questions, we also added the standard time survey question from the SOEP and the global preferences survey (e.g., Wagner et al., 2007; Vischer et al., 2013; Falk et al., 2018), see Appendix C.1.

4.2 Health-Related Behavior

Prudence has been linked theoretically to the health domain, for example by studying multivariate risk taking (e.g., Eeckhoudt et al., 2007), prevention effort (e.g., Courbage and Rey, 2006), the demand for medical care (e.g., Dardanoni and Wagstaff, 1990), or medical treatment decisions (e.g., Bleichrodt et al., 2003; Krieger and Mayrhofer, 2012). Yet, we are not aware of any empirical study connecting higher order risk preferences with behavior that may risk one's health status, such as smoking, drinking or physical inactivity. This is surprising given the interpretation of prudence as downside risk aversion (Menezes et al., 1980) and aversion to left-skewness (Modica and Scarsini, 2005; Ebert, 2012). For example, as smoking addiction increases the probability of cardiovascular diseases, it may be seen as a typical example of a downside risk, or a situation where the distribution of risk is left-skewed: There may be a relatively small positive outcome with a high prob-

[24] We use a subset of these questions to account for our underage sample: some questions (e.g., having an affair with a married person or cheating on one's tax return) would induce low variation and seemed inappropriate to ask to adolescents as young as of age 10. Moreover, we adapted some questions (e.g., using a helmet when riding a bike instead of riding a motor bike) and added some more that might be relevant to our sample (e.g., having a date with someone that they have met via the internet/social media/apps).

ability resulting from enjoyment of smoking, which, however, is combined with a low-probability, but high-impact, negative outcome due to a cardiovascular disease.

To test the importance of higher order risk preferences for behavior putting one's health status at risk, we include several questions in our questionnaire to capture this behavior. Notably, we include a novel, self-constructed scale consisting of 6 questions to capture smartphone and social media addiction, as this kind of addictive behavior has gained tremendous attention over the last decade (e.g., Hormes et al., 2014; Andreassen, 2015; Andreassen et al., 2017; He et al., 2017). Yet, to our knowledge, it has been ignored in the risk taking literature so far. Our scale comprises questions such as "When I feel bad or when I face a difficult task, I distract myself with my smartphone.", "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.", or "I often check my phone while eating with my family to see if there are any news." In addition to abusive smartphone usage, our addictive behavior index comprises smoking and drinking behavior, which is assessed by the respective frequency. Behavior that generally is a risk to health additionally includes items on, e.g., the BMI or physical inactivity. See Section C.2 in Appendix C for all items addressing health-related behavior.

4.3 General Prevention and Eco-Friendly Behavior

Prevention in the sense of self-protection is understood as effort that lowers the probability of the occurrence of an adverse event (Ehrlich and Becker, 1972). It has been theoretically connected to prudence (Eeckhoudt and Gollier, 2005; Menegatti, 2009). In one-period models, the preventative effort and the potential loss are contemporaneous. Contrary to intuition, in this setting, prudence has a negative impact on the optimal level of prevention, since the prudent agent prefers to accumulate wealth to face future risks instead of investing in preventative effort (Eeckhoudt and Gollier, 2005). In two-period models, the preventative effort precedes the potential loss. In that setting, the relation between prudence and the optimal level of prevention is positive (Menegatti, 2009).[25] In the abstract setting of a laboratory experiment, Krieger and Mayrhofer (2016) find support for the predictions of the one-period models in the literature: Prudent subjects invest less money than non-prudent subjects to reduce the probability of a loss. We are unaware of any study empirically investigating prevention in a two-period framework and a field-setting.

[25]Strictly speaking, those are the predictions for a loss that occurs with the same probability as it does not occur. For other probabilities, the predictions are more nuanced, but perfectly opposed in the two models.

For the purpose of studying real-world preventative and pro-environmental efforts and their relation with prudence, we include several questions in our questionnaire. We distinguish between actions preventing an unwanted event that might arguably happen on the same day (one-period model) and those that aim to prevent events happening in the more distant future (two-period model). For example, agreement to the statement "Since 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." indicates effort provision in order to prevent hunger, an event that is likely to happen on the same day. Contrarily, agreement to the statement "Since 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." indicates effort provision in order to prevent a missed chance to become, e.g., a medical doctor or a lawyer; an event that will happen only with considerable time delay in the future. Additionally, we add questions on eco-friendly behavior. For example, we ask whether students separate their waste, use reusable coffee cups, bottles, or shopping bags, take their bike when possible (instead of a bus or car), or turn down the heating if leaving a room. See Section C.3 in Appendix C for the questionnaire on prevention effort and eco-friendly behavior.

4.4 Planning Behavior

In his seminal paper in which he developed the "theory of the optimal response of decision variables to risk (which includes precautionary saving as a subcase)", Kimball (1990, p. 54) *defined* prudence as the "sensitivity of the optimal choice of a decision variable to risk", thus the degree to which plans are adapted to risk.

We test if we find support for prudence in this more general sense as a measure of cautious planning. The decision situations considered in our questions are, e.g., preparation for a class test, or being on time for a meeting. We ask participants how much more time they invest in the given situations, if risk increases by, e.g., a more uncertain scope of the class test, or by red lights on the way to the meeting. See Section C.4 in Appendix C for the exact wording of the items used.

4.5 Preference for Competitive Income

In the context of the German reunification "experiment", Fuchs-Schündeln and Schündeln (2005) discuss self-selection of risk-averse individuals into low-risk occupations and its importance for precautionary savings. They compare the difference in precautionary savings between civil servants and the remaining population in the East of Germany with this difference in the West of Germany. From a larger difference in the East of Germany, where all occupations were basically risk free before reunification, they infer that risk-averse individuals self-select into jobs as civil servants in the West of Germany. They explicitly mention that their argument

builds on their (untested) assumption that – at least on average – risk aversion equals prudence.

As we have separate, direct measures of higher order risk preferences, we ask whether our participants would have preferred a fixed or a tournament payment instead of the piece-rate payment for the cognitive ability tasks; in addition, we ask two questions on the choice of risky/riskless occupations (e.g., "Would you like to become a civil servant, i.e., work for the state as a teacher, a policeman, in a city's administration or at the tax office, etc?"). See Section C.5 in Appendix C for the exact wording of the items used.

4.6 Financial Decision Making

A positive third derivative of the utility function was already linked by Leland (1968) to financial decision making, in particular precautionary saving, long before Kimball (1990) coined the term prudence and introduced the now well-known measure for its strength. Temperance is theoretically related with less risky investments as reaction to greater background risk (Kimball, 1992; Gollier and Pratt, 1996). The demand for insurance and its positive relation to risk aversion is usually already discussed in basic economics textbooks, and it has been connected to prudence and temperance in the presence of background risk (e.g., Eeckhoudt and Kimball, 1992). Theory remains inconclusive about the impact of prudence and temperance on insurance demand (e.g., Fei and Schlesinger, 2008), but empirical work suggests that the standard model of risk aversion might not adequately describe insurance demand (Sydnor, 2010). Using the elicitation method by Eeckhoudt and Schlesinger (2006), Noussair et al. (2014) were the first to explore the relation between higher order risk preferences and financial decision making with experimental risk measures. Based on a large Dutch sample, they find support for the relation between saving and prudence, and less risky investment and temperance. They report that there is no robust connection between insurance coverage and their risk measures.

Naturally, adolescents take less financial decisions than adults. However, their financial education has long started, and they all receive regular pocket money, cash gifts, or earn money in a side job that they have to manage themselves. To study decision making adjusted to their situations, and to investigate whether our sample follows the same general patterns as found for adults in Noussair et al. (2014), we include questions on saving, risky investments and insurance coverage: To collect field behavior on saving, we ask students for example what fraction of a gifted 50 Euro bill they would save, or how they handle their pocket money, where possible answers range from "I spend everything quickly" to "I save everything". The equivalent to adults' preference for risky investment is assessed with questions

like "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"? To address insurance demand, we ask for the possession of a bike or phone insurance, and whether students bought it themselves. See Section C.6 in Appendix C for the full list of items included in the questionnaire to target financial behavior.

4.7 Building Indices of Field Behavior

For most of the questionnaire, we obtained data for all 658 students.[26] We build indices for the different domains of behavior, involving between three (planning behavior) and 25 questions (adapted DOSPERT catalogue). Importantly, all indices contain all information that we have gathered for an individual in the respective domain. To aggregate results for a varying number of questions across age cohorts with possibly also differing meaning and variation, we first build indices per age cohort using weights from a principal component analysis (PCA). For example, whether or not adolescents turn off the TV if no one is watching becomes of higher importance for determining eco-friendly behavior of the older age cohorts, as the share of adolescents having an own TV increases; consequently, the weight associated with this question increases from about .05 for the sixth graders to .1 for the twelfth graders.[27] After building age cohort indices, we center and standardize them for final aggregation over the cohorts, in accordance with our pre-registered analysis plan.[28] Therefore, all of our dependent variables are centered, standardized and continuous and we thus use least squares regression.

5 Higher Order Risk Preferences and Field Behavior

Now we turn to the relationship between higher order risk preferences and field behavior. For comparison reasons, we standardize our measures of risk and time preferences. Hence, coefficients report the effect of a one standard deviation increase in these measures. Recall from Section 3 that our measures of risk preferences are considerably correlated. To ease interpretation of coefficients and to avoid multicollinearity, we orthogonalize these measures.[29] For example, we remove vari-

[26] Some questions, however, were not asked to the youngest students (e.g., drinking or dating behavior), in accordance with our agreements with participating schools and to get meaningful results (see the respective column for a question in Appendix C).

[27] Similarly, questions on risky investment and insurance demand differ in their meaning, which we additionally address by likewise analyzing our full sample and our older participants in isolation.

[28] One index, the preference for competitive income index, is not aggregated using PCA. As it consists of only 4 binary questions, the support of the index consists of only 16 elements. Here, using weights from a PCA per age cohort would shift the support for every age cohort marginally, thus introducing noise in the measure when aggregating the indices rather than precision. Therefore, we compute z-Scores for every item and add them.

[29] Ridge regression (e.g., Hoerl and Kennard, 1970) is another way of handling multicollinearity. Results (shown in Table S-6 in the supplementary material) are very similar, but as Ridge

Table 4: Summary of Theoretical and Empirical Literature on Higher Order Risk Preferences (HORP) and Field Behavior

Domain	**Outcome**	**Related HORP**	**Theoretical Argument or Model (Selection)**	**Theoretical Relation[a]**	**Prior Empirical Evidence (Field-Behavior)**	**Empirical Relation (Prior Work)[a]**
Health	Unhealthy Behavior (in General) Addictive Behavior Abusive Smartphone Usage	Prudence	Menezes et al. (1980); Modica and Scarsini (2005); Ebert (2012)[b]	−	N/A[c]	
Prevention and Environment	Short-Term Prevention (General)	Prudence	Eeckhoudt and Gollier (2005)[d]	−	N/A[e]	
	Long-Term Prevention (General)	Prudence	Menegatti (2009)[d]	+	N/A	
	Eco-friendly Behavior	Prudence	Bramoullé and Treich (2009)	−	N/A	
Planning	Cautious Planning	Prudence	Kimball (1990)	+	N/A	
Competition	Preference for Competitive Income	Prudence	Fuchs-Schündeln and Schündeln (2005)[f]	−	N/A	
Financial Decision Making	Saving	Prudence	Leland (1968); Kimball (1990)	+	Noussair et al. (2014)	+
	Risky Investment	Temperance	Kimball (1992); Gollier and Pratt (1996)	−		−
	Insurance Demand	Prudence	Fei and Schlesinger (2008)[g]	+ & −		?
		Temperance	Eeckhoudt and Kimball (1992)	+		?

[a] A plus sign (+) indicates a positive relationship, whereas 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] These works provide a theoretical link that is rather general and stems from the interpretation of prudence as downside risk aversion (Menezes et al., 1980) or aversion to left-skewed risks (Modica and Scarsini, 2005; Ebert, 2012), which is particularly relevant for health risks. Most health risks have a left-skewed distribution (unlikely, but extremely negative events combined with likely, moderately positive events).

[c] 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.

[d] These are the predictions in case the loss to be prevented by investing preventative effort happens with a chance of 50%. Both papers also cover more nuanced cases.

[e] 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.

[f] Fuchs-Schündeln and Schündeln (2005) document lower precautionary wealth in 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.

[g] The prediction from this model on uninsurable background risk and insurance demand for another, insurable loss differs depending on whether the loss to be insured occurs in the state of the world where there is no background risk (−) or whether it occurs in the state of the world where there is background risk (+).

ation from the temperance measure that is already explained by prudence and risk aversion. We do so by predicting temperance on the individual level, using each individual's risk aversion and prudence measure. Then we subtract this prediction from the actual measure to obtain the variation not yet explained by prudence and risk aversion. Before, the same is done iteratively for prudence and risk aversion. This way, either the risk aversion, the prudence, or the temperance measure remains as is, and we pick the measure that matters according to theory for the respective index (see the predictions summarized in Table 4).[30] In the regressions, we then include the untouched measure and the so residualized ones. Results from regressions on correlated measures are printed in the Appendix, see Table B-16. Note that due to multicollinearity both coefficients and standard errors are imprecise in that table. Results are, however, very similar to those presented in the main text.

Although our elicitation method inherently addresses measurement error by combining repeated measurement with a spline regression approach, we show the robustness of our results to additionally accounting for measurement error with the ORIV technique as suggested by Gillen et al. (2019); see Section 2.3 for details and Table B-19 in Appendix A and Table S-7 in the supplementary material for results.

Before turning to the relationship between our risk preferences and field behavior, we would also like to note that our experimental measures of (higher order) risk preferences are estimates (generated regressors) and thus involve some degree of error. We therefore use bootstrapped standard errors.[31]

5.1 General Risk Taking and Patience Behavior

As a kind of illustration of the effectiveness of our method in addressing measurement error, and in addition to the validation of our method (see Section 2.3), we start by analyzing how our measures relate to standard survey questions about risk attitudes and patience as well as the "count measure" of prudence building on the method by Eeckhoudt and Schlesinger (2006), as used for adolescents by Heinrich and Shachat (2020). Results are summarized in Table 5. In our sample of adolescents, our experimentally elicited intensity measure of prudence is significantly related ($p<0.001$) with the "count measure" based on the method by Eeckhoudt and Schlesinger (2006), while our measure of risk aversion is not. This result is robust to using the ORIV technique, where the coefficient of our prudence measure

regression yields biased coefficients (Gruber, 2017) and for ease of interpretation of coefficients, we present results from orthogonalized measures in the main text.

[30] If theory does not suggest one measure as the most important one, we take prudence, as it has the lowest correlation with the other two measures on average.

[31] Another way to deal with this is multiple imputation (e.g., Rubin, 1996). Yet, as our results are almost identical when using multiple imputation (see Table S-5 in the supplementary material), we report results from OLS regressions with bootstrapped standard errors in the main text.

Table 5: Survey Questions/Questionnaires on General Risk Taking and Patience

	Risk tolerance (Survey, SOEP)		DOSPERT (Adapted)		Patience (Survey, SOEP)		Prudence (à la Eeckhoudt & Schlesinger, 2006)	
OLS regression results								
Risk aversion (AP)	−1.063****	(0.241)	−0.288****	(0.062)	0.016	(0.187)	0.004	(0.070)
Prudence	−0.477****	(0.075)	−0.087*	(0.045)	−0.052	(0.067)	0.105****	(0.031)
Temperance	−0.224**	(0.107)	−0.023	(0.039)	0.067	(0.107)	−0.065	(0.048)
Impatience	0.020	(0.099)	0.107***	(0.033)	−0.474****	(0.086)	0.065*	(0.039)
Other Factors	13		13		13		13	
R^2	0.13		0.18		0.12		0.05	
Observations	653		658		653		658	

	Risk tolerance (Survey, SOEP)	DOSPERT (Adapted)	Patience (Survey, SOEP)	Prudence (à la Eeckhoudt & Schlesinger, 2006)
Raw Pearson correlation coefficients				
Risk aversion (AP)	−0.270****	−0.150****	0.022	0.008
Prudence	−0.202****	−0.099**	−0.017	0.094**
Temperance	−0.192****	−0.092**	0.015	0.012
Impatience	0.075*	0.184****	−0.240****	0.045
Risk tolerance (Survey)		0.270****		

Notes: Positive coefficients imply increasing risk tolerance (as measured by the standard survey question on willingness to take risk in general, see Dohmen et al., 2011), increasing general risk taking behavior, increasing patience (where 'survey' refers to the standard survey question on patience, see Vischer et al., 2013) and a higher number of prudent decisions in choice tasks à la Eeckhoudt and Schlesinger (2006) in the upper panel ("OLS regression results") and positive correlations in the lower panel ("Raw Pearson correlation coefficients"). Experimental risk and time measures in standard deviations in the upper panel. Outcome indices are formed using PCA weights and are standard normalized (single item survey questions and the "count measure" are used without transformation). Questions included in these indices are listed in Section C.1 of Online Appendix C. See Table B-4 for detailed regression results. Bootstrapped standard errors (1000 rep.) clustered at the session level in parentheses.
****/***/**/* denotes significance at the 0.1 / 1 / 5 / 10 percent level.

even increases (see Table S-7a in the supplementary material). We have added this regression result to Table 5 for convenience, although predicting our continuous intensity measure with the more crude "count measure" arguably makes more sense: On average, choosing the prudent option one more time in the decision tasks à la Eeckhoudt and Schlesinger (2006) increases our intensity measure of prudence significantly ($p<0.001$) by about .3 SD; see Table 2 for this (first column) and further validation results, as well as Table A-2 in Appendix A.

Turning to the relation with survey measures, we note that our experimentally elicited Arrow-Pratt measure of risk aversion significantly predicts the willingness to take risk as elicited via the survey question ($p<0.001$). Notably, both our experimental measures of prudence and temperance predict willingness to take risk significantly; the former even at the same significance level as risk aversion (see first column of Table 5). Our measures of prudence and risk aversion also significantly predict risk taking behavior as measured by our adapted DOSPERT scale (in the second column of Table 5). Interestingly, while the coefficient of risk aversion indicates that a one standard deviation lower risk aversion is associated with a one degree higher willingness to take risk on a scale from 0 to 10, the coefficient of prudence is still about half as large and the coefficient of temperance about a quarter as large.[32] This provides further evidence that general risk taking behavior, as we understand it in everyday language, might only be insufficiently captured by risk aversion alone (Jaspersen et al., 2020). The results regarding patience show the hypothesized relations between a single item survey question and our experimental measure of impatience ($p < 0.001$).

Via the use of Pearson correlation coefficients and the meta study METARET (Crosetto, 2021), we can compare our method with alternative experimental measures of risk aversion and their correlation with the SOEP survey question. Our results with respect to this question are shown in the column "Risk tolerance (Survey, SOEP)" (partial correlations in the upper panel, and pair-wise Pearson correlation coefficients in the lower panel of Table 5). In METARET, Pearson coefficients of experimental measures with the SOEP question range, on average, from .12 ($N = 3{,}463$) for the Bomb Risk Elicitation Task (BRET) to $-.04$ ($N = 983$) for the Certainty Equivalent price list, when using raw choices. Interestingly, the probably most widely used Holt and Laury method performs worse in this aspect than the BRET, with a correlation of .1 ($N = 7{,}552$). When estimating Arrow-Pratt risk aversion coefficients, no correlation coefficient exceeds .1. Our method has a close

[32]When using correlated preference measures, the (potentially biased) coefficient of prudence decreases from about 50% to about 35% of the (potentially biased) coefficient of risk aversion; the conclusion drawn from this comparison, however, remains the same.

Table 6: Health-Related Behavior

	Unhealthy Behavior		Addictive Behavior		Smartphone Addiction	
Risk aversion (AP)	0.024	(0.086)	0.007	(0.087)	−0.023	(0.086)
Prudence	−0.140****	(0.025)	−0.146****	(0.025)	−0.160****	(0.025)
Temperance	−0.011	(0.047)	−0.002	(0.047)	0.017	(0.045)
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	

Notes: Positive coefficients imply increasing engagement in unhealthy or addictive behavior. Risk and time measures in standard deviations. Non-orthogonalized measure: Prudence. Outcome indices are standard normalized and formed by using PCA weights. Questions included in these indices are listed in Section C.2 of Online Appendix C. See Tables B-5, B-6 and B-7 for detailed regression results. Bootstrapped SEs (1000 rep.) clustered at the session level in parentheses. ****/***/**/* denotes significance at the 0.1 / 1 / 5 / 10 percent level.

to thrice as large coefficient – namely .27 – as the pooled Pearson correlation of the best method surveyed by the METARET study, suggesting that it is very effective in addressing measurement error. In summary, our experimental measures reflect survey responses on general risk taking very well.

5.2 Health-Related Behavior

Results with respect to health-related behavior are summarized in Table 6.[33] Patience is positively correlated with healthy behavior, and so is prudence (in line with our expectations, see Table 4). These findings are robust to controlling for age and gender, amongst others (see Tables B-5 to B-7 for details), and apply to all three measures in Table 6. Notably, neither risk aversion nor temperance are predictive for any of the behavior investigated in Table 6. Moreover, these findings are robust to applying the ORIV technique (Gillen et al., 2019) to address measurement error (see Table S-7b in the supplementary material), suggesting that this observation is unlikely to be explained by measurement error. Taken together, this means that having a measure of prudence may be important for identifying subjects who put their health at risk. This seems particularly relevant with respect to smartphone usage: Excessive usage has not only been associated with depression, but it is also negatively correlated with scholarly achievement, as we can show in our sample – at an absolute magnitude comparable to that of cognitive abilities![34]

[33]Unfortunately, in the first school, we had to work with printed questionnaires which did not include most of our health questions. For this reason, the number of observations for which we obtained information on health-related behavior is slightly lower than in the other domains.

[34]For example, a one SD higher score on our smartphone addiction scale is associated with a .13 SD ($p < 0.001$) and .16 SD ($p < 0.001$) worse German and Math grade, respectively, while a one SD increase in cognitive abilities is associated with a .12 SD and .19 SD improved German and Math grade, respectively ($p < 0.001$ and $p < 0.002$).

5.3 General Prevention and Eco-Friendly Behavior

Table 7: Prevention and Eco-Friendly Behavior

	Prevention (Short Term)		Prevention (Long Term)		Eco-friendly Behavior	
Risk aversion (AP)	0.120*	(0.063)	0.167***	(0.060)	0.201**	(0.081)
Prudence	−0.119****	(0.033)	0.010	(0.042)	0.022	(0.027)
Temperance	0.017	(0.039)	0.078**	(0.034)	0.034	(0.057)
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	

Notes: Positive coefficients imply increasing prevention effort or increasing eco-friendly behavior. Risk and time measures in standard deviations. Non-orthogonalized measure: Prudence. Outcome indices are are standard normalized and formed by using PCA weights. Questions included in these indices are listed in Section C.3 of Online Appendix C. See Tables B-8, B-9 and B-10 for detailed regressions results. Bootstrapped SEs (1000 rep.) clustered at the session level in parentheses. ****/***/**/* denotes significance at the 0.1 / 1 / 5 / 10 percent level.

General prevention and eco-friendly behaviors are summarized in Table 7. Risk aversion is positively correlated with pro-environmental behavior (as predicted by Bramoullé and Treich, 2009) and prevention effort, irrespectively of the period in which the possible adverse event might happen. The relation with prudence, however, depends on the timing, which is in line with theory (see Table 4). For an adverse event that might happen in the same period, prudence is negatively correlated with prevention effort. For an adverse event that is separated from the preventative effort by some time delay, the coefficient of prudence is positive (but not significant in our sample). This is also the case for eco-friendly behavior, which might be seen as just a special case of a two-period prevention setting. Interestingly, temperance predicts long-term preventative effort, i.e., when effort precedes its effect. This is in line with our measure of temperance being interpreted as a measure for kurtosis aversion, i.e., aversion against adverse outcomes. Another interesting observation is that patience seems to have a similar relation to prevention as prudence: Patience is positively related with long-term prevention efforts, including eco-friendly behavior, but negatively with short-term prevention efforts. These results are robust to applying the ORIV technique to account for measurement error (see Table S-7c in the supplementary material).

5.4 Planning Behavior

The relation between higher order risk preferences and cautious planning behavior is reported in column 1 in Table 8. This intensity is measured by the additional time individuals plan to invest in certain situations, in which risk in the decision situation increases (such as for how long to prepare for an exam with uncertain scope).

Table 8: Planning Behavior and Preference for Competitive Income

	Cautious Planning		Preference for Competitive Income	
Risk aversion (AP)	−0.063	(0.050)	−0.070***	(0.023)
Prudence	0.028	(0.051)	0.054**	(0.024)
Temperance	0.170****	(0.045)	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 more cautious planning behavior or increasing preference for competitive income. Risk and time measures in standard deviations. Non-orthogonalized measure: Prudence (planning behavior) and risk aversion (preference for competitive income). Outcome indices are standard normalized and formed by adding z-Scores (preference for competetive income) or using PCA weights (planning behavior). Questions included in these indices are listed in Sections C.4 and C.5 of Online Appendix C. See Tables B-11 and B-12 for detailed regression results. Bootstrapped standard errors (1000 rep.) clustered at the session level in parentheses. ****/***/**/* denotes significance at the 0.1 / 1 / 5 / 10 percent level.

As predicted by theory (see Table 4), prudence is positively (but insignificantly) correlated with a more cautious planning behavior. The coefficient of temperance, however, is positive and significant, even when controlling for age and gender (see Table B-11 for detailed regression results) and when applying the ORIV technique to address measurement error (see Table S-7d in the supplementary material). As our temperance measure is a measure of kurtosis aversion (Denuit and Eeckhoudt, 2010), this result might be explained by participants deciding rather based on the perceived distribution of the risk than on proper optimization (similar to the case of the two-period prevention effort): The more they dislike adverse outcomes, the more they prepare to avoid these situations.

5.5 Preference for Competitive Income

Results on a preference for competitive income are reported in column 2 of Table 8. In line with earlier findings by Dohmen and Falk (2011), risk aversion is negatively correlated with a preference for a competitive income. The coefficient on prudence, however, has the opposite sign and is half as large as the coefficient of risk aversion. This result supports the identification strategy by Fuchs-Schündeln and Schündeln (2005) – in the aggregate. On the individual level, however, the mechanisms at play seem to be somewhat more nuanced, as prudence is not equal to risk aversion (see Section 3), and as the opposing signs of prudence and risk aversion indicate.

Our results are robust to controlling for age and gender, among others, and to accounting for measurement error using the ORIV technique (see Table S-7d in the supplementary material). Being female is associated with a lower preference for competitive income, independent of risk preferences, and the size of this association is comparable to an increase of about three SD in risk aversion (see Table B-12).

Table 9: Financial Decision Making

	Saving (with Debt)		Risky Investment		Insurance Demand	
Risk aversion (AP)	0.096	(0.105)	−0.082	(0.069)	−0.046	(0.066)
Prudence	0.058*	(0.034)	−0.055	(0.042)	−0.038	(0.036)
Temperance	0.087**	(0.035)	−0.037**	(0.015)	0.002	(0.071)
Impatience	−0.197***	(0.032)	0.016	(0.026)	−0.000	(0.039)
Other Factors	18		13		13	
R^2	0.19		0.17		0.06	
Observations	658		658		658	

Notes: Positive coefficients imply increasing engagement in saving, risky investment or insuring. Risk and time measures in standard deviations. Non-orthogonalized measure: Prudence (Saving and Insurance Demand), and temperance (Risky Investment). Outcome indices are standard normalized and formed by using PCA weights. Questions included in these indices are listed in Section C.6 of Online Appendix C. Influence factors controlled for are the same as in Section 3, except for Saving, where we additionally control for income risk, as the saving theory that we rely on here is about precautionary saving. See Tables B-13, B-14 and B-15 for detailed regression results. Bootstrapped standard errors (1000 rep.) clustered at the session level in parentheses. ****/***/**/* denotes significance at the 0.1 / 1 / 5 / 10 percent level.

5.6 Financial Decision Making

Results on financial decision making are reported in Table 9. Already for adolescents and as predicted by theory (c.f. Table 4), prudence matters for (net) saving, and temperance is negatively related to "risky investment", even when controlling for our, as it appears, important measure of time preferences (a control variable that is missing in the study by Noussair et al., 2014). Financial insurance demand is unrelated to risk aversion, and negatively related to prudence, although this result is not robust to inclusion of control variables in the corresponding regression (see Table B-15). However, this relation is predicted by theory for the case when the insured object is of uncertain value.[35] Moreover, temperance is significantly related to saving, as is impatience in a negative way.

Although these results are expected (see Table 4), they should be interpreted with some care, since in particular younger adolescents may only have limited exposure to and experience in certain domains of financial decision making – among them probably insurance and investment (despite our questions being adapted to our sample, and these results being unaffected by exclusion of the youngest two grades, see Table B-18). In this light, it might not be surprising that, e.g., temperance seems to be more important for saving than prudence, although theory posits an unambiguously positive relationship only for the latter (which we also find), or that the results regarding investment and insurance depend to a certain degree on the inclusion or exclusion of gender, age and financial control variables (see

[35]This assumption seems realistic given that we asked for smartphone and bike insurance, i.e., insurance for two items that are heavily used by adolescents and decline in value relatively fast, where the decline additionally depends on the treatment of the item.

Tables B-14 and B-15 in the Appendix). The fact that we find the theoretically predicted relations already among adolescents despite those limitations suggests that higher order risk preferences are robustly related to financial behavior (thus confirming the seminal results by Noussair et al., 2014).

6 When are Prudence and Temperance Important to Account for in Empirical Work?

So far, we have examined how experimental measures of risk aversion, prudence and temperance are related to many important domains of field behavior. Our inclusion of prudence and temperance has been motivated by theoretical predictions about their relationship with field behavior. In addition to significant estimates presented in the previous section (that are robust to accounting for measurement error as suggested by Gillen et al., 2019), variable selection with the Lasso approach also confirms their importance for predicting field behavior: When considering risk aversion, prudence, temperance and impatience as well as all control variables used throughout, prudence is selected for inclusion in the suggested model more often than risk aversion is (see Table B-17 in the Appendix). Yet, in the introduction we have argued that, up to date, the large majority of papers that relate experimental risk measures to field behavior do not consider the higher order risk preferences prudence and temperance, but are confined to (experimental or survey) measures of risk aversion. In this final section before concluding, we show that ignoring prudence and temperance might lead to wrong conclusions in empirical work.

6.1 Properly Assessing the Role of Risk Aversion

First, we illustrate that ignoring prudence and temperance can lead to wrong conclusions about the relationship of risk aversion to field behavior. We start by observing that in some regressions the coefficients of prudence and/or temperance have the opposite sign of the coefficient of risk aversion (e.g., with respect to health behavior, short-term prevention or the preference for a competitive income; see Tables 6 to 9). This already indicates that for some field behavior, controlling for both risk aversion and higher order risk preferences is important, because they work in different directions. Yet, this does not yet answer the question whether omitting higher order risk preferences would yield wrong conclusions about the role of risk aversion.

Therefore, we present in Table 10 all regressions from Tables 6 to 9 *without* including prudence and temperance. In this way, we can demonstrate how the estimated coefficients and their significance level for risk aversion react to the inclusion or exclusion of higher order risk preferences. If we first look at Table 10 from a mere descriptive point of view, we note from column [1] that in 9 out of 11 cases,

risk aversion is significant in regressions *without* prudence and temperance. So, the first impression suggests that risk aversion has a lot of predictive power in relation to field behavior – at least when a precise and continuous measure like ours is applied.[36] In column [2] of Table 10 we then reprint the coefficient for risk aversion from the regressions in Tables 6 to 9 that had accounted for prudence and temperance. Most importantly, column [3] highlights all cases in which the significance of risk aversion changed with the inclusion of higher order risk preferences. This column shows 6 cases where the significance of risk aversion disappears, plus one case (for prevention behavior in the short term) where risk aversion turns significant in column [2] (although only at a significance level of 10%) when it was insignificant in column [1]. Column [4] indicates four cases in which even the sign of risk aversion changes when higher order risk preferences are taken into account.

Table 10: Significance and Sign of Risk Aversion, Conditional on Estimation With or Without Higher Order Risk Preferences

Coefficient of risk aversion (columns [1] and [2] show estimated coefficients)	[1] EXCLUDING prudence and temperance	[2] INCLUDING prudence and temperance	[3] Change whether risk aversion is significant or not	[4] Change in sign of risk aversion
Health-Related Behavior (Table 6)				
Unhealthy Behavior	-0.077^{**}	0.024	YES	YES
Addictive Behavior	-0.080^{**}	0.007	YES	YES
Smartphone Addiction	-0.085^{**}	-0.023	YES	
Prevention and Eco-Friendly Behavior (Table 7)				
Prevention (Short Term)	-0.027	0.120^{*}	YES	YES
Prevention (Long Term)	0.086^{**}	0.167^{***}		
Eco-friendly Behavior	0.078^{*}	0.201^{**}		
Planning Behavior and Preference for Competitive Income (Table 8)				
Cautious Planning	0.088^{**}	-0.063	YES	YES
Preference for Competitive Income	-0.071^{***}	-0.070^{***}		
Financial Decision Making (Table 9)				
Saving (with Debt)	0.087^{**}	0.096	YES	
Risky Investment	-0.051^{**}	-0.082	YES	
Insurance Demand	-0.031	-0.046		
Total number of changes (out of 11 regressions)			7 (out of 11)	4 (out of 11)

Notes: Positive coefficients imply increasing preference for the respective behavior. Risk measures in standard deviations. Column [1] shows regression coefficients of risk aversion (non-orthogonalized measure) without including higher order risk preferences in the models; column [2] shows the regression coefficients when including higher order risk preferences in the models as shown in Tables 6 to 9. See Tables 6 to 9 for these results and additional notes on the respective models. Bootstrapped standard errors (1000 rep.) clustered at the session level in parentheses. ****/***/**/* denotes significance at the 0.1 / 1 / 5 / 10 percent level.

[36] Such a continuous measure, corrected for measurement error, naturally has a higher correlation with prudence. It can thus pick up variation that is actually due to prudence if the approriate control is missing. See Section 6.3 below for another example.

Thus, Table 10 reveals that it matters substantially whether one relates only risk aversion to field behavior or whether higher order risk preferences are also considered. Once prudence is taken into account, risk aversion turns insignificant for unhealthy behavior, addictive behavior and smartphone addiction, but also for cautious planning, savings and risky investment. We think this is particularly important with respect to health-related behavior, because this type of behavior accounts for two of the four cases where the coefficient of risk aversion even changes its sign when higher order risk preferences are included. When applying the ORIV technique (Gillen et al., 2019), the coefficient of risk aversion even becomes positive in all three regressions of health-related behavior once higher order risk preferences are accounted for (all other results from Table 10 are identical when using ORIV; see Table B-19 in Appendix A). Thus, as apparent from the formula for omitted variable bias, when failing to control for prudence (and to a lesser extent temperance), the true relation between risk aversion and field behavior is imprecise, and it might be even blurred further by the level of prudence that could be captured by a risk aversion measure.[37] Consequently, the estimated relation between risk preferences and, e.g, health-related behavior is vastly underestimated (compare the coefficients of risk aversion in column [1] of Table 10 with those of prudence in Table 6). This can lead to unwanted consequences, especially if the applied measure of risk aversion is unrelated with prudence, as the following analysis illustrates.

6.2 Properly Investigating the Relation of Risk Preferences in General to Field Behavior

We now demonstrate how neglecting higher order risk preferences may lead to wrong conclusions not only about the relation between risk aversion to field behavior, but also about the relation between field behavior and risk preferences in general when using experimental measures. To do so, we re-estimate in Table 11 the models from Tables 6 to 9 and include the same explanatory control variables (i.e., impatience, cognitive abilities, gender, age, ...), but this time we limit the set of explanatory risk preferences to one single risk preference only, as is typically done in empirical work (see, e.g., Galizzi et al., 2016; Charness et al., 2020; Samek et al., 2021). To mirror such a typical approach as closely as possible, we additionally use only simple, but commonly used measures for risk aversion and prudence – which we then contrast with our results on (jointly accounting for) risk aversion, prudence and temperance from Section 5.

[37] See, e.g., the correlation of prudence and the standard one-item survey question on general willingness to take risk in Table 5, and the consequences discussed below.

Table 11: Significance of Risk Preferences in Predicting Field Behavior

Coefficient of risk preferences (columns show estimated coefficients)	Simple Measures		Measures from Section 5 (Joint Estimation)			Survey Measure
	[1] Simple Risk (r^{CE})	[2] Simple Prudence (counting measure)	[3] Risk Aversion (Arrow-Pratt)	[4] Prudence (Crainich-Eeckhoudt)	[5] Temperance (Denuit-Eeckhoudt)	[6] Survey Measure (Risk Tolerance)
Health-Related Behavior (Table 6)						
Unhealthy Behavior	−0.031	−0.085***	0.024	−0.140****	−0.011	0.080*
Addictive Behavior	−0.036	−0.085***	0.007	−0.146****	−0.002	0.084*
Smartphone Addiction	−0.038	−0.071**	−0.023	−0.160****	0.017	0.080
Prevention and Eco-Friendly Behavior (Table 7)						
Prevention (Short Term)	0.020	−0.027	0.120*	−0.119****	0.017	0.173****
Prevention (Long Term)	0.124****	−0.008	0.167***	0.010	0.078**	−0.032
Eco-friendly Behavior	0.094**	−0.011	0.201**	0.022	0.034	−0.013
Planning Behavior and Preference for Competitive Income (Table 8)						
Cautious Planning	0.062*	−0.044	−0.063	0.028	0.170****	−0.082**
Preference for Competitive Income	−0.075****	0.008	−0.070***	0.054**	0.032	0.039
Financial Decision Making (Table 9)						
Saving (with Debt)	0.067	0.011	0.096	0.058*	0.087**	−0.057
Risky Investment	−0.056	0.010	−0.082	−0.055	−0.037**	0.135***
Insurance Demand	−0.048	0.014	−0.046	−0.038	0.002	0.082

Notes: Positive coefficients imply increasing preference for the respective behavior. All measures in standard deviations. Column [1] shows the regression coefficients of "risk preferences" from the models shown in Table 6 to 9 when only using the simple risk aversion measure as measure of risk preferences; column [2] shows coefficients when using a counting measure of prudence instead. Columns [3] to [5] show results when measuring risk preferences with the continuous intensity measures of risk aversion, prudence and temperance as resulting from our method in the same model, as done in Tables 6 to 9 (see the legends to these tables for additional notes on the respective models). Column [6] shows regression coefficients when risk preferences are measured with the one-item survey question on willingness to take risk in general. Bootstrapped standard errors (1000 rep.) clustered at the session level in parentheses.
****/***/**/* denotes significance at the 0.1 / 1 / 5 / 10 percent level.

In column [1] of Table 11 we use the simple measure of risk aversion (r^{CE}), as defined in Section 2.3 (and used, e.g., in Sutter et al., 2013), as the single measure of risk preferences. As can be seen at the top of the table, one would conclude that risk preferences do not matter for any of our scales on health-related behavior. Yet, such a conclusion would have to be refuted if one took a simple count measure of prudence (à la Eeckhoudt and Schlesinger, 2006, see Section 2.3) as explanatory variable, as is done in column [2]. By comparing the coefficients in column [2] to those in column [4] one would then see, however, that the simple count measure of prudence underestimates the influence of risk preferences considerably when compared to our intensity measure of prudence. This shows that risk preferences matter for health-related behavior, even though this can only be revealed when using a measure of

prudence – even the most simple one – rather than a measure of risk aversion.[38] Regarding (short-term) prevention behavior and financial decision making (in the middle and at the bottom of Table 11), one could also (wrongly) conclude that risk preferences do not matter if one took the simple risk measure from column [1]. Yet again, risk preferences matter, but this time this can only be seen when prudence (for the case of prevention and saving) and temperance (for the case of investment and saving) are used as explanatory variables. Finally, with respect to cautios planning, using the simple risk measure would indicate a relation with risk aversion, whereas in reality, temperance is the relevant behavior.

6.3 Properly Interpreting Survey Measures of Risk Tolerance

Columns [1] to [5] in Table 11 have used experimental measures of risk preferences to explain different domains of field behavior. Many papers on the relation of risk preferences to field behavior use survey measures, however (e.g., Barsky et al., 1997; Caliendo et al., 2007; Dohmen et al., 2011). In our study, we have also collected such a survey measure of risk tolerance, which is often interpreted as a measure of risk aversion. Importantly, we have seen already in Table 5 that this one-item survey measure on risk tolerance is not only related to risk aversion, but also – and significantly so – to prudence and temperance. This raises the question what can be inferred from a significant relation of such a survey-based measure of risk tolerance to field behavior. To discuss this, we present, first, in column [6] of Table 11 the estimates for the survey measure. Disregarding for the moment the low precision of some results that are significant only at the 10% level, the survey measure is actually more often successful in predicting field behavior (5 significant estimations) than the simple measures of risk aversion (4 significant relations in column [1]) or the simple count measure of prudence (3 significant relations in column [2]). Thus, the survey measure seems to capture more dimensions of risk preferences than risk aversion. However, precisely because of this feature, caution is warranted in interpreting the established relations with the survey measure: In none of the five cases where the survey measure on risk tolerance predicts field behavior, interpreting this measure as a measure of risk aversion is appropriate: In three of the five cases, the survey measure picks up the relation with prudence (health-domain and, considering the sign of the coefficient, also short-term planning), while in the remaining two cases, the relation with temperance is captured (cautious planning and risky investment). Also noteworthy, in three of the four cases where the simple risk measure (from column [1]) leads to a significant relation with field behavior (that are confirmed

[38] Including a measure of risk preferences that is higher correlated with prudence might already show an association, see the previous and following subsections. Note, however, that neglecting prudence still biases the estimation towards zero.

by the Arrow-Pratt measure resulting from our method in column [3]), the survey measure fails to relate to these outcomes (long-term prevention, eco-friendly behavior, and preference for a competitive income). These observations provide a novel perspective on the inconsistent findings on risk preferences and their relation to field behavior mentioned in the introduction: While all the mentioned failures rely on experimental measures of risk preferences and neglect higher order risk preferences (Sutter et al., 2013; Galizzi et al., 2016; Charness et al., 2020; Samek et al., 2021), all of the mentioned successes (except for Anderson and Mellor, 2008) rely on the survey measure that is obviously capable of capturing additional dimensions of risk preferences on top of risk aversion alone.

So, overall our Tables 10 and 11 have shown that ignoring higher order risk preferences – by not capturing them at all, or by not properly accounting for them – leads frequently to wrong conclusions, most prominently with respect to health-related behavior (also see, e.g., Galizzi and Miraldo, 2017).[39] Taking into account higher order risk preferences has also shed new light on how to interpret the significant relationship of survey measures on risk tolerance to field behavior (see, e.g., Caliendo et al., 2007; Dohmen et al., 2011; Becker et al., 2012). Obviously, these survey measures capture a relationship to field behavior. We have shown, however, that this might be the case because they are related to higher order risk preferences and not only to risk aversion. In fact, when the survey measure on risk tolerance is a significant predictor for field behavior in our sample, this is never due to risk aversion driving the relationship, but always because of prudence and temperance.

7 Conclusion

In this paper, we have analyzed how experimentally elicited intensity measures of risk aversion, prudence and temperance as well as of time preferences relate to field behavior concerning decisions with uncertain outcomes such as health-related behavior, eco-friendly behavior, planning behavior, or financial decision making.

We have used a novel – and in-sample validated – method, developed specifically for this and its companion paper, Schneider et al. (2021), to measure risk aversion and the higher order risk preferences prudence and temperance, that is, to quantify their intensities non-parametrically with utility-based measures. As the method accounts for measurement error and the resulting measures are the ones used in theoretical work, this enables us to empirically study field behavior closely guided by theory. In our sample of 658 students from sixth to twelfth grade in German

[39] Of course, this observation extends to using risk preference measures as control variables: If the relevant control variable is missing in a regression, standard errors of other estimates might increase, but risk preferences might also be attributed to other characteristics. Both might dramatically alter conclusions (see, e.g., Gillen et al., 2019).

schools, we have found clear evidence for risk aversion, prudence and temperance in the aggregate. These findings are in line with studies on adult populations (e.g., Noussair et al., 2014; Deck and Schlesinger, 2014; Ebert and Wiesen, 2014). We have found no significant age effects for any of our preferences. We find females exhibiting more risk averse, more prudent and more temperant behavior. Cognitive abilities and prudence are unrelated, while cognitive abilities are negatively related to risk aversion and temperance in our study with adolescents.

The most important findings of our paper concern the relationship of experimental measures and field behavior and, in particular, the importance of prudence and temperance with respect to understanding behavior that involves some sort of risk. Starting with the most general result, our measures and the single-item willingness-to-take-risk question (Dohmen et al., 2011) are highly correlated. Comparing regression coefficients of standardized risk measures, prudence (and to a lower degree temperance) seems to have a sizeable influence on general risk taking, representing about 35% to 50% of the influence of risk aversion. This suggests that risk tolerance, as used in everyday language, is only insufficiently captured by measures of risk aversion alone, extending related results by Jaspersen et al. (2020). This finding is corroborated in a regression exercise showing that the survey measure of risk tolerance is successful in predicting field behavior precisely because it is related to prudence and temperance. In fact, none of the documented relations is driven by risk aversion. We thereby provide new insights on how to interpret the significant relationship of survey measures on risk tolerance to field behavior.

We provide further evidence on the importance of higher order risk preferences by illustrating that – even when addressing measurement error by additionally using ORIV (Gillen et al., 2019) – the significance (and even the sign) of our risk aversion parameter depends in a large number of cases on whether or not we include prudence and temperance. The most striking case refers to behavior in the health domain: Unhealthy and addictive behavior, in particular our smartphone addiction scale, is strongly related to imprudence, but not to risk aversion, even though using only our risk aversion measure (and excluding prudence and temperance) seems to suggest that risk aversion and this health-related behavior are linked (although, due to omitted variable bias, with coefficients of an order of magnitude lower).

Neglecting prudence and temperance might be even more consequential when in addition the measure of risk preference is uncorrelated with prudence: When using a simple measure of risk aversion to account for risk preferences and neglecting higher order risk preferences, health-related behavior even seems unrelated to risk preferences at all. The picture changes completely, however, once a measure of prudence is used instead, even if it is a simple count measure. Similarly, short-term

prevention efforts and financial decision making seem to be unrelated to risk preferences when using a simple measure of risk aversion. For these kinds of behavior, however, prudence and temperance are needed to establish the relationship.

In conclusion, this is not to say that risk aversion never matters, but our findings advocate caution when drawing inferences from studies that ignore prudence and temperance. More generally, our results demonstrate that some behavior is only predicted by prudence, such as health-related behavior, whereas other behavior seems to depend on a combination of risk aversion, prudence, and, to a lesser extent, temperance. Thus, whether or not a certain behavior is related to risk attitudes depends on the nature of the risk. The absence of a correlation with the attitude towards a symmetric gamble, which would be captured by classical risk aversion, does not necessarily rule out that individuals perceive a certain behavior as risky. It might just also be the case that prudence is the better (and sometimes only) predictor for that kind of behavior.

Our study provides further evidence to the results by Gillen et al. (2019) in illustrating the importance to account for measurement error in experimental work. Besides illustrating the importance of higher order risk preferences, our last section also illustrates the shortcomings of imprecise or coarse measures, and how results can change by using measures like ours inherently accounting for measurement error: As estimation precision is improved, additional relations can be uncovered, and findings can be documented at significance levels that have been increasingly advocated for recently (such as $p < 0.005$). These results also illustrate the effectiveness of our method in addressing measurement error, which is further underlined by correlation coefficients to standard survey questions that exceed common values in the literature by an order of magnitude.

Another important aspect of our paper is the fact that we have been able to provide empirical support for theoretical predictions emerging from several models that have lacked empirical evidence so far. The model by Bramoullé and Treich (2009), for example, suggesting that uncertainty might alleviate the commons problem, posits that risk aversion decreases pollution due to uncertainty. In fact, we find support for the claim that risk averters behave more eco-friendly. Our results with respect to prevention effort of a possibly contemporaneous unwanted event (negative relation) are in line with theory (Eeckhoudt and Gollier, 2005), while we find indication (but not significantly) for the prediction that the relation flips when the possible unwanted event follows the effort only with some time delay (Menegatti, 2009). Moreover, we provide support for the theoretical predictions related to financial decision making and higher order risk preferences (e.g., Leland, 1968; Kimball, 1990, 1992) already among adolescents. Although they certainly have less training

in this area compared to adults, already adolescents have to make financial decisions on a daily basis in dealing, e.g., with their pocket money. We have shown that decisions of that kind are related to higher order risk preferences, and thus, as a minor result of our study, can extend the seminal finding for adults by Noussair et al. (2014) to our sample of adolescents.

To conclude with a potential policy implication of our study, our results suggest that higher order risk preferences could be used for an efficient identification of adolescents that might be prone to problematic health-related behavior, in particular addictive behavior. The latter aspect makes our results particularly timely, given the age of our sample and the growing evidence that adolescence is crucial in developing addictive behavior. For example, smoking experimentation of any level before adulthood, including only a few puffs, is associated with an at least 26% increased risk of being a smoker 20 years later compared to those who never smoked (Paul et al., 2008). The age of smoking the first cigarette during childhood and adolescence is a highly significant predictor of smoking status, nicotine dependence, and monthly cigarette consumption at age 22 (Buchmann et al., 2011). Similar patterns have been documented for alcohol (e.g., Grant and Dawson, 1997; DeWit, 2000), hard drugs (e.g., Lynskey, 2003; Chen et al., 2009) and gambling addiction (e.g., Jiménez-Murcia et al., 2010; Black et al., 2015). The conclusion from this literature is always the same: Early prevention is key! So, given that we find no age effects, but rather stable relationships of prudence and temperance to field behavior across the whole age range studied in this paper, our experimental measures of higher order risk preferences could be very helpful in identifying youths at risk of developing harmful habits and field behavior.

References

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.

Andreassen, C. S. (2015). Online social network site addiction: A comprehensive review. *Current Addiction Reports 2*(2), 175–184.

Andreassen, C. S., S. Pallesen, and M. D. Griffiths (2017). The relationship between addictive use of social media, narcissism, and self-esteem: Findings from a large national survey. *Addictive Behaviors 64*, 287–293.

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. *The 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.

Becker, A., T. Deckers, T. Dohmen, A. Falk, and F. Kosse (2012). The relationship between economic preferences and psychological personality measures. *Annual Review of Economics 4*(1), 453–478.

Black, D. W., M. Shaw, W. Coryell, R. Crowe, B. McCormick, and J. Allen (2015). Age at onset of DSM-IV pathological gambling in a non-treatment sample: Early- versus later-onset. *Comprehensive Psychiatry 60*, 40–46.

Bleichrodt, H., D. Crainich, and L. Eeckhoudt (2003). The effect of comorbidities on treatment decisions. *Journal of Health Economics 22*(5), 805–820.

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.

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

Breaban, A., G. van de Kuilen, and C. N. Noussair (2016). Prudence, emotional state, personality, and cognitive ability. *Frontiers in Psychology 7*, 1688–1698.

Buchmann, A. F., D. Blomeyer, C. Jennen-Steinmetz, M. H. Schmidt, G. Esser, T. Banaschewski, and M. Laucht (2011). Early smoking onset may promise initial pleasurable sensations and later addiction. *Addiction Biology 18*(6), 947–954.

Caballé, J. and A. Pomansky (1996). Mixed risk aversion. *Journal of Economic Theory 71*(2), 485–513.

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.

Chapman, J., M. Dean, P. Ortoleva, E. Snowberg, and C. Camerer (2018). Econographics. NBER Working Paper No. 24931.

Charness, G., C. Eckel, U. Gneezy, and A. Kajackaite (2018). Complexity in risk elicitation may affect the conclusions: A demonstration using gender differences. *Journal of Risk and Uncertainty 56*(1), 1–17.

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.

Chen, C.-Y., C. L. Storr, and J. C. Anthony (2009). Early-onset drug use and risk for drug dependence problems. *Addictive Behaviors 34*(3), 319–322.

Courbage, C. and B. Rey (2006). Prudence and optimal prevention for health risks. *Health Economics 15*(12), 1323–1327.

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

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

Crosetto, P. (2021). METARET: A Meta Analysis of the external validity of Risk Elicitation Tasks. OSF (http://doi.org/10.17605/OSF.IO/H2Z56) and https://paolocrosetto.shinyapps.io/METARET/. Accessed: July 9, 2021.

Crosetto, P. and A. Filippin (2015). A theoretical and experimental appraisal of four risk elicitation methods. *Experimental Economics 19*(3), 613–641.

Croson, R. and U. Gneezy (2009). Gender differences in preferences. *Journal of Economic Literature 47*(2), 448–474.

Dardanoni, V. and A. Wagstaff (1990). Uncertainty and the demand for medical care. *Journal of Health Economics 9*(1), 23–38.

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

Deck, C. and H. Schlesinger (2014). Consistency of higher order risk preferences. *Econometrica 82*(5), 1913–1943.

Deck, C. and H. Schlesinger (2017). On the robustness of higher order risk preferences. *Journal of Risk and Insurance 85*(2), 313–333.

Denuit, M. M. and L. Eeckhoudt (2010). Stronger measures of higher-order risk attitudes. *Journal of Economic Theory 145*(5), 2027–2036.

DeWit, D. J. (2000). Age at first alcohol use: A risk factor for the development of alcohol disorders. *American Journal of Psychiatry 157*(5), 745–750.

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

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

Dohmen, T., A. Falk, D. Huffman, and U. Sunde (2018). On the relationship between cognitive ability and risk preference. *Journal of Economic Perspectives 32*(2), 115–134.

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.

Ebert, S. (2012). Moment characterization of higher-order risk preferences. *Theory and Decision 74*(2), 267–284.

Ebert, S. and D. Wiesen (2011). Testing for prudence and skewness seeking. *Management Science 57*(7), 1334–1349.

Ebert, S. and D. Wiesen (2014). Joint measurement of risk aversion, prudence, and temperance. *Journal of Risk and Uncertainty 48*(3), 231–252.

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

Eeckhoudt, L. and M. Kimball (1992). Background risk, prudence, and the demand for insurance. In G. Dionne (Ed.), *Contributions to Insurance Economics*, pp. 239–254. Dordrecht: Springer Netherlands.

Eeckhoudt, L., B. Rey, and H. Schlesinger (2007). A good sign for multivariate risk taking. *Management Science 53*(1), 117–124.

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

Eeckhoudt, L., H. Schlesinger, and I. Tsetlin (2009). Apportioning of risks via stochastic dominance. *Journal of Economic Theory 144*(3), 994–1003.

Eeckhoudt, L. R., R. J. A. Laeven, and H. Schlesinger (2020). Risk apportionment: The dual story. *Journal of Economic Theory 185*, 104971.

Ehrlich, I. and G. S. Becker (1972). Market insurance, self-insurance, and self-protection. *Journal of Political Economy 80*(4), 623–648.

Eilers, P. H. and B. D. Marx (1996). Flexible smoothing with B-splines and penalties. *Statistical Science 11*(2), 89–102.

Epper, T. and H. Fehr-Duda (2018). The missing link: Unifying risk taking and time discounting. Zurich Working paper series / Dep. of Economics No. 96.

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

Familien-wegweiser.de (2018). Taschengeld. http://www.familien-wegweiser.de/wegweiser/stichwortverzeichnis,did=38294.html (new address: https://familienportal.de/familienportal/lebenslagen/kinder-jugendliche/taschengeld). Accessed: June 11, 2018.

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. *The Quarterly Journal of Economics 120*(3), 1085–1120.

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, London School of Economics and Political Science, Department of Social Policy, 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).

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.

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

Grant, B. F. and D. A. Dawson (1997). Age at onset of alcohol use and its association with DSM-IV alcohol abuse and dependence: results from the national longitudinal alcohol epidemiologic survey. *Journal of Substance Abuse 9*, 103–110.

Gruber, M. H. J. (2017). *Improving efficiency by shrinkage*. Routledge.

Haering, A., T. Heinrich, and T. Mayrhofer (2020). Exploring the consistency of higher order risk preferences. *International Economic Review 61*(1), 283–320.

Harrison, G. W. (1986). An experimental test for risk aversion. *Economics Letters 21*(1), 7–11.

He, Q., O. Turel, and A. Bechara (2017). Brain anatomy alterations associated with Social Networking Site (SNS) addiction. *Scientific Reports 7*(1).

Heinrich, T. and J. Shachat (2020). The development of risk aversion and prudence in Chinese children and adolescents. *Journal of Risk and Uncertainty 61*(3), 263–287.

Hoerl, A. E. and R. W. Kennard (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics 12*(1), 55–67.

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

Hormes, J. M., B. Kearns, and C. A. Timko (2014). Craving Facebook? Behavioral addiction to online social networking and its association with emotion regulation deficits. *Addiction 109*(12), 2079–2088.

Jaspersen, J. G., M. A. Ragin, and J. R. Sydnor (2020). Linking subjective and incentivized risk attitudes: The importance of losses. *Journal of Risk and Uncertainty 60*(2), 187–206.

Jiménez-Murcia, S., E. M. Álvarez-Moya, R. Stinchfield, F. Fernández-Aranda, R. Granero, N. Aymamí, M. Gómez-Peña, N. Jaurrieta, F. Bove, and J. M. Menchón (2010). Age of onset in pathological gambling: Clinical, therapeutic and personality correlates. *Journal of Gambling Studies 26*(2), 235–248.

Kimball, M. S. (1990). Precautionary saving in the small and in the large. *Econometrica 58*(1), 53–73.

Kimball, M. S. (1992). *New Palgrave Dictionary of Money and Finance*, Chapter Precautionary Motives for Holding Assets, pp. 158–161. MacMillan, London.

Krieger, M. and T. Mayrhofer (2012). Patient preferences and treatment thresholds under diagnostic risk. Ruhr Economic Paper No. 321.

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

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *The Quarterly Journal of Economics 112*(2), 443–478.

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

Lynskey, M. T. (2003). Escalation of drug use in early-onset cannabis users vs co-twin controls. *JAMA 289*(4), 427.

Maier, J. and M. Rüger (2011). Experimental evidence on higher-order risk preferences with real monetary losses. Working Paper, University of Munich.

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

Menezes, C., C. Geiss, and J. Tressler (1980). Increasing downside risk. *The American Economic Review 70*(5), 921–932.

Modica, S. and M. Scarsini (2005). A note on comparative downside risk aversion. *Journal of Economic Theory 122*(2), 267–271.

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.

Paul, S. L., L. Blizzard, G. C. Patton, T. Dwyer, and A. Venn (2008). Parental smoking and smoking experimentation in childhood increase the risk of being a smoker 20 years later: The Childhood Determinants of Adult Health Study. *Addiction 103*(5), 846–853.

Pedroni, A., R. Frey, A. Bruhin, G. Dutilh, R. Hertwig, and J. Rieskamp (2017). The risk elicitation puzzle. *Nature Human Behaviour 1*(11), 803–809.

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

Przybylski, A. K. and N. Weinstein (2017). A large-scale test of the Goldilocks Hypothesis: Quantifying the relations between digital-screen use and the mental well-being of adolescents. *Psychological Science 28*(2), 204–215.

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.

Rothschild, M. and J. E. Stiglitz (1970). Increasing risk: I. A definition. *Journal of Economic Theory 2*(3), 225–243.

Rubin, D. B. (1996). Multiple imputation after 18+ years. *Journal of the American Statistical Association 91*(434), 473–489.

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

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.

Schildberg-Hörisch, H. (2018). Are risk preferences stable? *Journal of Economic Perspectives 32*(2), 135–154.

Schneider, S. O. and G. Baldini (2021a). oTree: Ready-made apps for higher order risk preference elicitation methods. Mimeo.

Schneider, S. O. and G. Baldini (2021b). utilityFunctionTools. R package, available via CRAN: https://cran.r-project.org/package=utilityFunctionTools.

Schneider, S. O., M. Ibanez, and G. Riener (2021). Measuring utility – An application to higher order risk. Mimeo.

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

Sutter, M., C. Zoller, and D. Glätzle-Rützler (2019). Economic behavior of children and adolescents – A first survey of experimental economics results. *European Economic Review 111*, 98–121.

Sydnor, J. (2010). (Over)insuring modest risks. *American Economic Journal: Applied Economics 2*(4), 177–199.

Tarazona-Gomez, M. (2004). Are individuals prudent? An experimental approach using lottery choices. EHESS Toulouse Working Paper.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological) 58*(1), 267–288.

Trautmann, S. T. and G. van de Kuilen (2018). Higher order risk attitudes: A review of experimental evidence. *European Economic Review 103*, 108–124.

Twenge, J. M., T. E. Joiner, M. L. Rogers, and G. N. Martin (2018). Increases in depressive symptoms, suicide-related outcomes, and suicide rates among u.s. adolescents after 2010 and links to increased new media screen time. *Clinical Psychological Science 6*(1), 3–17.

Vischer, T., T. Dohmen, A. Falk, D. Huffman, J. Schupp, U. Sunde, and G. G. Wagner (2013). Validating an ultra-short survey measure of patience. *Economics Letters 120*(2), 142–145.

Wagner, G. G., J. R. Frick, and J. Schupp (2007). The German Socio-Economic Panel Study (SOEP) – Evolution, scope and enhancements. *Schmollers Jahrbuch: Journal of Applied Social Science Studies / Zeitschrift für Wirtschafts- und Sozialwissenschaften 127*(1), 139–169.

Weber, E. U., A.-R. Blais, and N. E. Betz (2002). A domain-specific risk-attitude scale: Measuring risk perceptions and risk behaviors. *Journal of Behavioral Decision Making 15*(4), 263–290.

A Online Appendix: Details: Methods and Design

Table A-1: Illustration of the Staircase Method

		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 2 for details.

This exemplary illustration of the staircase method for certainty equivalents also 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 (lowest possible future equivalent), and H_1 with 140 (highest possible future equivalent). The immediate payment stays 100 Taler throughout the four iterations.

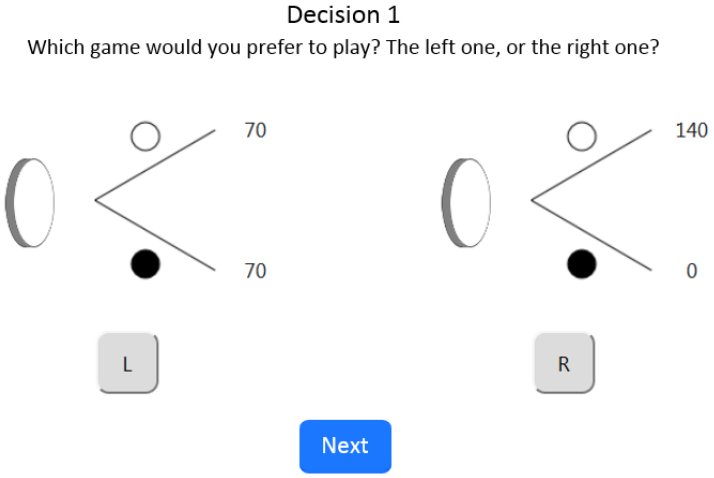


Figure A-1: Elicitation of (Higher Order) Risk Preferences via Certainty Equivalents: Exemplary Decision Screen with a Animated Rotating Coin

A.1 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. 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). One of the 658 individuals chose the gambling path throughout all six certainty equivalents (compared to 28 who always choose the safe path). 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.

A.2 Details on our Method

The basic 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 defined 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, and consequently obtaining only one coefficent for each term for the whole interval as in a polynomial regression, spline regression regresses on a large number of these terms that are only piece-wise defined (and zero otherwise). This allows to obtain a different coefficient for, e.g., the squared term, 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 steep (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 the increas(s) thereof). Increasing the order of polynomials while sticking to globally 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) has pointed out already at the beginning of the last century.

To overcome the issue of subjectively deciding about the number of sub-intervals as well as the position of their boundaries, penalized regression relies on an abundant number of piece-wise defined basis functions. Overfitting is then tackled in a data-driven way by using a penalization term that balances the trade-off between smoothness and fit to the data, and ensures that just the right amount of flexibility is used. In that way, the coefficients of, e.g., two adjacent basis functions of the same degree will only differ if it favours the fidelity/smoothness-trade-off. Here, the weight of the penalty term, i.e., the solution to this trade-off, is determined by optimizing the function's predictive quality via cross-validation. We adapt this approach for the use with utility functions by incorporating value constraints, such that the utility of having nothing can be set to zero, and the utility of receiving the heighest amount can be set to 1. Moreover, we incorporate a penalty term that allows for joint smoothing of several derivatives. For the statistical details on these non-trivial adaptations, see our companion paper, Schneider et al. (2021). We use this approach with 20 sub-intervals, and basis functions up to degree six $(x, x^2, \ldots, x^6)$, such that the fourth derivative is a possibly quadratic, continuous and differentiable function. Figure A-2 in the Appendix depicts an exemplary estimated utility function with derivatives.

The vast advantage of this approach over a simple linear interpolation is its ability of error correction: The regression approach considers the overall shape of the utility function instead of just the change from one point to another while neglecting all other data (i.e., most of the available data), as linear interpolation does. This impedes error correction, but what is more important, it would also be possible to classify a subject's utility curve as convex, if, e.g., the majority of changes in the function's slopes imply an increase in the slope, although the overall shape of the function is clearly concave, because the minority of changes is implying a decrease in the slope, but at a higher degree as the majority of changes (for an example, see our companion paper, Schneider et al., 2021). Lastly, linear interpolation would only allow for a crude intensity measure containing considerable noise, which is unlikely to possess enough precision for predicting field behavior and an empirical investigation of theory.

Another crucial advantage of the spline approach is that from utility functions established in that way, derivatives can be calculated analytically with a closed form solution without the need for additional numerical computation. Based on the derivatives of the utility function, attitude measures can be calculated as follows.

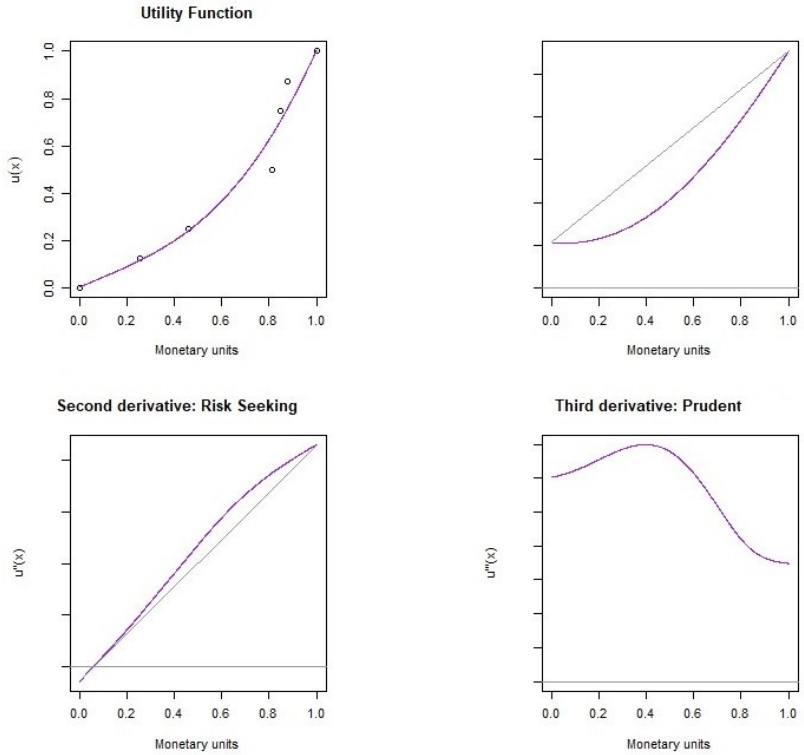


Figure A-2: Estimation of Utility Functions (Including Derivatives) from Utility Points: Example

A.3 Mixed Risk Aversion and Mixed Risk Loving Behavior

Dividing our sample into risk seekers ($r < 0$) and risk averters ($r > 0$) reveals that the sign of the correlation between risk aversion and prudence changes: For risk seekers, the degree of prudence increases as the degree of risk aversion and temperance *decreases*, because the correlation coefficient of risk aversion is negative (see central panel of Table A-3). Contrarily, for risk averters prudence increases when risk aversion and temperance *increase* (see the right panel of Table A-3). The relation between risk aversion and temperance, however, is positive and independent of being risk averse or not. These patterns are in line with the existence of both, mixed risk aversion (Caballé and Pomansky, 1996) and mixed risk-loving behavior (Crainich et al., 2013): Mixed risk aversion, nowadays mostly referred to as a preference for combining good with bad (Eeckhoudt et al., 2009), was originally defined via alternating signs of the derivatives of the utility functions (i.e., risk averse, prudent, and temperant), whereas mixed risk-loving behavior would correspond to positive derivatives of the utility function throughout (i.e., risk seeking, prudent, and intemperant). Note that the positive correlation of prudence with both risk aversion and temperance in the full sample is not a contradiction to the existence of these types; it is merely a result of the share of risk averters (with a positive correlation) in our sample being about twice as large as the share of risk seekers (with

Table A-2: Validation of Elicitation Method: Identification of Risk

	Risk Aversion (Intensity$^\sharp$)		Risk Aversion (Classification†)	
Simple Risk (Sutter et al.)	0.205****	(0.030)	0.764****	(0.100)
N	658		658	

Notes: Positive coefficients imply increased intensity measures of risk aversion as resulting from our method (expressed in standard deviations) or an increased likelihood of being classified as risk averse. Regression coefficients resulting from OLS regression (without constants) in the first column and probit regression in the second column. Bootstrapped standard errors (1000 reps.) clustered at the session level in parentheses.
$^\sharp$We use the Arrow-Pratt measure r as used in most of the empirical part of the paper, i.e., after having removed the variation explained by prudence. See *Attitude Measures* in Section 2.3 for details and definition of the measure. †We classify subjects as risk averse if their Arrow-Pratt measure is positive, i.e., if $r > 0$.
****/***/**/* denotes significance at the 0.1 / 1 / 5 / 10 percent level.

Table A-3: Correlation Between (Higher Order) Risk Preferences and Impatience

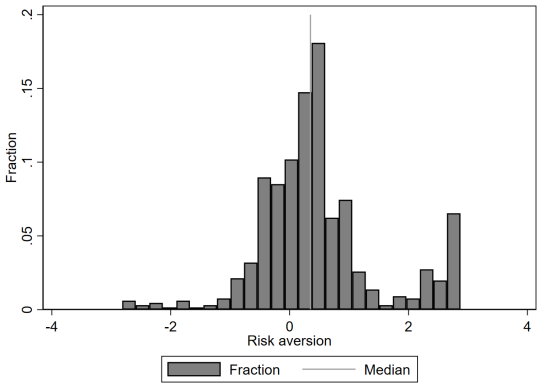
	Full Sample			Risk Seeking Subjects		Risk Averse Subjects	
	Risk aversion	Prudence	Temperance	Risk aversion	Prudence	Risk aversion	Prudence
Prudence	0.559***			−0.876***		0.928***	
Temperance	0.867***	0.652***		0.846***	−0.878***	0.917***	0.954***
Impatience	−0.133***	−0.0926**	−0.112***				
Observations	658			198		460	

Notes: Pearson correlation coefficients reported; *** denotes significance at the 1 percent level.

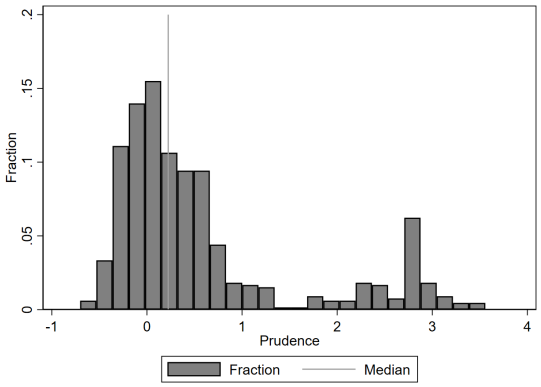
a negative correlation). This can also be seen in the lower (absolute) correlation in the whole sample compared to the higher correlations – with opposite signs – for the risk averse ($N = 460$) and risk seeking ($N = 198$) subsamples - see the central and right panels of Table A-3.

Thus, already among adolescents we find support for the existence of mixed risk averters (Caballé and Pomansky, 1996) and mixed risk lovers (Crainich et al., 2013), which is not yet well documented among adults, but in line with findings by Deck and Schlesinger (2014), Ebert and Wiesen (2014), and Haering et al. (2020). The results by Deck and Schlesinger (2017) and Haering et al. (2020) suggest that mixed risk aversion and mixed risk loving behavior results mainly from the compound lottery design used in these studies. Our results speak against this hypothesis, as we use a completely different elicitation method with binary outcomes only, suggesting that the pattern is more robust to presentation and elicitation method as commonly assumed, and even translates to the internal margin.

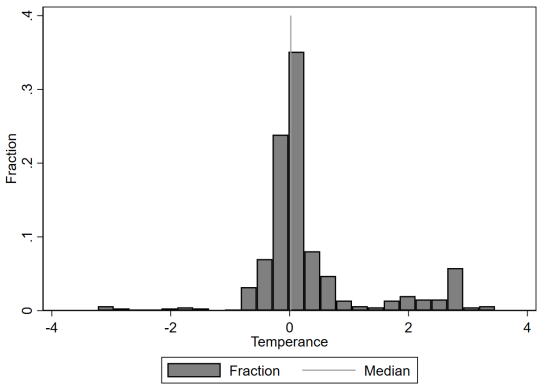
B Online Appendix: Additional Tables and Figures



(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 B-1: Prevalences of (Higher Order) Risk Preferences (in SD)

Table B-1: Influence Factors of Risk Aversion (Arrow-Pratt Measure)

	Dependent Variable: Risk Aversion			
	(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. 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 fulltime', and 'Catholic' for religion. For 24 participants, some demographic information has been imputed with 0, the variable's mean value, and the 'other' category for binary, continuous and categorical variables, respectively. We controlled for imputation with indicator variables. See Table S-2 in the supplementary material on our website for regression results excluding participants that reported problems with handling their tablets during our study. 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).

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

Table B-2: Influence Factors of Prudence (Crainich-Eeckhoudt Measure)

	Dependent Variable: Prudence							
	(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. 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 fulltime', and 'Catholic' for religion. For 24 participants, some demographic information has been imputed with 0, the variable's mean value, and the 'other' category for binary, continuous and categorical variables, respectively. We controlled for imputation with indicator variables. See Table S-3 in the supplementary material on our website for regression results excluding participants that reported problems with handling their tablets during our study. 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).

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

Table B-3: Influence Factors of Temperance (Denuit-Eeckhoudt Measure)

	Dependent Variable: Temperance							
	(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. 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 fulltime', and 'Catholic' for religion. For 24 participants, some demographic information has been imputed with 0, the variable's mean value, and the 'other' category for binary, continuous and categorical variables, respectively. We controlled for imputation with indicator variables. See Table S-4 in the supplementary material on our website for regression results excluding participants that reported problems with handling their tablets during our study. 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).

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

Table B-4: DOSPERT (Adapted)

	Dependent Variable: DOSPERT (Adapted)			
	(1)	(2)	(3)	(4)
Risk aversion (AP)		−0.283**** (0.064)	−0.259**** (0.065)	−0.288**** (0.062)
Prudence	−0.098* (0.052)	−0.098** (0.049)	−0.083* (0.049)	−0.087* (0.045)
Temperance		−0.047 (0.047)	−0.032 (0.044)	−0.023 (0.039)
Impatience			0.165**** (0.044)	0.107*** (0.033)
Cognitive ability				−0.081**** (0.025)
Age (in years)				0.001 (0.018)
Female (=1)				−0.082 (0.062)
Pocket money per week				0.010**** (0.003)
Math grade				−0.132*** (0.042)
German grade				−0.127*** (0.048)
Number of siblings				0.057* (0.031)
Migration background (=1)				0.175** (0.083)
Education mother: A-Levels (=1)				0.004 (0.047)
Education father: A-Levels (=1)				−0.069 (0.098)
BMI				0.019 (0.012)
Parents Occupation				
Full-time and part-time				−0.221** (0.087)
One full-time				−0.067 (0.151)
Don't work/other regularity				−0.219 (0.152)
Religion				
Protestant				−0.084 (0.066)
Other or no religion				0.021 (0.073)
R^2	0.01	0.03	0.06	0.18
Observations	658	658	658	658

Notes: Positive coefficients imply increasing general risk taking behavior. Experimental risk and time measures are 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 fulltime', 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 C.1 of the questionnaire in Online Appendix C. Bootstrapped standard errors (1000 repetitions) clustered at the session level in parentheses.
****/***/**/* denotes significance at the 0.1 / 1 / 5 / 10 percent level.

Table B-5: Unhealthy Behavior

	Dependent Variable: Unhealthy Behavior							
	(1)		(2)		(3)		(4)	
Risk aversion (AP)			0.030	(0.088)	0.053	(0.087)	0.024	(0.086)
Prudence	−0.136****	(0.035)	−0.136****	(0.034)	−0.123****	(0.034)	−0.140****	(0.025)
Temperance			−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. 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 fulltime', 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 C.2 of the questionnaire in Online Appendix C. Bootstrapped standard errors (1000 repetitions) clustered at the session level in parentheses.
****/***/**/* denotes significance at the 0.1 / 1 / 5 / 10 percent level.

Table B-6: Addictive behavior

	Dependent Variable: Addictive Behavior			
	(1)	(2)	(3)	(4)
Risk aversion (AP)		0.016 (0.087)	0.039 (0.086)	0.007 (0.087)
Prudence	−0.142**** (0.034)	−0.142**** (0.034)	−0.129**** (0.033)	−0.146**** (0.025)
Temperance		−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. 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 fulltime', 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 C.2 of the questionnaire in Online Appendix C. Bootstrapped standard errors (1000 repetitions) clustered at the session level in parentheses.
****/***/**/* denotes significance at the 0.1 / 1 / 5 / 10 percent level.

Table B-7: Addictive Usage of Smarthpone and Social Media

	Dependent Variable: Smartphone Addiction			
	(1)	(2)	(3)	(4)
Risk aversion (AP)		0.002 (0.089)	0.024 (0.087)	−0.023 (0.086)
Prudence	−0.152**** (0.033)	−0.152**** (0.033)	−0.139**** (0.032)	−0.160**** (0.025)
Temperance		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. 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 fulltime', 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 these indices are listed in Section C.2 of the questionnaire in Online Appendix C. Bootstrapped standard errors (1000 repetitions) clustered at the session level in parentheses.
****/***/**/* denotes significance at the 0.1 / 1 / 5 / 10 percent level.

Table B-8: Prevention (Short-Term)

	Dependent Variable: Prevention (Short Term)			
	(1)	(2)	(3)	(4)
Risk aversion (AP)		0.168** (0.079)	0.178** (0.078)	0.120* (0.063)
Prudence	−0.117*** (0.036)	−0.117*** (0.037)	−0.110*** (0.038)	−0.119**** (0.033)
Temperance		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. 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 fulltime', 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 C.3 in the questionnaire in Online Appendix C. Bootstrapped standard errors (1000 repetitions) clustered at the session level in parentheses.
****/***/**/* denotes significance at the 0.1 / 1 / 5 / 10 percent level.

Table B-9: Prevention (Long-Term)

	Dependent Variable: Prevention (Long Term)			
	(1)	(2)	(3)	(4)
Risk aversion (AP)		0.245*** (0.077)	0.233*** (0.077)	0.167*** (0.060)
Prudence	0.047 (0.056)	0.047 (0.053)	0.040 (0.054)	0.010 (0.042)
Temperance		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. 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 fulltime', 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 C.3 in the questionnaire in Online Appendix C. Bootstrapped standard errors (1000 repetitions) clustered at the session level in parentheses.
****/***/**/* denotes significance at the 0.1 / 1 / 5 / 10 percent level.

Table B-10: Eco-friendly behavior

	Dependent Variable: Eco-friendly Behavior							
	(1)		(2)		(3)		(4)	
Risk aversion (AP)			0.208**	(0.100)	0.188**	(0.095)	0.201**	(0.081)
Prudence	0.044	(0.035)	0.044	(0.036)	0.031	(0.034)	0.022	(0.027)
Temperance			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. 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 fulltime', 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 C.3 in the questionnaire in Online Appendix C. Bootstrapped standard errors (1000 repetitions) clustered at the session level in parentheses.
****/***/**/* denotes significance at the 0.1 / 1 / 5 / 10 percent level.

Table B-11: Planning Behavior

	Dependent Variable: Cautious Planning							
	(1)		(2)		(3)		(4)	
Risk aversion (AP)			−0.033	(0.048)	−0.034	(0.048)	−0.063	(0.050)
Prudence	0.040	(0.047)	0.040	(0.045)	0.040	(0.043)	0.028	(0.051)
Temperance			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. 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 fulltime', 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 C.4 in the questionnaire in Online Appendix C. Bootstrapped standard errors (1000 repetitions) clustered at the session level in parentheses.
****/***/**/* denotes significance at the 0.1 / 1 / 5 / 10 percent level.

Table B-12: Preference for Competitive Income

	Dependent Variable: Preference for Competitive Income			
	(1)	(2)	(3)	(4)
Risk aversion (AP)	−0.094**** (0.023)	−0.093**** (0.023)	−0.096**** (0.023)	−0.070*** (0.023)
Prudence		0.052** (0.026)	0.052** (0.026)	0.054** (0.024)
Temperance		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. 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 fulltime', 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 C.5 in the questionnaire in Online Appendix C. Bootstrapped standard errors (1000 repetitions) clustered at the session level in parentheses.
****/***/**/* denotes significance at the 0.1 / 1 / 5 / 10 percent level.

Table B-13: Saving (with Debt)

	Dependent Variable: Saving (with Debt)			
	(1)	(2)	(3)	(4)
Risk aversion (AP)		0.093 (0.114)	0.060 (0.111)	0.096 (0.105)
Prudence	0.060* (0.031)	0.060** (0.030)	0.040 (0.030)	0.058* (0.034)
Temperance		0.110** (0.048)	0.090* (0.047)	0.087** (0.035)
Impatience			−0.222*** (0.034)	−0.197*** (0.032)
Pocket money risky + (=1)				0.063 (0.044)
Pocket money risky − (=1)				−0.037 (0.133)
Earnings side job stable (=1)				0.099** (0.046)
Earnings side job per week				0.001 (0.002)
Earnings side job self-det. (=1)				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. 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 fulltime', 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 C.6 of the questionnaire in Online Appendix C. Bootstrapped standard errors (1000 repetitions) clustered at the session level in parentheses.
****/***/**/* denotes significance at the 0.1 / 1 / 5 / 10 percent level.

Table B-14: Risky Investment

	Dependent Variable: Risky Investment			
	(1)	(2)	(3)	(4)
Risk aversion (AP)		−0.161** (0.074)	−0.159** (0.075)	−0.082 (0.069)
Prudence		−0.103** (0.048)	−0.102** (0.048)	−0.055 (0.042)
Temperance	−0.064*** (0.023)	−0.064*** (0.024)	−0.062*** (0.023)	−0.037** (0.015)
Impatience			0.013 (0.025)	0.016 (0.026)
Cognitive ability				−0.052 (0.035)
Age (in years)				−0.000 (0.015)
Female (=1)				−0.521*** (0.061)
Pocket money per week				0.008** (0.004)
Math grade				0.075** (0.031)
German grade				−0.026 (0.033)
Number of siblings				−0.046** (0.020)
Migration background (=1)				0.002 (0.060)
Education mother: A-Levels (=1)				0.212*** (0.065)
Education father: A-Levels (=1)				0.189** (0.085)
BMI				0.006 (0.010)
Parents Occupation				
Full-time and part-time				0.096 (0.080)
One full-time				−0.140 (0.119)
Don't work/other regularity				−0.152 (0.097)
Religion				
Protestant				−0.042 (0.078)
Other or no religion				−0.223*** (0.051)
R^2	0.00	0.02	0.02	0.17
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. 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 fulltime', 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 C.6 of the questionnaire in Online Appendix C. Bootstrapped standard errors (1000 repetitions) clustered at the session level in parentheses.
****/***/**/* denotes significance at the 0.1 / 1 / 5 / 10 percent level.

Table B-15: Financial Insurance Demand

	Dependent Variable: Insurance Demand							
	(1)		(2)		(3)		(4)	
Risk aversion (AP)			−0.064	(0.075)	−0.061	(0.076)	−0.046	(0.066)
Prudence	−0.062**	(0.025)	−0.062***	(0.024)	−0.060**	(0.026)	−0.038	(0.036)
Temperance			−0.010	(0.070)	−0.008	(0.071)	0.002	(0.071)
Impatience					0.021	(0.041)	−0.000	(0.039)
Cognitive ability							0.021	(0.027)
Age (in years)							−0.010	(0.015)
Female (=1)							−0.124*	(0.066)
Pocket money per week							0.008***	(0.003)
Math grade							−0.029	(0.039)
German grade							0.029	(0.055)
Number of siblings							−0.044**	(0.022)
Migration background (=1)							0.180***	(0.050)
Education mother: A-Levels (=1)							−0.018	(0.054)
Education father: A-Levels (=1)							−0.006	(0.067)
BMI							0.015	(0.012)
Parents Occupation								
Full-time and part-time							−0.118	(0.088)
One full-time							−0.149	(0.122)
Don't work/other regularity							−0.385***	(0.098)
Religion								
Protestant							−0.076	(0.093)
Other or no religion							−0.141	(0.105)
R^2	0.00		0.00		0.01		0.06	
Observations	658		658		658		658	

Notes: Positive coefficients imply increasing likelihood to possess an insurance. Risk and time measures are 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 fulltime', 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 C.6 of the questionnaire in Online Appendix C. Bootstrapped standard errors (1000 repetitions) clustered at the session level in parentheses.
****/***/**/* denotes significance at the 0.1 / 1 / 5 / 10 percent level.

Table B-16: Regression Results Using Correlated Risk Measures

(a) General Survey Questions/Questionnaires on General Risk Taking and Patience (see Table 5)

	Risk tolerance (Survey)		DOSPERT (adapted)		Patience (Survey)		Prudence (à la E & S, 2006)	
Risk aversion (AP)	−1.063****	(0.241)	−0.288****	(0.062)	0.016	(0.187)	0.004	(0.070)
Prudence	−0.341***	(0.119)	−0.075	(0.054)	−0.095	(0.096)	0.147****	(0.041)
Temperance	0.704***	(0.253)	0.228****	(0.066)	0.053	(0.152)	−0.068	(0.083)
Impatience	0.020	(0.099)	0.107***	(0.033)	−0.474****	(0.086)	0.065*	(0.039)
Other Factors	13		13		13		13	
R^2	0.13		0.18		0.12		0.05	
Observations	653		658		653		658	

(b) Health-Related Behavior (see Table 6)

	Unhealthy Behavior		Addictive Behavior		Smartphone Addiction	
Risk aversion (AP)	0.024	(0.086)	0.007	(0.087)	−0.023	(0.086)
Prudence	−0.133****	(0.034)	−0.145****	(0.034)	−0.171****	(0.032)
Temperance	−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	

(c) Prevention and Environmentally-Friendly Behavior (see Table 7)

	Prevention (Short Term)		Prevention (Long Term)		Eco-friendly behavior	
Risk aversion (AP)	0.120*	(0.063)	0.167***	(0.060)	0.201**	(0.081)
Prudence	−0.130****	(0.034)	−0.040	(0.049)	0.001	(0.047)
Temperance	−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 (see Table 8)

	Cautious Planning		Preference for Competitive Income	
Risk aversion (AP)	−0.063	(0.050)	−0.124***	(0.042)
Prudence	−0.084	(0.063)	0.046*	(0.027)
Temperance	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	

(e) Financial Decision Making (see Table 9)

	Saving (with Debt)		Risky Investment		Fin. Insurance	
Risk aversion (AP)	0.096	(0.105)	−0.082	(0.069)	−0.046	(0.066)
Prudence	0.002	(0.040)	−0.055	(0.042)	−0.039	(0.066)
Temperance	0.003	(0.105)	0.071	(0.070)	0.042	(0.106)
Impatience	−0.197****	(0.032)	0.016	(0.026)	−0.000	(0.039)
Other Factors	18		13		13	
R^2	0.19		0.17		0.06	
Observations	658		658		658	

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 and time measures. Consequently, coefficient estimates might be biased and standard errors incorrect. See Tables 5 to 9 for results from orthogonalized risk measures and 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 B-17: Selected Variables Using Lasso

	Risk Aversion	Prudence	Temperance	Impatience
Risk tolerance (Survey)	x	x		
DOSPERT (adapted)	x	x	x	x
Patience (Survey)				x
Prudence (à la E & S, 2006)		x		
Saving (with Debt)	x			x
Risky Investment	x	x		
Fin. Insurance	x	x		
Unhealthy Behavior		x	x	x
Addictive Behavior		x		x
Smartphone Addiction		x		x
Prevention (Short Term)		x		x
Prevention (Long Term)	x	x		x
Eco-friendly behavior	x			x
Preference for Competitive Income	x			
Cautious Planning			x	

Notes: An "x" in the row of an outcome indicates that the respective risk/time measure (see table head) is selected as independent variable for a linear regression of the outcome according to the Lasso approach (Tibshirani, 1996). The method was performed for a linear model using the built-in Stata command with default parameters. See Tables 5 to 9 for results from regressions with all risk and time measures (orthogonalized) and additional notes on the respective models. Note that inclusion in the model is not equivalent to significance. It rather suggests that inclusion of the variable increases the model's predictive quality.

Table B-18: Financial Decision Making - Subsample of Adolescents in Grades 10 and 12 (vast majority 15 years and older)

	Risky Investment		Insurance Demand	
Risk aversion (AP)	0.023	(0.092)	−0.002	(0.084)
Prudence	−0.019	(0.045)	0.028	(0.051)
Temperance	−0.047**	(0.022)	−0.018	(0.128)
Impatience	0.041	(0.031)	0.017	(0.061)
Other Factors	13		13	
R^2	0.25		0.06	
Observations	336		336	

Notes: Positive coefficients imply increasing likelihood to invest in risky assets or possess an insurance. Risk and time measures in standard deviations. Non-orthogonalized measure: Temperance (Risky Investment) and prudence (Insurance Demand). Outcome indices are standard normalized and formed by using PCA weights. Questions included in these indices are listed in Section C.6 of Online Appendix C. Influence factors controlled for are the same as in Section 3. See Table 9 for an overview table of regressions using the full sample, as well as Tables B-14 and B-15 for detailed regression results using the full sample. Bootstrapped standard errors (1000 repetitions) clustered at the session level in parentheses.
****/***/**/* denotes significance at the 0.1 / 1 / 5 / 10 percent level.

Table B-19: Significance and Sign of Risk Aversion, Conditional on Estimation With or Without Higher Order Risk Preferences - ORIV Estimation

Coefficient of risk aversion (columns [1] and [2] show estimated coefficients)	[1] EXCLUDING prudence and temperance	[2] INCLUDING prudence and temperance	[3] Change whether risk aversion is significant or not	[4] Change in sign of risk aversion
Health-Related Behavior (Table 6)				
Unhealthy Behavior	−0.064*	0.043	YES	YES
Addictive Behavior	−0.068*	0.028	YES	YES
Smartphone Addiction	−0.072*	0.011	YES	YES
Prevention and Eco-Friendly Behavior (Table 7)				
Prevention (Short Term)	−0.005	0.135**	YES	YES
Prevention (Long Term)	0.123****	0.204****		
Eco-friendly Behavior	0.100**	0.189***		
Planning Behavior and Preference for Competitive Income (Table 8)				
Cautious Planning	0.088**	−0.038	YES	YES
Preference for Competitive Income	−0.085****	−0.119****		
Financial Decision Making (Table 9)				
Saving (with Debt)	0.104**	0.076	YES	
Risky Investment	−0.062*	−0.084	YES	
Insurance Demand	−0.046	−0.070		
Total number of changes (out of 11 regressions)			7 (out of 11)	5 (out of 11)

Notes: Positive coefficients imply increasing preference for the respective behavior. Risk and time measures in standard deviations. Column [1] shows regression coefficients from the models shown in Tables 6 to 9 when excluding higher order risk preferences; column [2] shows the regression coefficients from the models when including higher order risk preferences. See Tables 6 to 9 for these results and additional notes on the respective models. See *Measurement Error* in Section 2.3 for implementation details. Bootstrapped standard errors (1000 repetitions) clustered at the session level in parentheses.
****/***/**/* denotes significance at the 0.1 / 1 / 5 / 10 percent level.

C Online Appendix: Questionnaire

Answer possibilities are listed 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 occassion/never", or "very often/never", depending on the item.

C.1 General Risk Taking and Patience Behavior

One-Item Survey Questions

- Compared to others, are you generally willing to renounce something to benefit from that in the future? Or are you, compared to others, not willing to do so? Please tick one of the boxes on the scale, whereby the value 0 means: "not at all willing to do so", and the value 10 means: "very willing to do so". With the values in between you can graduate your assessment. [0-10]
- 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]
- In general, are you also ready to take risks even when something really bad can happen or do you try to avoid risks like that? 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 occassion)]
- [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 neighbour, 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]

General Impatience Scale

- I tend to procrastinate activities. [0-3]
- I always do my homework as early as possible. [0-3]
- Playing an instrument (e.g., in music school, band, at home...) ["never", "less than 1x per month", "1x per month", "2x per month", "1x per week", "2x per week", "more than 2x per week"]

C.2 Health Related Behavior

- 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, regulary (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]

C.3 General Prevention and Eco-Friendly Behavior

General Prevention (Short Term)

- 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)

- 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

- 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 noone 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 seperate 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]

C.4 Planning Behavior

- 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"]

C.5 Preference for Competitive Income

- 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 allot 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]

C.6 Financial Decision Making

Saving w./ Debt

- 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 credit card? [y, n]

Risky Investment

- 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

- 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]

C.7 Demographic Information

- 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]
- How often do you attend religious festivities (e.g. mass, mosque attendance, ...) ["never", "less than 1x per month", "1x per month", "2x per month", "1x per week", "2x per week", "more than 2x per week"]
- 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 regulary 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]