Experimental Economics (2026), 1–16
doi:10.1017/eec.2026.10041

economic science association

CAMBRIDGE UNIVERSITY PRESS

ORIGINAL PAPER

Eliciting time, risk, and social preferences in children: a validated survey

Helene Willadsen[1] and Marco Piovesan[2]

[1]Department of Economics, University of Copenhagen and the National Research Center for the Working Environment, Copenhagen, Denmark
[2]Department of Economics, University of Verona, Verona, Italy
Corresponding author: Helene Willadsen; Email: hew@nfa.dk

(Received 2 May 2024; revised 23 April 2025; accepted 19 January 2026)

Abstract
We develop and validate a survey instrument to elicit six key economic preferences in children: undefined time preferences, risk preferences, altruism, positive reciprocity, negative reciprocity, and trust. The survey was administered to a sample of 339 nine-year-old children, for whom we also collected behavioral data through incentivized choice experiments targeting the same preferences. Our econometric analysis allows us to identify a set of 14 survey items that best predict children's experimental behavior. For each preference, we also compare the predictive power of this 14-item validated survey to a shorter 9-item self-evaluation version. Our results demonstrate that these surveys provide a simple and reliable tool for measuring individual preferences in children – enabling researchers to account for heterogeneity when designing and evaluating policies targeting younger populations.

Keywords: Altruism; experiments for children; reciprocity; risk preferences; survey; time preferences

JEL Codes: C70; C90; C83; J24

1. Introduction

Any decision we make in daily life involves a combination of risk (our willingness to take a gamble in exchange for a possible reward), time (how we value the future relative to the present), and social preferences (how much we care about the wellbeing of others). Understanding the heterogeneity of these preferences is crucial to explaining differences in life outcomes and the persistent inequalities observed across individuals and societies. A growing body of research suggests that these undefined preferences, when formed in childhood, are not only predictive of but may also causally influence long-term outcomes such as educational attainment, financial stability, health behaviors, and prosocial engagement (Alan & Ertac, 2018; Sutter et al., 2013; Mischel et al., 1989).

To investigate how such preferences develop during childhood and adolescence, economists have traditionally relied on incentivized experimental tasks, designed to elicit intertemporal, risk-related, and prosocial behaviors under controlled conditions (for a survey, see Sutter et al., 2019; List et al., 2021). A substantial literature documents how these preferences evolve across childhood and adolescence. Patience increases with age (Sutter et al., 2015) and children from low socioeconomic status (SES) families make more impatient choices (Falk et al., 2021b). Older children are more generous than the younger ones (Fehr et al., 2008, 2013) and prosocial behaviors such as fairness and altruism become more structured and norm-based during adolescence (Almås et al., 2010). Risk aversion decreases or stabilizes depending on context, and, on average, girls are more risk averse than boys

(Sutter et al., 2019). Gender differences emerge around puberty – particularly in risk and social preferences – while SES and cultural background consistently predict variation in time and social preferences from early childhood (Fehr et al., 2008; Sutter et al., 2013; Blake et al., 2015; Andreoni et al., 2020).

While incentivized experiments offer incentive compatibility and fine control, they are often time-consuming, costly, and logistically demanding, especially in large-scale or repeated settings. In contrast, Psychology has long used questionnaires for eliciting personality traits (for example, Ashton & Lee, 2009; McCrae & Costa, 2008) and risk taking (for example, Gullone et al., 2000). Another example is the Strength and Difficulties Questionnaire (Goodman, 1997), which elicits prosociality with self-rated questions like '*I usually share with others, for example, CDs, games, food*', that resembles economists' understanding of social preferences. However, a major concern remains whether surveys capture the same underlying preferences as incentivized tasks. Among adults, recent advances have validated short surveys that correlate well with experimental measures (Enke et al., 2022; Falk et al., 2023), and such tools have been used to map global distributions of preferences (Falk et al., 2018). In children and adolescents, however, the evidence is more limited and the correlation between survey and behavioral measures appears much weaker, raising concerns about measurement equivalence across age groups. For instance, Samek et al. (2021) elicited risk and time preferences using both experiments and survey questions and discovered only a modest correlation coefficient of 0.09 between experimentally elicited time preferences and survey-based measures of the same among adolescents, and with a slightly higher coefficient of 0.11 for risk preferences. These correlations are lower than for adults. Dohmen et al. (2011) report a more substantial correlation coefficient of 0.26. Falk et al. (2023) report correlations of 0.37 for negative reciprocity, 0.41 for risk preferences, 0.59 for time preferences and 0.67 for trust. This discrepancy is important to consider when relying on survey-based tools for measuring preferences in children, particularly in large-scale or policy-oriented studies where experimental elicitation is not feasible.

Our study addresses this gap by developing, testing, and validating a novel and scalable survey instrument for measuring time, risk, and social preferences in children and adolescents. We began by designing candidate questions, spanning domains such as time, risk and social Preferences – trust, positive and negative reciprocity. These questions were asked in a sample of children aged 9 years: 160 children answered a survey measuring their preferences regarding risk, time, and altruism and 179 answered a survey focusing on positive reciprocity, negative reciprocity, and trust. Either one week prior to or following the survey, these children participated in a series of incentivized choice experiments, targeting the exact same set of preferences. By comparing children's survey responses with their behavior in the experimental tasks, we identified the subset of items that best predicted choices in incentivized choice experiments (in terms of maximizing R^2). To strike a balance between explanatory power and brevity, we evaluate all possible survey combinations using both the Bayesian Information Criterion (BIC) and Root Mean Squared Error (RMSE). While RMSE tends to favor larger models, we find that BIC selects fewer survey items with only marginal losses in predictive accuracy – making it the preferred approach for constructing a brief and practical instrument. Our final instrument consists of just 14 survey items yet explains between 11.8% and 34.8% of the variance in behavior across domains, making it a practical and scalable alternative to experimental elicitation. Our findings indicate that hypothetical versions of incentivized tasks outperform general self-evaluation items in predicting children's behavior. Items that closely resemble the experimental setup or describe concrete, age-relevant situations – such as decisions involving toys, peers, or classroom settings – tend to yield better predictive accuracy. This underscores the importance of contextual relevance in item design and opens new avenues for research on the effectiveness of different item types across developmental stages and cultural contexts.

To validate our selection, we conducted an out-of-sample prediction exercise with a separate group of children for risk and time preferences as well as altruism. This new sample confirms that the

https://doi.org/10.1017/eec.2026.10041 Published online by Cambridge University Press

selected items retained substantial predictive accuracy in a new population aged 9–16 years old. With this older sample we also investigate self-evaluation questions, closer to what has been used with adults in Falk et al. (2023), and we show that these are most useful for eliciting time preferences.

Our contributions are fourfold. First, we construct a novel set of over 100 survey items explicitly designed to measure economic preferences in children and adolescents. These items were carefully tailored to be developmentally appropriate, contextually relevant, and suitable for diverse research settings. The full item pool – along with its child-friendly graphical design – serves as a flexible toolbox for researchers aiming to study preference heterogeneity across populations and institutional contexts. Second, we validate a 14-item short-form survey that reliably approximates behavior in incentivized choice experiments. This is the first validated instrument of its kind that spans risk, time, and social preferences for children and adolescents, and it is comparable in structure and performance to established adult surveys (e.g., Cappelen et al., 2025; Falk et al., 2023). Our results enable researchers to incorporate robust measures of economic preferences into large-scale, cross-national, and longitudinal studies, even when experimental methods are not feasible. Third, we design the survey to be accessible and practical for use in schools, families, and field research. All questions are concise, easy to understand, and accompanied by child-friendly illustrations. The instrument can be administered with minimal supervision and is suitable for both individual and group settings, making it well-suited for use with younger populations. Fourth, we use a transparent and replicable validation procedure. Researchers can apply the same method to identify and validate the best-performing subset of items within their own samples – across different age groups, cultural settings, or socioeconomic contexts. This adaptability makes the instrument and its underlying methodology broadly applicable. By bridging survey and experimental methods in childhood, our study lays the groundwork for a new, scalable approach to measuring economic preferences in developmentally and policy-relevant contexts. Since early-formed preferences play a key role in shaping life trajectories, accurately identifying them is critical for designing effective interventions – whether in education (e.g., improving time management or fostering self-regulation), health (e.g., promoting long-term planning for nutrition and exercise), or financial literacy (e.g., encouraging saving and delayed gratification).

The remainder of the paper is structured as follows: Section 2 reviews the literature on preference elicitation in youth. Section 3 describes our experimental design and survey construction. Section 4 presents the validation strategy and results. Section 5 discusses implications and limitations. Section 6 concludes.

2. Data

We collected our data by visiting 21 third-grade classes in nine public schools located just outside of Copenhagen. Participation in the study was voluntary, and therefore we could not obtain a representative sample. The participating schools are in seven municipalities: in four of them, the average household income is below the national average by 1–10%, while in the remaining three it is higher than the national average by 1–84%. Each class was visited twice, one week apart, for a total of approximately 180 minutes. We contacted 484 parents and received consent from 407 of them (84%). Our final sample consists of 339 observations, as 68 children were absent during one of the two visits. In total, 160 children participated in incentivized choice experiments and answered a 59-item survey on risk preferences, time preferences, and altruism, and 179 children answered a 50-item survey and incentivized experiments measuring positive reciprocity, negative reciprocity, and trust. To avoid spurious interdependencies, we elicited these preferences with the incentivized choice experiments and with the hypothetical questions at two separate visits, one week apart. Given the bifurcated structure of the study, children were exposed to a maximum of 36 survey questions in any single session.

During the activities, children earned tokens that could be exchanged for prizes such as toys or school materials. We paid out one decision for each experiment to ensure a clear distinction between the incentivized experimental games and the non-incentivized hypothetical survey items.

https://doi.org/10.1017/eec.2026.10041 Published online by Cambridge University Press

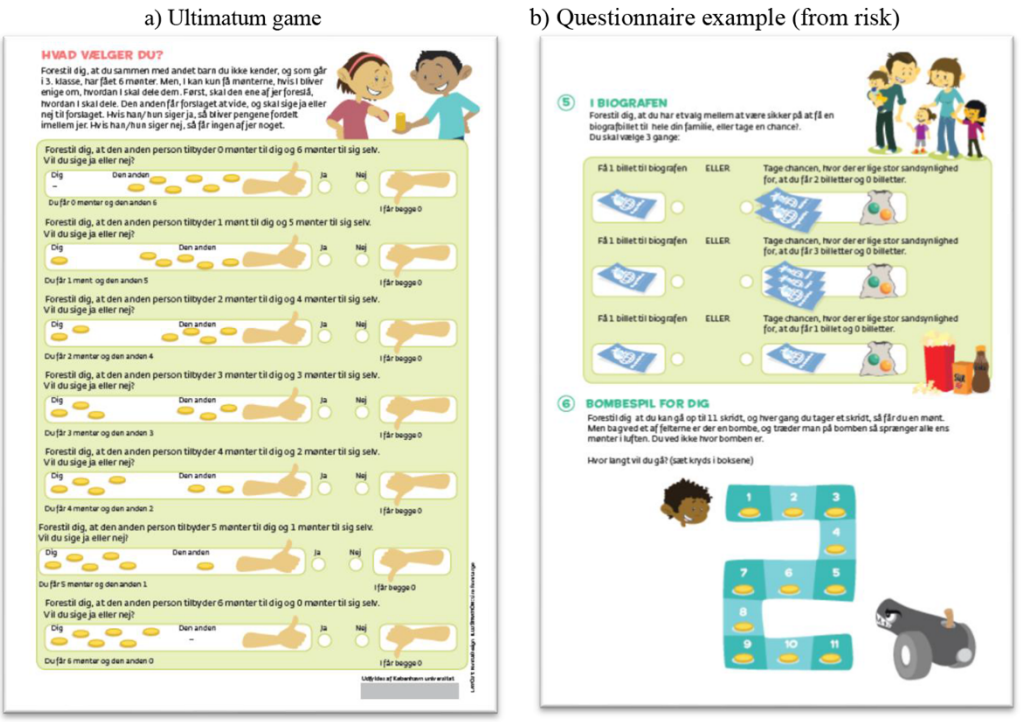

Fig. 1 Examples of decision sheets: (a) Ultimatum game. (b) Questionnaire example (from risk)

Emphasizing the role of incentives in the experimental setting allowed us to clearly communicate the difference between the incentivized and non-incentivized items.

We explained to the children that earning more tokens would result in receiving more gifts. However, we deliberately did not disclose the exact exchange rate (1 token = 2 DKK) to minimize social comparison and avoid strategic behavior. On average, children earned 9.5 tokens – 7.5 from the incentivized choice experiments and 2 tokens as a participation fee. We also took great care in designing decision sheets, collaborating with a professional graphic designer. Figure 1 reports examples of decision sheets. Appendix B describes our procedures in detail, including our script and decision sheets.

Our data have two main components: children's decisions in the six incentivized choice experiments, and their answers to the 109 hypothetical questions we designed to elicit preferences.

2.1. Incentivized choice experiments

We elicit *time preferences* by asking children to choose between an amount today (4 tokens) or a larger amount (5, 6, 7, or 8 tokens) in a week.[1] We randomly pick one of these four choices to determine the payoff.[2] We use the answers to construct a measure of *patience* using the switching point from today to the delayed payoff. For instance, when a child chooses 4 tokens today rather than 5 in a week, but she chooses 6 tokens in a week instead of 4 today, we assign a patience value of 3. For consistent children,

[1]The weekly rate of return in our incentivized task varies from a weekly interest rate of $r = 0.25$ – $r = 1$ (a doubling of the here and now endowment) similar to what has been used in Alan and Ertac (2018) and comparable to Angerer et al. (2015). In both Alan and Ertac (2018) and Angerer et al. (2015), the subjects are approximately the same age.

[2]To illustrate the random draw to children we gave each decision a color and let the teacher – in front of children – draw out the color from a bag of colored bricks to determine the decision to be paid out, see also Appendix B.

https://doi.org/10.1017/eec.2026.10041 Published online by Cambridge University Press

this patience measure equals the number of patient decisions. In the case of multiple switching points (39 observations), we took the average of the switching points.[3] Appendix A, Table A.14 reports our main results restricting the sample to consistent individuals and, alternatively, using the number of patient choices as a robustness measure of patience.

We elicit *risk preferences* by asking children to choose between a safe option (an increasing number of tokens from 1 and 11) or a gamble (0 or 10 tokens with the same probability, i.e., the expected value of 5). The gamble was illustrated as a draw between two colors (green or orange), and children were asked to choose which color would represent the winning outcome. At the end of the experiment, we randomly pick one of these 11 choices to determine the payoff, and the winning color.[4] Then, we construct a measure of risk tolerance using the switching point from the safe option to the gamble. For instance, when a child chooses 7 tokens for sure rather than the gamble but chooses the gamble rather than 6 tokens for sure, we assign a risk-loving value of 6. For consistent children, this risk measure equals the number of risky decisions. As before, in the case of multiple switching points (39 observations), we took the average of the switching points. In Appendix A (Table A.15), we report our main results using only consistent individuals, as well as alternative measures of risk preferences based on the number of risky choices and the first switching point.

We elicit *altruism* by using a dictator game in which children have to allocate 6 tokens between themselves or another unknown child. Our measure of *altruism* corresponds to the number of tokens allocated to the other child. For instance, when a child gives 3 tokens, we assign a value of 3.

We elicit *negative reciprocity* by using an ultimatum game: the first mover has 6 tokens to allocate, and the second mover can choose 'accept' or 'reject' the allocation. We ask children to play as the second mover first and state each possible allocation if they want to 'accept' or 'reject'.[5] Then, children play as the first mover. At the end of the experiment, we pair each child with an anonymous third grader from another school and use children's decisions to determine their payoffs. Our measure of *negative reciprocity* corresponds to the lowest number of accepted tokens. For instance, if a child declines the offer '0 to self; 6 to the other' but accepts the offer '1 to self; 5 to the other', we assign a value of 1.

We elicit *positive reciprocity* and *trust* by using a sequential trust game: the first mover has 4 tokens, and she has to choose how many of these tokens she wants to send to the second mover. Then, tokens are tripled and are sent to the second mover. Finally, the second mover can choose how many of these tokens she wants to send back to the first mover. We ask children to play first as the second mover and state how many tokens they want to send back if the first mover sends 1, 2, 3, or 4 tokens. Then, children play as the first mover and choose how many tokens they want to send to the second mover.[6] At the end of the experiment, we pair each child with an anonymous third grader from another school and use children's decisions to determine their payoffs. We construct a measure of *positive reciprocity* using the average number of returned tokens as the second mover. For instance, if a child never sends back any tokens, we assign a value of 0. If a child always sends back all the tokens, we assign a value of $(3 + 6 + 9 + 12)/4 = 7.5$. We construct a measure of *trust* using the amount sent as the first mover. For instance, if a child sends 2 tokens as the first mover, we assign a value of 2.

[3] 24% of the sample had multiple switching points. This is similar to Khachatryan et al. (2015) in which 19% of children have multiple switching points. Willadsen et al. (2024) investigates the determinants of inconsistency in children choice experiments, particularly the role of cognitive abilities.

[4] To illustrate the random draw to children we let the teacher – in front of children – draw out a numbered card from a deck numbered 1-11 to determine the decision row to be paid out. Similarly, the teacher determined the winning color, see also Appendix B.

[5] We use the strategy method: subjects must make contingent decisions for all nodes at which they may have to play. This approach enhances participants' understanding of the game structure. This design is common in experimental economics, and Brandts and Charness (2011) provide a comprehensive review of the literature showing that the strategy method induces results similar to those induced by the direct-response method.

[6] We again use the *strategy method*. See footnote 2.

https://doi.org/10.1017/eec.2026.10041 Published online by Cambridge University Press

For both the ultimatum game and the trust game, we informed children that either their decision as first mover or as second mover would be selected for payment. This design introduces role uncertainty as children did not know which decision would be payoff relevant. Previous work has shown that role uncertainty may shift behavior toward more altruistic and less spiteful behavior (Iriberri & Rey-Biel, 2011). For context, we therefore relate our results to previous experimental studies with children in the results section.

2.2. Hypothetical questions

Our survey includes 109 items: 23 to elicit time preferences, 16 for risk preferences, 20 for altruism, 14 for negative reciprocity, 14 for positive reciprocity, and 22 for trust. The development of these items followed a multi-step process. First, we adapted questions from existing instruments used with adults, such as the Preference Survey Module (Falk et al., 2023) and psychological tools like the Strengths and Difficulties Questionnaire (Goodman, 1997). Second, we tailored the items to align with the cognitive and emotional development of children. This involved rephrasing questions to reflect situations that children are likely to encounter in everyday life – such as decisions involving toys, school, or peer interactions – making them more relatable and easier to understand. Questions that referred to adult-specific domains, such as driving, alcohol, sex, or work, were excluded. We also developed hypothetical versions of standard experimental tasks, including dictator games, trust games, and intertemporal choice tasks (e.g., Harbaugh et al., 2003; Sutter et al., 2013). Finally, recognizing that children's behavior is shaped in part by socialization, we included items reflecting adult feedback (e.g., from parents and teachers) as well as questions that varied the target of the behavior – for instance, whether it was directed toward family, friends, or unfamiliar peers – a distinction that is especially relevant in the context of social preferences (Harbaugh et al., 2003).

For each preference, we have four types of questions: qualitative (for instance, *Are you a child who runs a lot of risks, or do you try to avoid risks?*), scenario questions (for instance, *Imagine that you are in a queue. Somebody skips the queue and takes your turn. Will you now tell a teacher instead of taking your turn? YES/NO*), other's view (for instance, *Do your parents tell you to be more cautious? YES/NO*) and the hypothetical version of incentivized choice experiments (for instance, *Now you have to imagine that you have six tokens. You have to choose how many of these tokens you want to keep and how many you want to give to another child. The other child is your age and lives in Denmark, but is not someone you know*). In short, we prepared our questions specifically for children: the content of each question reflects situations children commonly experience in their daily lives, and the graphical layout is child-friendly. All survey questions can be found in Appendix A, Table A.21.

The data has been collected and stored in accordance with the rules set by the General Data Protection Regulation (GDPR) covering the European Union. As a copy of this data exists as personally identifiable, we are unable to anonymize the data as defined in the GDPR by the Danish Data Protection Agency[7] (DPA). We are therefore unable to publicly share the data. Instead, researchers interested in accessing these data must apply alongside the University of Copenhagen for approval from the DPA and adhere to the specified conditions.[8] This process ensures that data sharing aligns with the highest standards of ethical and legal compliance.

3. Results

Figure 2 presents the distribution of children's preferences elicited through the incentivized choice experiments. For each preference, the panel includes a boxplot displaying the mean (×) and median (●). Overall, children's behavior aligns with previous findings in the literature. Panel (a) shows time preferences (patience): on average, children in our sample choose the patient option in 70% of the

[7] www.datatilsynet.dk/english/legislation.
[8] www.datatilsynet.dk/regler-og-vejledning/blanketter-og-skabeloner.

https://doi.org/10.1017/eec.2026.10041 Published online by Cambridge University Press

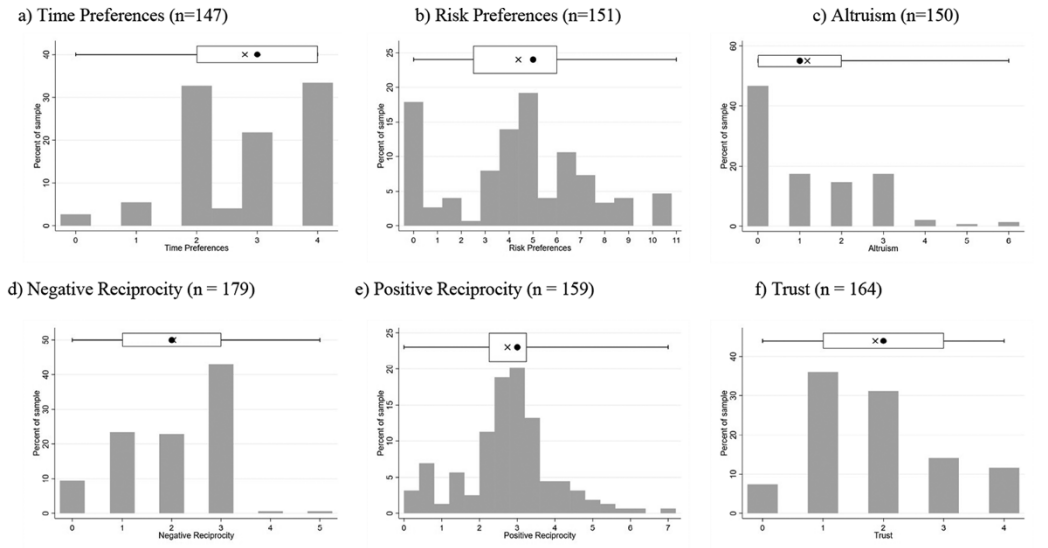

Fig. 2 Distributions of decisions for each type of preference elicited in the incentivized choice experiments. (a) Time Preferences ($n = 147$). (b) Risk Preferences ($n = 151$). (c) Altruism ($n = 150$). (d) Negative Reciprocity ($n = 179$). (e) Positive Reciprocity ($n = 159$). (f) Trust ($n = 164$)

Note: The Figure shows the distribution of decisions in the incentivized choice experiments. The top bar shows the interquartile range, mean (×), and median (●). For time preferences, the *x*-axis indicates *patience*, i.e., the switching point from today to the delayed reward. For risk preferences, the *x*-axis indicates *risk-loving*, i.e., the switching point from the safe option to the gamble. For altruism, the *x*-axis indicates the number of tokens donated. For negative reciprocity, the *x*-axis indicates the lowest number of tokens accepted. For positive reciprocity, the *x*-axis indicates the average number of tokens returned in the trust game. For Trust, the *x*-axis indicates the number of tokens sent by the first mover in the trust game.

decisions, similar to the results in Falk et al. (2021b), but higher than in Angerer et al. (2015), who use a longer delay. In Alan and Ertac (2018), around 40% of participants make 9 or 10 patient choices (out of 10), whereas in our data, 33% of children choose the patient option in all 4 decisions. Panel (b) shows risk preferences (risk tolerance): children choose the risky option in 40% of cases, comparable to the 45% reported by Khachatryan et al. (2015). The estimated average coefficient of risk aversion in our sample is $r = 0.57$, matching the estimate from Sutter et al. (2013). Panel (c) displays altruism: children donate, on average, 19% of their endowment, which is lower than in Harbaugh and Krause (2000), Bettinger and Slonim (2006), and Blake et al. (2015), who report average donations of 29%, 27%, and 25%, respectively. Panel (d) shows negative reciprocity: 10% of children accept all offers, 23% accept offers of at least 1 token, another 23% accept offers of at least 2 tokens, and 43% only accept offers of 3 tokens or more. Murnighan and Saxon (1998) report that 27% of third graders accept offers of 1 token, while Sutter (2007) finds that 35% accept unequal offers of 2 out of 10 tokens. Panel (e) displays positive reciprocity: when the first mover sends 1, 2, 3, or 4 tokens, children return, on average, 1.0, 2.11, 3.32, and 4.51 tokens, respectively. These return amounts are higher and more responsive than the return amounts reported in Harbaugh et al. (2003), where the returns are 1.63, 2.13, 2.45, and 2.24. Panel (f) shows trust: children in our sample send 47% of their endowment, which is notably higher than the 18% reported in Harbaugh et al. (2003), and also higher than the estimates in Sutter and Kocher (2007), who find that second graders send 20% and sixth graders send 36%. In Appendix (Tables A.8–A.13), we show preferences by gender and age (above and below median age). To compare these results with other experimental findings with children we refer to Sutter et al. (2019).

To find the survey questions that provide the best (linear) prediction of individual behavior in the incentivized choice experiments, we use OLS to regress the experimental outcome on various

https://doi.org/10.1017/eec.2026.10041 Published online by Cambridge University Press

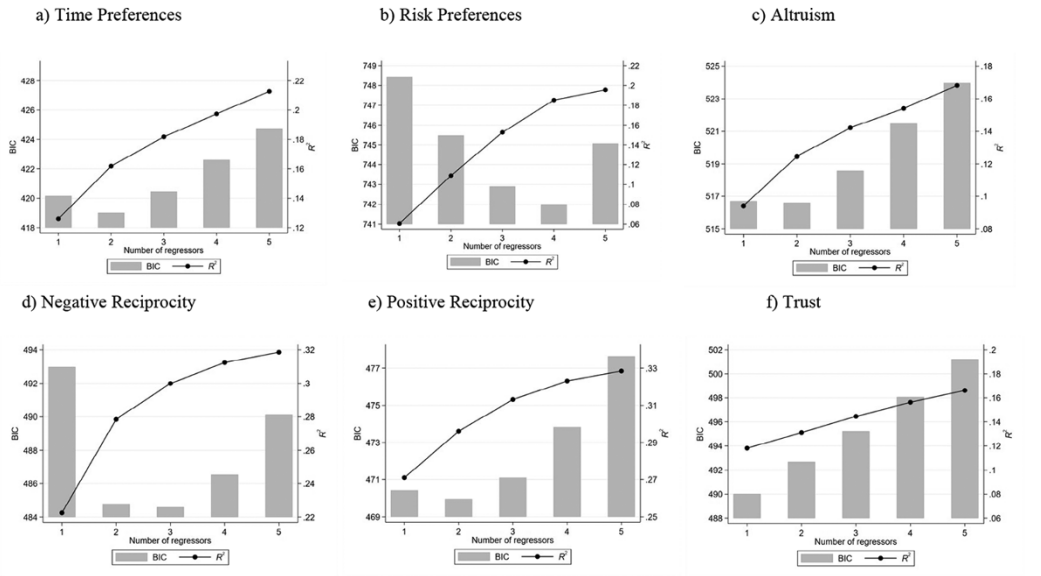

Fig. 3 R^2 and BIC for 1–5 regressors. (a) Time Preferences. (b) Risk Preferences. (c) Altruism. (d) Negative Reciprocity. (e) Positive Reciprocity. (f) Trust

Note: The figure shows the R^2 and the BIC for the models of highest R^2 from OLS-regressions with all combinations of survey items and 1–5 regressors. The preferred models of the ones that are displayed here are those that minimize the BIC penalty term. For comparison with an alternative selection criterion, the Appendix (Table A.4) reports 10-fold cross-validated RMSE for models with 1–5 items. RMSE tends to favor slightly larger models than BIC, but differences are small; we therefore retain the BIC-optimal specifications shown here.

combinations of survey items. Specifically, our analysis considers all possible combinations of 1–5 survey items and uses (adjusted) R^2 to compare models with a similar number of predictors.

To compare models with a different number of items, we consider both the Bayesian Information Criteria (BIC),[9] which penalizes the inclusion of additional regressors, and RMSE from 10-fold cross-validation. Alternative approaches, such as the Obviously Related Instrumental Variables (ORIV) approach (Gillen et al., 2019), could in principle address measurement error[10] but in our design predictive validity is the most informative benchmark.

As per construction, R^2 increases as more items are added to explain the experimental measure, but the BIC does not increase monotonically.[11] Figure 3 shows both the R^2 and BIC for the best combinations of survey items (1-5) for all preferences. The Figure shows, for example for time preferences how R^2 increases from 0.126 when including one survey item, up to 0.213 when including five items. However, the BIC is minimized with just two items, and this model selection criterion implies choosing the specification with the lowest BIC, i.e. two survey items for time preferences. Note that BIC depends on sample size, so we remove all observations with one or more missing values.

The best set of survey items, i.e., the set of items with the lowest BIC, comprises 14 items: 2 for time preferences ($R^2 = 0.162$), 4 for risk preferences ($R^2 = 0.185$), 2 for altruism ($R^2 = 0.124$), 3 for negative reciprocity ($R^2 = 0.300$), 2 for positive reciprocity ($R^2 = 0.348$) and 1 for trust ($R^2 = 0.118$).

To evaluate the models using RMSE, we perform a 10-fold cross-validation procedure, which randomly partitions the data into 10 subsets, and uses 9 of these folds as training data. For each

[9]BIC: $-2\log(L) + k^*\log(n)$, where L is the log-likelihood, k is the number of independent variables and n is the number of observations.

[10]The ORIV approach, however, requires multiple correlated measures of the same latent construct. Since our experimental preferences were elicited with only one incentivized measure per domain, and our survey items were designed for predictive validity rather than internal consistency, instrument relevance is weak. ORIV therefore adds little in this specific setting.

[11]The BIC was the method of model selection in the working paper version of Falk et al. (2023) (see Falk et al. 2021a).

https://doi.org/10.1017/eec.2026.10041 Published online by Cambridge University Press

fold, we estimate the model using the training data and compute the RMSE on the remaining dataset. We repeat this procedure 100 times, as the data is randomly partitioned and the results vary across partitions. The average RMSE across the 100 repetitions and 10-folds are presented in the Appendix (Tables A.4). As with the BIC, the objective is to minimize the RMSE. The results show that for time preferences, the RMSE is the smallest for the 5-item model, whereas the BIC selections chose the 2-item model. For risk preferences, the RMSE is the smallest for the 5-item model, where the BIC selections chose the 4-item model. For altruism, the RMSE is the smallest for the 3-item model, but the BIC selection chose the 2-item model. For negative reciprocity, the RMSE is the smallest for the 5-item model, whereas the BIC selection chose the 3-item model. For positive reciprocity, the RMSE is the smallest for the 3-item model, where the BIC selection chose the 2-item model. For trust, the RMSE is the smallest for the 3-item model, where the BIC selection chose the 1-item model. In sum, the RMSE from the cross-validation in Table A.4 generally favors larger models than those chosen by the BIC. We notice though, that the increase in RMSE is small. For instance, for time preferences, the RMSE increases by 2.9% when going from the worst predictive model (1 item) to the most predictive model (5 items).

As our goal is to prioritize model simplicity, we choose the BIC-optimal models. We visualize the procedure in Figure 2 that displays the R^2 and BIC for the combinations of survey items that maximize the R^2 for each preference.

Table 1 reports the estimated coefficients from the preferred regressions discussed above: all selected survey items are statistically significant at the 5-percent level. Appendix A presents similar results using standardized measures, which account for differences in the scales between the experimental tasks and the survey items.

The results in Table 1 show that the hypothetical version of our incentivized choice experiments is included in our survey for all preferences, similar to Falk et al. (2023). The survey items are all positively associated with the experimental measures, with one exception: *Risk-16* is negatively correlated (see Table 2) with the experimental measure. Risk-16 reads '*Now, think about what others think of you. Do your friends think that you are brave?*' While one might expect a positive correlation between perceived bravery and risk tolerance, our results contradict this assumption, reinforcing the importance of empirical validation. For the specific correlation, we do not have a good explanation, and we refrain from speculative interpretation, but we chose to retain the item in accordance with our predefined methodology.

Table 2 shows the chosen survey items for each preference. Table A.20 in the Appendix shows all the survey items and their correlations with the decisions in our incentivized choice experiments. In bold, we indicate the selected items. In all but one case, the most correlated items are those chosen by our validation exercise.

Finally, we test the validity of the results, and we compare our findings with Falk et al. (2023). In Falk et al. (2023), the authors assess the quality of their survey and report the correlation between the behavior in the incentivized choice experiments predicted by the survey and actual behavior in the experiment. The correlations in Falk et al. (2023) vary from 0.37 (negative reciprocity) to 0.67 (trust). In our data, the correlations vary from 0.34 (trust) to 0.59 (positive reciprocity). Interestingly, Falk et al. (2023) report the correlations between two repetitions of the exact same incentivized choice experiment conducted one week apart. The within-subject correlations of behavior vary from 0.59 (risk) to 0.82 (time). Frey et al. (2017) report a test-retest correlation below 0.5 for risk preferences elicited using multiple price lists. In conclusion, it is not accurate or fair to evaluate the correlations between survey and incentivized choice experiments against a benchmark of 1.

We perform a final check to see whether random survey answering can produce significant associations between the survey items and the experiment, by generating new survey answers at random but keeping the same distribution as the real survey answers. While the simulated data preserves the marginal distribution of each survey item, it breaks any correlation structure between items and behavioral outcomes, as well as among survey items themselves. This ensures that any

https://doi.org/10.1017/eec.2026.10041 Published online by Cambridge University Press

Table 1 Regression results for all preferences

Variables	Time	Risk	Altruism	Neg. Rec.	Pos. Rec.	Trust
Time-1	0.298*** (0.067)					
Time-11	0.133*** (0.054)					
Risk-5		0.250*** (0.083)				
Risk-6		1.011*** (0.323)				
Risk-15		0.384** (0.159)				
Risk-16		−0.595*** (0.219)				
Altruism-9			0.309** (0.137)			
Altruism-20			0.207*** (0.066)			
NegRec-4				0.417*** (0.063)		
NegRec-8				0.479*** (0.136)		
NegRec-12				0.321** (0.139)		
PosRec-1					0.580*** (0.065)	
PosRec-5					0.208** (0.097)	
Trust-11						0.387*** (0.083)
Constant	1.773*** (0.213)	2.939*** (1.081)	0.401* (0.222)	1.378*** (0.313)	0.328 (0.452)	0.233*** (0.159)
Observations	147	151	150	179	159	164
R^2	0.162	0.185	0.124	0.300	0.348	0.118

Note: Table shows estimation coefficients from an OLS regression of the experimental measure on the chosen survey items for all six preferences. The number of observations changes because we had to exclude children who did not answer an item of the survey or who left the class during the experiment. Standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

predictive power observed with the real survey is not an artifact of item-level distributions or random alignment. Then, we regress the random survey items with the experimentally elicited preferences. We repeat this procedure 1000 times. Table 3 shows the parameter estimates based on the simulated survey answers. Reassuringly, the mean estimate is zero for all randomly generated survey answers (see column 1). The minimum and maximum values are symmetrical around zero (see columns 2 and 3), and the distributions are dense around zero with a low standard deviation (see column 4). Hence, data that simulates random answering of the survey does not explain the behavior in the incentivized choice experiments.

3.1. Alternative short survey

Running the hypothetical experiments can be tedious and challenging to do. To investigate the importance of the hypothetical games for predicting behavior in the incentivized experiments, we redo

https://doi.org/10.1017/eec.2026.10041 Published online by Cambridge University Press

Table 2 Correlations between survey question and behavior in the experiment

ID Item	*Preference*	*Question*	*Correlation*	*p-Value*
Time-1	Time	Hypothetical version of the choice experiment	0.355	<0.001
Time-11	Time	*I really don't like to wait for the red light to turn green.*	0.213	0.0096
Risk-5	Risk	Hypothetical version of the choice experiment.	0.230	0.0045
Risk-6	Risk	*Imagine that you have a choice between being sure that you get one family movie ticket or taking a chance. You have to choose 3 times. (Number of risky choices)* *– Do you take one ticket for sure or the chance when the chance offers you an equal probability of winning 2 or 0 tickets?* *– Do you take one ticket for sure or the chance when the chance offers you an equal probability of winning 3 or 0 tickets?* *– Do you take one ticket for sure or the chance when the chance offers you an equal probability of winning 1 or 0 tickets?*	0.246	0.0023
Risk-15	Risk	*Now, think about what others think of you. Does your mother or father tell you to be more cautious?*	0.170	0.0364
Risk-16	Risk	*Now, think about what others think of you. Do your friends think that you are brave?*	− 0.215	0.0079
Altruism-9	Altruism	*Imagine that you have one week of vacation, and for each day, you have ONE ticket for TIVOLI. The tickets have to be used on separate days. Now you have to choose who gets the tickets. How many for you, and how many for a friend?*	0.257	0.002
Altruism-20	Altruism	Hypothetical version of the choice experiment.	0.307	<0.001
NegRec-4	Negative Reciprocity	Hypothetical version of the choice experiment, minimal accepted offer	0.472	<0.001
NegRec-8	Negative Reciprocity	*Imagine that there is a box of candy in your classroom. The candy should be divided between everyone in the class. During a break, only you and one other pupil are in the classroom. The other pupil takes 2 pieces of candy and gives one of them to you. Will you tell your teacher, who will then take the candy away from both of you?*	0.241	0.001
NegRec-12	Negative Reciprocity	*Imagine that you and a friend are both supposed to be quiet. If you are both quiet, you will get 1 piece of candy, and your friend will get 9 pieces of candy. Will you be quiet or say something?*	0.273	<0.001
PosRec-1	Positive Reciprocity	Hypothetical version of the choice experiment	0.574	<0.001
PosRec-5	Positive Reciprocity	*Do you want to help, even if it is costly to do so, or do you not want to, if you have to help your friends?*	0.128	0.107
Trust-11	Trust	Hypothetical version of the choice experiment	0.344	<0.001

Note: Pearson correlation coefficient between the answer in the survey question and the behavior in the incentivized game for the questions chosen in the validation method, by maximizing R^2 and minimizing the Bayesian Information Criterion. Survey items are translated from Danish.

Table 3 Comparison real vs. simulated data (1000 repetitions)

Variables	(1) Mean	(2) Min	(3) Max	(4) SD
Time-1-rand	0.000172	−0.229	0.218	0.0703
Time-11-rand	−0.0000957	−0.192	0.199	0.0569
Risk-5-rand	−0.00118	−0.253	0.258	0.0872
Risk-6-rand	−0.0129	−1.0246	1.033	0.334
Risk-15-rand	−0.00945	−0.910	0.769	0.246
Risk-16-rand	−0.00631	−0.536	0.588	0.184
Altruism-9-rand	−0.00130	−0.227	0.232	0.0700
Altruism- 20-rand	−0.00264	−0.531	0.365	0.137
NegRec-4-rand	0.00153	−0.231	0.263	0.0747
NegRec-8-rand	0.000899	−0.486	0.559	0.167
NegRec-12-rand	0.000922	−0.473	0.569	0.164
PosRec-1-rand	0.000149	−0.224	0.231	0.0769
PosRec-5-rand	0.00360	−0.403	0.442	0.124
Trust-11-rand	−0.00296	−0.276	0.309	0.0893

Note: Table shows the mean, maximum, minimum, and standard deviation of the estimated coefficients when regressing the experimental measure on simulated survey data. We generate the simulated data randomly with the same distribution as in the real data.

the analysis, excluding answers to the hypothetical experiments. The resulting survey has only nine items. These nine items include two for time preferences, three for risk, two for altruism, and two for social preferences (reciprocity and trust). The selection is based on the same BIC-guided validation approach used for the full survey. The complete list of items is reported in the Appendix (Table A.20).

The explanatory power for each preference is lower than in the survey with hypothetical experiments. The explanatory power ranges between 1.6% (trust) and 13.4% (risk).[12] Hence, excluding the hypothetical games drastically lowers the explanatory power. We also notice that the short survey performs worse than the long survey in terms of BIC. In Appendix A, we present the resulting survey along with the regression results (Table A.7), and an illustration of the relationship between R^2 and the BIC (Figure A.1), as well as an overview of the development of RMSE (Table A.19). In the Appendix (Table A.3), we also present the results if only the hypothetical experiments are included.

3.2. Validation with a different sample

As a final test, we investigate how our survey questions perform when using a different sample of children and adolescents (3rd graders, 5th, 7th, and 8–9th graders). We elicited time, risk, and altruism using incentivized-choice experiments, a hypothetical version of the game, and our previously validated survey questions. In addition, we asked three self-evaluation survey questions similar to those in Falk et al. (2023): '*Are you someone who is ready to give up something today to benefit from it in the future?*' for time preferences. '*Are you a child that takes risks, or do you try to avoid risks?*' for risk Preferences, and *'Are you a child who mostly shares with other children without expecting anything in return?*' for altruism.

The resulting dataset allows us to compare three different versions of a survey: A *Long Survey* that includes all items selected by our initial validation procedure; a *Short Survey* that excludes the hypothetical game and a *Self-evaluation Survey* that includes only the self-evaluation question. Table 4 reports our estimates.

[12]Time: $R^2 = 4.5\%$; Risk: $R^2 = 13.4\%$; Altruism: $R^2 = 6.6\%$; Negative Reciprocity: $R^2 = 12.2\%$; Positive Reciprocity: $R^2 = 1.7\%$; Trust: $R^2 = 1.6\%$.

https://doi.org/10.1017/eec.2026.10041 Published online by Cambridge University Press

Table 4 Regression results, 2018 sample, Time, Risk, and Altruism

VARIABLES	(1) Time Long survey	(2) Time Short survey	(3) Time Self-evaluation	(4) Risk Long survey	(5) Risk Short survey	(6) Risk Self-evaluation	(7) Altruism Long survey	(8) Altruism Short survey	(9) Altruism Self-evaluation
Time_1	0.184***								
	(0.033)								
Time_11	0.011	−0.001							
	(0.046)	(0.047)							
Time_2			0.106*						
			(0.061)						
Risk_5				0.287***					
				(0.040)					
Risk_6				0.208**	0.412***				
				(0.106)	(0.107)				
Risk_15				0.188***	0.205***				
				(0.064)	(0.067)				
Risk_16				−0.169*	−0.218**				
				(0.088)	(0.092)				
Risk_1						0.145			
						(0.091)			
Altruism_9							0.177**	0.350***	
							(0.079)	(0.081)	
Altruism_20							0.355***		
							(0.041)		
Altruism_1									0.104**
									(0.049)
Constant	3.293***	3.968***	3.564***	3.007***	4.119***	4.096***	1.008***	1.509***	1.765***
	(0.180)	(0.138)	(0.238)	(0.378)	(0.362)	(0.295)	(0.156)	(0.154)	(0.184)
Observations	552	552	552	518	518	518	536	536	536
R^2	0.054	0.000	0.005	0.143	0.056	0.005	0.152	0.033	0.008
BIC	1910	1934	1931	2119	2163	2178	1594	1657	1671

Note: All regressions are OLS regressions. The hypothetical experiments are Time_1, Risk_5, and Altruism_20. Standard errors in parentheses: ***$p < 0.01$, **$p < 0.05$, *$p < 0.10$.

The results show that the *Long Survey* using the hypothetical version of the game is the most predictive model with a higher R^2 than the other surveys. When comparing the *Short Survey* and the *Self-evaluation Survey*, we find that the *Short Survey* performs better (higher R^2), for Risk Preferences and Altruism. However, the *Self-evaluation Survey* performs better for time preferences.

3.3. Scalability

In the Appendix (Tables A.8–A.13), we present a variety of robustness tests, splitting the sample by boys and girls, as well as by age (born in the first two quarters vs. the two latter quarters). We observe differences across gender and age. For example, for time preferences, the survey item Time_11 is significant only for girls and the younger ages, but the hypothetical experiment is significant and large across all these subgroups. We recommend using the long survey version to ensure comparability across groups defined by age, gender and socioeconomic background.

Our work not only provides a validated instrument for measuring economic preferences in children but also offers a generalizable framework for survey development and validation. Since our full item set and validation procedures are openly available, researchers in other cultural or institutional contexts can use our approach to construct customized short scales that are tailored to their target population. Further, our child-friendly illustrations and figure files are available upon reasonable request for non-commercial scholarly use. This modular structure enhances the scalability of our tool and supports broader efforts to collect preference data at scale in settings where incentivized experiments are not feasible.

4. Conclusion

This study develops and validates a short-form survey instrument designed to elicit six core economic preferences in children: time preferences, risk preferences, altruism, negative reciprocity, positive reciprocity, and trust. We show that a set of 14 child-friendly survey items predicts behavior in incentivized choice experiments, with explanatory power comparable to validated adult surveys. Our findings demonstrate that the survey captures meaningful individual differences and can be implemented in a scalable, developmentally appropriate way across various research and policy contexts.

The main contribution of our work lies in bridging the methodological gap between experimental and survey-based measures of economic preferences in childhood. While incentivized experiments provide high internal validity, their implementation is resource-intensive and often unfeasible in large-scale, repeated, or field settings. Our validated survey offers a practical alternative that maintains predictive accuracy while reducing cost and complexity. This tool can facilitate preference measurement in schools, field interventions, and longitudinal studies – allowing researchers and practitioners to incorporate individual heterogeneity into the design and evaluation of programs targeting children and adolescents.

Our validated survey instruments offer a promising tool for capturing children's preferences, which are integral to understanding behaviors with economic and social implications. While the measures demonstrate robustness and reliability within experimental settings, their application to real-world behavior warrants further exploration. Experimentally elicited preferences such as risk aversion, time patience, and prosociality should manifest in field settings – for instance, through educational choices or peer interactions. However, in the case of children, parental influence plays a significant role in shaping and mediating these behaviors, complicating direct causal interpretations. Future research could explore mediation effects via parental influence using structural modeling or matched registered data, to better understand how early preferences translate into real-world behavior such as academic achievement, health-related decisions, and social participation. Although this data linkage is ongoing, initial findings suggest promising avenues for scaling and generalizing our

https://doi.org/10.1017/eec.2026.10041 Published online by Cambridge University Press

instrument across broader populations. Addressing the interplay between individual preferences and parental decision-making will be essential for interpreting these measures and applying them to the design of effective policies and interventions.

Overall, our work establishes a foundation for scalable and reliable preference measurement in childhood. By enabling researchers to track how individual preferences evolve and how they relate to important life outcomes, this survey instrument contributes to the broader agenda of understanding economic behavior across the life course. It also opens the door to more personalized and effective interventions – whether in education, health, or social policy – tailored to the behavioral profiles of younger populations.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/eec.2026.10041.

Acknowledgements. We thank the Carlsberg Foundation for generous financial support (Project: 108579). We thank the Department of Economics, the Center for Economic Behavior and Inequality (CEBI), and the Centre for Healthy Ageing (CEHA) at the University of Copenhagen for their financial support. Helene Willadsen gratefully acknowledges financial support from the ERC project *DISTRACT* (project No. 834540). Numerous collaborators and students have helped us prepare and run this experiment; a special thanks to our dedicated research assistant Nina Maria Pedersen. We thank Thomas Dohmen and participants from CEBI seminars and the CAM/CEBI workshop 2018 for helpful comments. During the preparation of this manuscript, the authors used ChatGPT for proofreading assistance. The authors subsequently reviewed and edited the content as necessary and take full responsibility for the final version of the publication. The replication material for the study is available at: https://doi.org/10.17605/OSF.IO/4QTDZ.

Conflict of interest. The authors have no conflict of interest.

References

Alan, S., & Ertac, S. (2018). Fostering patience in the classroom: Results from randomized educational intervention. *Journal of Political Economy*, *126*(5), 1865–1911. https://doi.org/10.1086/699007

Almås, I., Cappelen, A. W., Sørensen, E. Ø., & Tungodden, B. (2010). Fairness and the development of inequality acceptance. *Science*, *328*(5982), 1176–1178. https://www.science.org/doi/10.1126/science.1187300

Andreoni, J., Di Girolamo, A., List, J. A., Mackevicius, C., & Samek, A. (2020). Risk preferences of children and adolescents in relation to gender, cognitive skills, soft skills, and executive functions. *Journal of economic behavior & organization*, *179*, 729–742.

Angerer, S., Lergetporer, P., Glätzle-Rützler, D., & Sutter, M. (2015). How to measure time preferences in children: A comparison of two methods. *Journal of the Economic Science Association*, *1*, 158–169. https://doi.org/10.1007/s40881-015-0016-0

Ashton, M. C. & Lee, K. (2009). The HEXACO–60: A short measure of the major dimensions of personality. *Journal of Personality Assessment*, *91*(4), 340–345. https://doi.org/10.1080/00223890902935878

Bettinger, E., & Slonim, R. (2006). Using experimental economics to measure the effects of a natural educational experiment on Altruism. *Journal of Public Economics*, *90*(8-9), 1625–1648. https://doi.org/10.1016/j.jpubeco.2005.10.006

Blake, P., McAuliffe, K., Corbit, J., Callaghan, T., Barry, O., Bowie, A., Kleutsch, L., Kramer, K., Ross, E., Vongsachang, H., & Wrangham, R. (2015). The ontogeny of fairness in seven societies. *Nature*, *528*, 258–261. https://doi.org/10.1038/nature15703

Brandts, J., & Charness, G. (2011). The strategy versus the direct-response method: A first survey of experimental comparisons. *Experimental Economics*, *14*(3), 375–398. https://doi.org/10.1007/s10683-011-9272-x

Cappelen, A. W., Enke, B., & Tungodden, B. (2025). Universalism: Global evidence. *American Economic Review*, *115*(1), 43–76. https://doi.org/10.1257/aer.20230038

Daniels, B. (2013). *Crossfold: Stata Module to perform k-fold Cross-Validation*. Department of Economics. https://ideas.repec.org/c/boc/bocode/s457426.html

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., & Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, *9*(3), 522–550. https://doi.org/10.1111/j.1542-4774.2011.01015.x

Enke, B., Rodriguez-Padilla, R., & Zimmermann, F. (2022). Moral universalism: Measurement and economic relevance. *Management Science*, *68*(5), 3590–3603. https://doi.org/10.1287/mnsc.2021.4086

Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D., & Sunde, U. (2018). Global evidence on economic preferences. *The Quarterly Journal of Economics*, *133*, 1645–1692. https://doi.org/10.1093/qje/qjy013

Falk, A., Becker, A., Dohmen, T., Huffman, D., & Sunde, U. (2021a). The preference survey module: A validated instrument for measuring risk, time, and social preferences. https://www.iza.org/publications/dp/9674/the-preference-survey-module-a-validated-instrument-for-measuring-risk-time-and-social-preferences

Falk, A., Becker, A., Dohmen, T., Huffman, D., & Sunde, U. (2023). The preference survey module: A validated instrument for measuring risk, time, and social preferences. *Management Science*, *69*(4), 1935–1950.

Falk, A., Kosse, F., Pinger, P., Schildberg-Hörisch, H., & Deckers, T. (2021b). Socioeconomic status and inequalities in children's IQ and economic preferences. *Journal of Political Economy*, *129*(9), 2504–2545.

Fehr, E., Bernhard, H., & Rockenbach, B. (2008). Egalitarianism in young children. *Nature*, *454*, 1079–1083.

Fehr, E., Glätzle-Rützler, D., & Sutter, M. (2013). The development of egalitarianism, Altruism, spite and parochialism in childhood and adolescence. *European Economic Review*, *64*, 369–383.

Frey, R., Pedroni, A., Mata, R., Rieskamp, J., & Hertwig, R. (2017). Risk preference shares the psychometric structure of major psychological traits. *Science Advances*, *3*(10), e1701381.

Gillen, B., Snowberg, E., & Yariv, L. (2019). Experimenting with measurement error: Techniques with applications to the caltech cohort study. *Journal of Political Economy*, *127*(4), 1826–1863.

Goodman, R. (1997). The strengths and difficulties questionnaire: A research note. *Journal of Child Psychology and Psychiatry*, *38*, 581–586.

Gullone, E., Moore, S., Moss, S., & Boyd, C. (2000). The adolescent risk-taking questionnaire: Development and psychometric evaluation. *Journal of Adolescent Research*, *15*(2), 231250.

Harbaugh, W. T., & Krause, K. (2000). Children's Altruism in public good and dictator experiments. *Economic Inquiry*, *38*, 95–109.

Harbaugh, W. T., Krause, K., Liday, S. G., & Vesterlund, L. (2003). Trust in children. In E. Ostrom, & J. Walker (Eds.), *Trust and reciprocity: interdisciplinary lessons for experimental research* (pp. 302–322). Russell Sage Foundation.

Iriberri, N., & Rey-Biel, P. (2011). The role of role uncertainty in modified dictator games. *Experimental Economics*, *14*(2), 160–180.

Khachatryan, K., Dreber, A., Von Essen, E., & Ranehill, E. (2015). Gender and preferences at a young age: Evidence from Armenia. *Journal of Economic Behavior & Organization*, *118*, 318–332.

List, J. A., Petrie, R., & Samek, A. (2021). *How Experiments with Children Inform Economics* (No. w28825). National Bureau of Economic Research. https://doi.org/10.1257/jel.20211535

McCrae, R. R., & Costa, P. T. Jr. (2008). The five-factor theory of personality. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: theory and Research* (pp. 159–181). The Guilford Press.

Mischel, W., Shoda, Y., & Rodriguez, M. L. (1989). Delay of gratification in children. *Science*, *244*(4907), 933–938.

Murnighan, J. K., & Saxon, M. S. (1998). Ultimatum bargaining by children and adults. *Journal of Economic Psychology*, *19*, 415–445. https://doi.org/10.1016/S0167-4870(98)00017-8

Samek, A., Gray, A., Datar, A., & Nicosia, N. (2021). Adolescent time and risk preferences: Measurement, determinants and field consequences. *Journal of Economic Behavior & Organization*, *184*, 460–488. https://doi.org/10.1016/j.jebo.2020.12.023

Sutter, M. (2007). Outcomes versus intentions: On the nature of fair behavior and its development with age. *Journal of Economic Psychology*, *28*, 69–78. https://doi.org/10.1016/j.joep.2006.09.001

Sutter, M., & Kocher, M. G. (2007). Trust and trustworthiness across different age groups. *Games and Economic Behavior*, *59*, 364–382. https://doi.org/10.1016/j.geb.2006.07.006

Sutter, M., Kocher, M. G., Glätzle-Rüetzler, D., & Trautmann, S. T. (2013). Impatience and un- certainty: Experimental decisions predict adolescents' field behavior. *American Economic Review*, *103*, 510–531. https://doi.org/10.1257/aer.103.1.510

Sutter, M., Yilmaz, L., & Oberauer, M. (2015). Delay of gratification and the role of defaults—An experiment with kindergarten children. *Economics Letters*, *137*, 21–24.

Sutter, M., Zoller, C., & Glätzle-Rützler, D. (2019). Economic behavior of children and adolescents – A first survey of experimental economics results. *European Economic Review*, *111*, 98–121. https://doi.org/10.1016/j.euroecorev.2018.09.004

Willadsen, H., Zaccagni, S., Piovesan, M., & Wengström, E. (2024). Measures of cognitive ability and choice inconsistency. *Journal of Economic Behavior & Organization*, *220*, 495–506. https://doi.org/10.1016/j.jebo.2024.02.029

Cite this article: Willadsen, H., & Piovesan, M. (2026). Eliciting time, risk, and social preferences in children: a validated survey. *Experimental Economics*, 1–16. https://doi.org/10.1017/eec.2026.10041