LATIN AMERICAN ECONOMIC REVIEW

Volume 34: Issue on Experimental Economics | Article 2 | 2024

Inconsistent Choices Among Adolescents in El Salvador

Gladis Gonzales* Adriana Gaviria†

Submission received: January 16, 2024
Final version received: October 28, 2024
Accepted: November 30, 2024
Published: January 24, 2025

Abstract

This paper examines inconsistencies in the decision-making of a sample of 2,248 adolescents from El Salvador when completing two classic experiments: temporal discounting and risk preferences. Inconsistency in responses is a significant issue when collecting experimental data because it implies a loss of the sample, as these data come from subjects who do not respond to the task as they should and could indicate, for instance, a lack of understanding of the task. To mitigate this problem, we reduced the number of decisions, designed tasks with a strong visual component, and adapted them to the context with the assistance of a local pedagogical team. Despite these adaptations, we first observe participants' significant difficulties in avoiding errors such as multiple switching. Secondly, we investigate whether developmental factors (age), cognitive skills, and task repetition improve consistency. Lastly, we explore whether inconsistency in decision-making somehow shapes their immediate and long-term educational expectations.

Keywords: El Salvador, teenagers, inconsistencies, developmental decision-making, field experiment.
JEL codes: C91, D81.

*Loyola Behavioral Lab, Universidad Loyola Andalucía, email: ghgonzales@uloyola.es
†Loyola Behavioral Lab, Universidad Loyola Andalucía, email: agaviria@uloyola.es

1. Introduction

In recent times, economic experiments have started to involve young populations, typically within school or high school settings. These experimental studies with non-standard subject pools (Exadaktylos et al., 2013; Henrich et al., 2010), fundamentally address three research questions: i) the emergence of economic preferences (Fehr et al. (2013); Brocas and Carrillo (2020b,a)); ii) the exploration of potential relations between these preferences and academic as well as non-academic outcomes (Sutter et al., 2013); and iii) the examination of the feasibility of intervening in these preferences in early ages (Lührmann et al., 2018).

It is important to underscore that experimental studies undertaken with students offer a unique opportunity to gather insights beyond the conventional school metrics. Collecting supplementary data from students contributes to more effective school management and the improvement of educational processes and school environment. For instance, social network studies in schools are useful for launching alcohol and drug prevention programs, among other initiatives (Paluck et al. (2016); Starkey et al. (2009); Valente et al. (2007)).

While the body of literature involving children and adolescents has grown rapidly in recent years, the majority of these studies rely on samples drawn from developed nations, namely the United States and Europe. Although there are exceptions[1], this prevailing trend not only limits our understanding of students in different contexts but, more importantly, hinders the applicability of findings from these settings to environments where such knowledge could be particularly relevant.

An important limitation to transporting experimental tools to developing countries lies in the uncertainty surrounding their utility in more complex socioeconomic contexts. It is well-established that many of these countries are characterized by marked educational gaps, exacerbated by economic constraints and high violence levels that deter school attendance and human capital acquisition ((Adelman and Szekely, 2016; Dinarte-Diaz and Egana-delSol, 2023)).

The primary objective of this paper is precisely to assess the applicability of a tool initially deployed among 5,000 Spanish students to a cohort of 2,248 Salvadoran students. These tasks were meticulously tailored to adolescents (refer to Alfonso et al. (2023a)). The only difference between the Spanish and Salvadoran versions pertains to a linguistic adaptation process conducted by a local pedagogical team to ensure clarity of language while preserving the integrity of the task illustrations.

To test the applicability of these experimental tools, we focused on two classical experiments: first, a task designed to measure time preferences (as outlined in Alfonso et al. (2023a)), and second, a task intended to measure risk preferences (see Vasco and Vazquez (2023)). Both are very simplified tasks. In each case, students make 6 decisions on a set of vignettes where they have to decide between A and B. The success of the portability of these tasks is measured in terms of the percentage of students who make consistent decisions in the tasks. Inconsistency may arise due to an array of factors, including a lack of understanding of the task or simply lack of attention.

Inconsistency poses a serious challenge; when a student responds inconsistently, it means that their decisions fail to accurately reflect their underlying preferences. The immediate consequence of such inconsistency lies in the loss of data reliability. Consequently, the proportion of students offering inconsistent responses promptly translates into a segment of the sample that is unusable. This paper builds on a substantial body of

[1]For example, some studies have measured social preferences among children in El Salvador (Bonan et al., 2023) and Turkey (Alan et al., 2021; Ugur, 2021); competitiveness and risk preferences in Colombia (Cárdenas et al., 2012), Perú (Castillo, 2020), Brazil (Moreira et al., 2010), Uganda (Munro et al., 2014), and Armenia (Khachatryan et al., 2015); and cooperative behavior in Taiwan (Fan, 2000). See Sutter et al. (2019) for a detailed review.

literature examining inconsistent behavior in economic experiments both with standard (e.g., Amador-Hidalgo et al. (2021)) and non-standard subject pools (e.g., Jacobson and Petrie (2009); Tymula (2019)). By extending the analysis to a novel subject pool of Salvadoran adolescents, our study contributes to the literature by providing evidence on how experimental tools perform when applied to a less typical population than university students, who tend to be more motivated and mathematically proficient.

The principal findings of this study can be summarized as follows: First, the percentage of students that make inconsistent decisions in either task is nearly half of the cohort. More importantly, only one quarter of the subjects exhibit consistent performance in both tasks. These outcomes bear significant implications concerning sample size reduction. Furthermore, we do not observe distinctive patterns among schools, with all institutions exhibiting comparably poor performance. Notably, there are no substantial gender differences.

Secondly, we delve into the question of whether developmental factors exert influence over the consistency of responses, specifically whether inconsistency diminishes with age. We find that age has a negative and significant impact on risk inconsistencies, but not on time inconsistencies.

Third, we explore the influence of cognitive skills (including reflection, financial analyses, and probability assessment) in the consistency of their decision-making processes. We show that students with higher cognitive abilities are also less likely to exhibit inconsistencies, especially in risk preferences.

Also, taking advantage of the fact that more than one thousand participants had taken part in a second wave, we test whether repeating the tasks a few months later improves consistency. The results are not particularly encouraging. While many become consistent, there is also a percentage who were consistent in the first wave but cease to be in the second.

Lastly, we analyze whether this consistency problem is associated with students' immediate and long-term educational outcomes, namely: performance at school, the expectations to return to school next year, to attend university, and whether they would like to go to university. We find that the inconsistent students have lower school performance, worse expectations, and even less interest in attending university.

After this introduction, the next sections of this paper are organized as follows: the second section describes the data collection procedures, the tasks, the sample composition, and defines the measures that underpin our analytical framework. The third section presents the findings and the fourth section discusses the implications, while the fifth section concludes.

2. Protocol, tasks and dataset

2.1. Protocol

The study was approved by the Ethical Committee of Universidad Loyola Andalucía and the entire experiment was pre-registered. The project was funded by an international development agency. The experiment was conducted during 2 years in 12 schools located in 5 departments of El Salvador. Subjects were students from middle school (7^{th} and 8^{th} grades) and high school (9^{th}, 10^{th}, 11^{th} and 12^{th} grades). The data was collected in three waves.[2] While the first wave was collected between March and June 2022, the second wave was run between March and May 2023. Respectively, we obtained 2,600 and 2,649 observations for waves 1 and 2.

[2]For this research, only waves 1 and 2 are analyzed. Wave 3 was running in 2023.

Participants were recruited with the support of a local partner who had an established network with public schools across the country. This local partner worked with a field coordinator under the supervision of the research team in Spain. The experiment was adapted to the context with the help of a local pedagogical team, and the adaptations were tested with a group of students from El Salvador.

The recruitment process was identical for both waves. Participants were not self-selected. School directors signed an agreement to integrate the experiment into their pedagogical curriculum and to conduct it as a class activity. This eliminated the need of parental consent for subjects under 14 years old, decreasing the non-response rate. Experimental sessions were scheduled with the support of the local partner, and the experiment was conducted as a lab-in-the-field on a platform called SAND[3] that was developed to provide greater control over data privacy and adaptability. Subjects were informed that all responses were anonymous and were provided with general information and instructions, as well as the necessary tools to complete the experiment: a tablet, internet connection, username, and temporary password to access SAND. Once subjects logged in, they completed the experiment by navigating through screens containing specific instructions for each task.[4] Payments were hypothetical, knowing they elicit similar results to monetary payments in adolescent subject pools (refer to Alfonso et al. (2023a)).

2.2. *The truck and the gumball tasks*

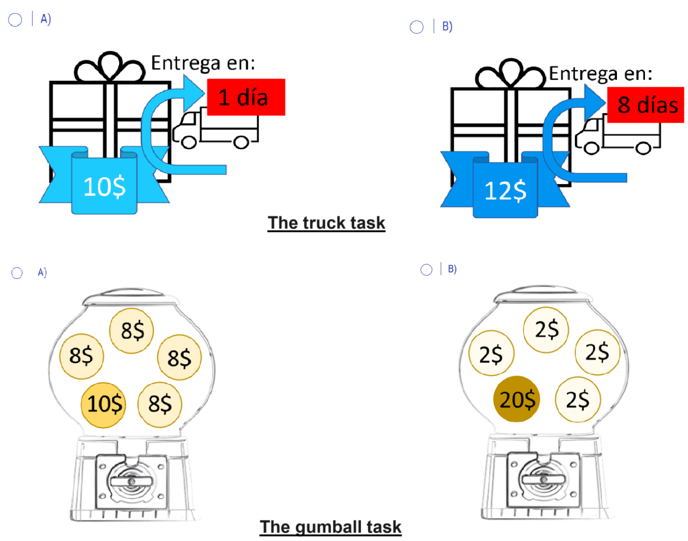
This study uses tailored tasks designed explicitly for the non-adult population. *The Truck task* is a visual version of the Multiple Price List task of Coller and Williams (1999) to elicit time preferences developed by Alfonso et al. (2023a). *The Gumball Machine* is a graphic version of Holt and Laury (2002) test to measure risk preferences introduced by Vasco and Vazquez (2023). Figure 1 shows the first screen for participants of the truck (top) and the gumball (bottom) tasks.

In both tasks, subjects made six consecutive decisions. For the truck task, they were asked to choose in each scenario between 10$ *tomorrow* or 10$+$x$ (being $x = 0, 2, 4, 6, 8, 10$ dollars) one *week later*. For the gumball, they choose between two paired lotteries (A and B) with high and low payoffs. Lottery A is initially better than Lottery B until p_{high} becomes sufficiently large, and lottery B is more attractive (and rewarding). We explain both tasks in detail in Appendix A.

[3] Acronym for Social Analysis and Network Data. This platform is offered by Kampal company (https://www.kampal.com/).

[4] In 8 schools, participants first completed the temporal discounting (*The Truck*) and then they completed risk preferences (*The Gumball*). The task order was inverted for the last 4 schools. We control for this through school fixed effects.

Figure 1: *The truck and gumball tasks*



Under *consistency*[5], the truck task lets us compute several measures of time preferences: *number* of future allocations ($\#Future$, from 0 to 6 choices of one week later) and three dummies: subjects who allocate all in present ($AllPresent$), future ($AllFuture$) or those who use interior allocations at least once ($Interior$).

Similarly, the gumball task lets us compute several measures of risk preferences under *consistency*[6]: $\#Risky$ counts from 0 to 6 the number of risky allocations (option B vs A) and three dummies: $Averse$ if the subject chooses at least three A's in the first five decisions –i.e. AAABBB, AAAABB, AAAAAB–, $Neutral$ if she/he chooses AABBBB and $Lover$ if she/he chooses ABBBBB.

2.3. Other tasks

To study the role of individual abilities on how subjects make choices we used two different measures: cognitive and probability abilities. We used two complementary tasks to study cognitive abilities: the Cognitive Reflection Test (CRT) adapted for teenagers and a financial numeracy test ($FinAb$) comprising three mathematical questions related to basic operations and interest rates. To study probability abilities we used an extended version of the Delavande test (Delavande and Kohler, 2009) adapted by Brañas-Garza et al. (2021) that allows assessing subjects' accuracy and consistency when handling probabilities. Both tasks were previously tested with adolescents in Spain (Alfonso et al., 2023a). See Appendix A.3, A.4 and A.5 for details.

We also collected data on students' educational outcomes, expectations, and aspirations. For short-term educational outcomes, students reported their performance from the previous academic year, specifically indicating the number of subjects in which they received excellent or good grades. They were also asked about their likelihood of continuing their studies in school the following year.[7] Additionally, we assessed long-term

[5] As we explain in Section 2.5, consistent subjects in the truck task are those who do not switch back.

[6] There are three types of inconsistency: i) *lack of understanding* when subjects selected lottery B in the first decision, which had no uncertainty, and B is dominated by A; ii) *Switch back*, when subjects switched back from lottery B to A; and iii) *inattention* when subjects chose lottery A in the sixth decision which had no uncertainty, and A is dominated by B. See Section 2.5 for more details.

[7] As noted in 2.6, 12th-grade students were excluded from our analysis, meaning that all students in the sample were expected, in theory, to continue studying in school the next year.

educational expectations and aspirations by asking students to rate their likelihood of attending university and whether they desired to go to university.[8]

2.4. Schools

In Table 1, we summarize the main characteristics of the 12 schools that participated in the study. All are public institutions and are evenly distributed across urban and rural areas.[9] Most schools offer both middle and high school education, with a few exceptions.[10] The total number of middle and high school students varies significantly, with the smallest school having 70 students and the largest 560. Additionally, we gathered data on reported violent incidents in the municipalities where the schools are located, often linked to gang activity.[11]

Table 1: *School description*

School	Rural/ Urban	Students (2022)	Violence index* (2011 to 2021)	Municipality	Department
1	R	184	0.11	Coatepeque	Santa Ana
2	R	129	0.19	Chalchuapa	Santa Ana
3	R	239	0.07	Tacuba	Ahuachapán
4	U	240	1.00	San Salvador	San Salvador
5	R	143	0.07	Tacuba	Ahuachapán
6	R	70	0.48	San Ana	Santa Ana
7	U	381	0.51	San Miguel	San Miguel
8	U	235	0.09	Acajutla	Sonsonate
9	U	279	0.16	Tonacatepeque	San Salvador
10	U	236	1.00	San Salvador	San Salvador
11	U	560	0.48	Soyapango	San Salvador
12	U	395	0.07	El Congo	Santa Ana

*Note: The Violence index is based on official data by the Tripartite Roundtable of El Salvador composed of the Legal Medicine Institute, Attorney General's Office and the Police. This information was recorded prior to the exception regime that began on March 27^{th}, 2022. The index corresponds to the normalized number of violent events reported: 1 represents the greatest number of violent events (2,218) in the period 2011 and 2021. Published by La Prensa Gráfica (2021) and complemented with Instituto de Medicina Legal (2021).

2.5. Definitions

[8] Expectations and aspirations were measured on a scale from 0 to 100.

[9] Information provided by the local partner.

[10] Schools 10 and 12 only provide classes up to 9^{th} grade.

[11] These gangs, locally known as *maras*, are part of criminal structures that dominate economic and social activities within certain territories. Their presence in schools has various impacts: students face violence, including threats, murders, and sexual assault, and are at risk of dropping out due to gang recruitment efforts. Furthermore, gang presence generates tensions within the school environment. For more details, refer to Dinarte-Diaz and Egana-delSol (2023).

Inconsistency

Along this paper we consider three types of inconsistent behavior: subjects who make wrong choices in *time, risk* and *both* ($time \cup risk$). For simplicity, we do consider whether the individual makes or not an inconsistent choice in the respective domain. Observe that we might also consider *intensity* (number of inconsistencies) in the same task. For instance, in Holt-Laury elicitation we may find at least three types of inconsistencies: *lack of understanding* (playing B in the first choice), *switching* and *inattention* (playing A in the last choice, strictly dominated) (see Amador-Hidalgo et al. (2021)). Therefore, a subject might be inconsistent in *time*, $i_i^t = 1$, in *risk* $i_i^r = 1$, or in *both*, $I_i = i_i^t + i_i^r = 0, 1, 2$. To sum up:

Time: refers to multiple switching along the temporal discounting task.

Risk: refers to choosing dominated strategies and/or multiple switching along the risk task.

Both: refers to being inconsistent in time and/or risk task.

In this paper we explore inconsistency by task (i_i^t, i_i^r) and in aggregate (I_i). While the former measures are useful to test whether a particular domain may impact differently, the later helps us to identify subjects with strong problems in decision-making ($I_i = 2$).

Individual abilities

We constructed different indexes to study the role of individual abilities in decision-making. The first one is Cognitive abilities and reflects subjects' ability to think deliberately and carefully but also some analytical skills, particularly in financial maths: $CogAb = CRT + FinAb$.

On the other side, we created a Probability abilities index based on an extended version of Delavande test (see Brañas-Garza et al. (2021)). This index is composed of the number of correct probability estimations ($accuracy$) and the consistent decisions between sets ($noviolations$)[12]: $ProbAb = accuracy + noviolations$.[13] Our two measures of individual abilities, $CogAb$ and $ProbAb$, are highly correlated ($\varphi = 0.33, p = 0.00$).

Finally, we built an index from 0 to 12 according to students' reports of the number of school subjects in which they obtained excellent or good grades during the previous school year. We refer to this variable as *Self-reported GPA*.

2.6. *Students*

Along this section, we summarize the main features of the 2,248 students who participated in the study.[14] Descriptive statistics are presented in Table 2. Half of the sample (49%) were female and the mean age was 15.46 years old. The average $CogAb$ and $ProbAb$ scores were 1.49 and 2.5, respectively. In terms of school

[12] An example of a violation would be, for example, reporting that the likelihood of choosing an apple is higher when there are ten apples in a basked than where there are five apples in a basket.

[13] Both $CogAb$ and $ProbAb = 0, 1, ..., 6$.

[14] From a total of 2,600 observations obtained, the following criteria were taken to obtain the final sample: task completion (truck, gumball, CRT, financial abilities and Delavande test), *age* between 12 and 20 years, being in the school years from 7^{th} to 11^{th} grade and having specified their gender as *female* (or male).

performance, the mean *self-reported GPA* was 2.05.[15] The questions on long-term expectations show that the average likelihood of going to university is 63.2% and studying next year is 86.48%. Finally, the mean desire to go to university is 85.25%.[16] Table 2 also shows that the proportion of 7^{th}, 8^{th} and 9^{th} grade students is similar (approximately 25%), while 13.79% are 10^{th} graders and 13.21% are 11^{th} graders.

Table 2: *Descriptive statistics*

Variables	n	Mean	Min.	Max.
Female	2,248	0.49		
Age	2,248	15.46	12	20
CogAb	2,248	1.49	0	6
ProbAb	2,248	2.50	0	6
Self-reported GPA	2,248	2.05	0	12
Likelihood of going to university	2,248	63.2	0	100
Likelihood of studying next year	1,160	86.48	0	100
Wants to go to university	1,160	85.25	0	100
Distribution across grades:				
7th grade	573	25.49%		
8th grade	540	24.02%		
9th grade	528	23.49%		
10th grade	310	13.79%		
11th grade	297	13.21%		
Total	2,248			

3. Results

3.1. Overview

Figure 2 shows the observed aggregate inconsistencies for *time* preferences, *risk* preferences, and *both* (within-subjects). Numbers are not positive: 45% of students made at least one inconsistent choice along the truck task (left), 57% were inconsistent when faced the gumball machine (center), and more concerning, only 27% were consistent in both tasks (right side).

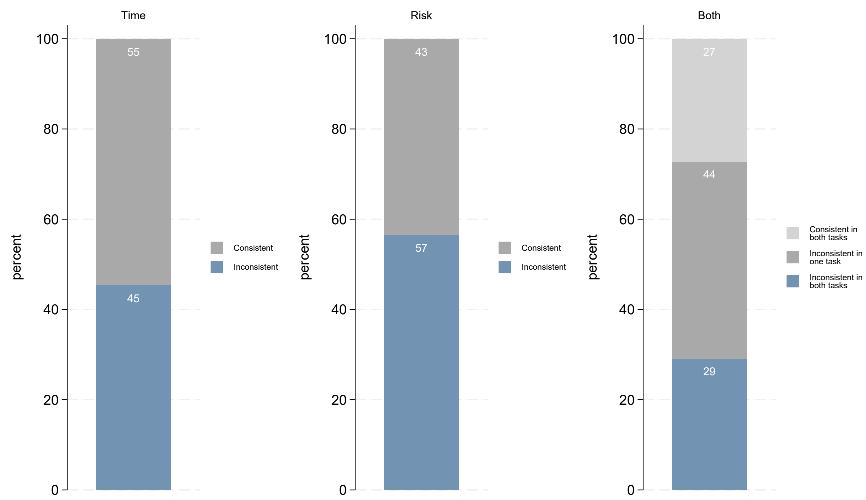
Compared to the outcomes of Spanish students who used identical tasks (see Alfonso et al. (2024)) with average levels of inconsistency of about 15%, these results are extremely concerning.[17] Recall that the only difference between Alfonso et al. (2024) and these data is that the language of the instruments was adapted by a local pedagogical team.

[15]The distribution of cognitive abilities and school performance variables can be found in Appendix B.

[16]There are fewer observations for the questions on likelihood of studying next year and wanting to go to university because they were included during data collection.

[17]Inconsistency levels are also high compared to previous studies in other contexts. For time preferences, Castillo et al. (2011) used an instrument that is different from the truck task (but can also be used to assess inconsistencies) and found an inconsistency rate of 31% among 8^{th} graders in the United States. Using a similar task with a sample between 5 and 16 years old, Bettinger and Slonim (2007) found that 34% was inconsistent. For risk preferences, in a recent study Filippin and Crosetto (2016) found an average inconsistent behavior of 14% across 41 studies. It is important to note, however, that Jacobson and Petrie (2009) found a similar level of inconsistency in Rwanda (55%).

Figure 2: *Inconsistency across tasks*



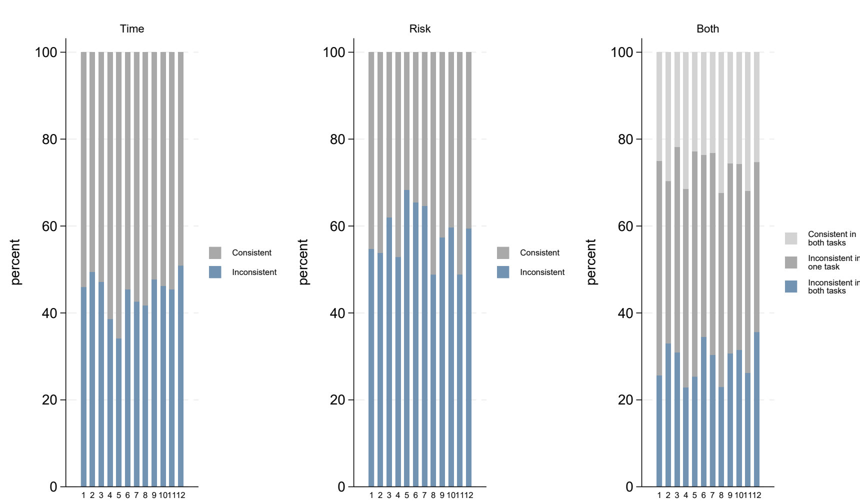
Result 1: *45% of students are inconsistent in time, 57% in risk while only 27% were consistent in both tasks.*

3.2. Inconsistencies by schools

In this section, we compare results across schools. The main reason why we run this comparison is to assess whether the previous results (Figure 2) are mostly due to certain schools with severe problems or, more concerning if adolescents' difficulties when facing the tasks are general. Figure 3 shows the results.

Regarding inconsistencies in *time* preferences (left side) we do not observe substantial differences among schools. In most cases, the deviation from the average is not larger than 5%. The only exceptions are schools #4 and #5, with inconsistency levels of 39% and 33% respectively. For *risk* preferences (center) we observe a similar pattern: with few exceptions, most schools show similar numbers. Interestingly, in school #5 we observed a *lower* level of inconsistency in time preferences, but *higher* inconsistency in risk preferences.

Figure 3: *Inconsistency across schools*



The graph on the right shows the fraction of inconsistent (within) subjects in both tasks. By inspection, we can see that, on average, schools are pretty similar with a fraction of around 27% of students who performed both tasks consistently.

Table 3 shows a series of models to estimate the effect of each school.[18] Models 1, 4 and 7 study inconsistencies in *time* (also introducing *CogAb* and *ProbAb*). Models 2, 5 and 8 analyze inconsistencies in *risk* (with the same specifications). Similarly, models 3, 6 and 9 focus on inconsistency on *both* tasks (being no-inconsistency=0, inconsistent in one task=0.5, and 1 when the subject fails in both tasks). The reference group is school #11.

School effects are rather limited. Similarly to results presented in Figure 3, Table 3 shows that school #5 does better in *time* preferences ($p < 0.01$) but worse in *risk* preferences ($p < 0.01$, like schools #6, #7 and #12) but on aggregate (*both*) these differences balance each other in the majority of cases with some weak differences like #6, #7 and #12 ($p < 0.05$). When we introduce *CogAb* these differences decrease[19] and when we control for *ProbAb* differential behavior across schools completely vanishes.[20]

Result 2: *The level of inconsistency across schools of our sample is not different.*

We cannot generalize these results to all the schools in the country since our sample was not collected with the aim of being representative. However, our sample is composed of 12 schools with different sizes, locations (urban/rural), economic backgrounds, and violence exposure (presence of *maras*). See Section 2.4 for details.

3.3. Inconsistencies by gender

Now, we study whether girls and boys performed the tasks differently. Figure 4 shows the fraction of inconsistent students for *time*, *risk*, and both tasks. The figure suggests that girls and boys performed similarly on the *time* preferences task (graph on the left), with 45% of both groups making at least one inconsistent choice. Models 1, 4, and 7 (Table 3) confirm this result, showing no statistically significant difference ($p > 0.1$ in all cases).

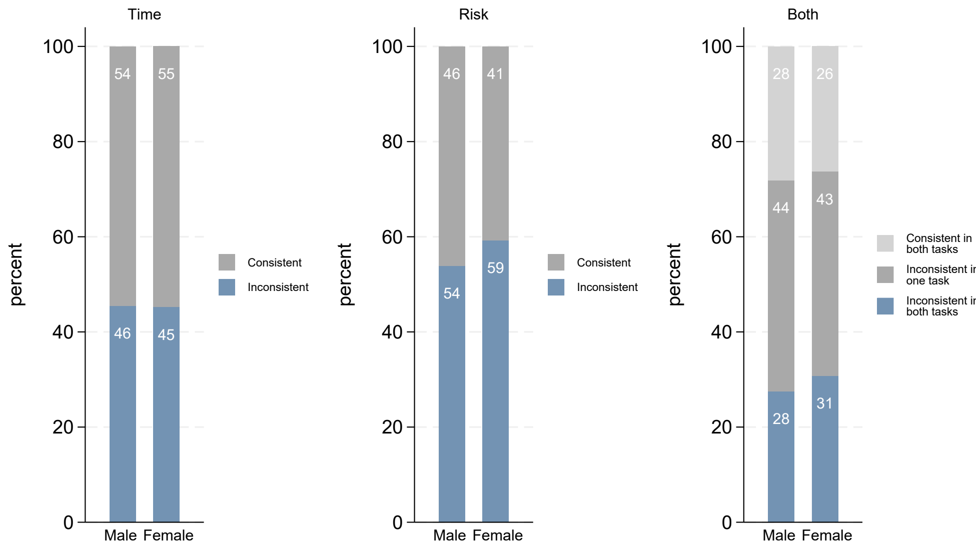
For *risk* preferences (graph in the center) the situation differs. Girls made more errors than boys (59% vs. 54%) and the difference is strongly significant in model 2 ($p < 0.01$) – Table 3 – but becomes weakly significant ($p < 0.05$) after adding controls for abilities (models 5 and 8). Finally, the right side shows that there are no gender differences for *both* tasks, as 26% (28%) of the girls (boys) were consistent in *both* ($p > 0.1$). Models 3, 6 and 9 show that gender differences in *both* tasks vanish.

[18] The table presents robust standard errors with clustering at the class level. As a robustness check, Table C.0.1 in the Appendix replicates this table with clustering at the individual level. The results described in the paper are substantially the same.

[19] The only exception is #5 for *time* which stays similar.

[20] The only exceptions are: school #5 at 1% for *time* and school #7 at 1% for *risk*, schools #5, #6 and #12 at 5% for *risk*, and school #12 at 5% for *both*.

Figure 4: *Inconsistency by gender*



Result 3: *Inconsistency in time preferences is not different between boys and girls, but girls are more inconsistent in risk preferences. However, gender differences vanishes in both tasks.*

3.4. Developmental consistency

Now we explore whether inconsistencies evolve with age, that is, if developmental factors exert influence over the consistency of responses, specifically whether inconsistency diminishes with age.

Figure 5 shows the average level of inconsistency in *time* (left), *risk* (center) and *both* (right) by age. Apparently, the performance in these tasks seems to improve with age (since inconsistency seems to decrease). However, in order to have a more precise picture Table 3 explores the developmental side of these behaviors in detail using regressions (with and without controls for cognitive abilities).

Models 1, 4 and 7 show that age has no impact on *time* inconsistencies ($p > 0.1$) implying that subjects do not learn to complete the task properly as they become more mature. Models 2 and 5 show that age has a negative and significant ($p < 0.01$) impact on *risk* inconsistencies, that is, as subjects become older (more mature) they are more likely to make consistent choices under uncertainty. Model 8 shows that the effect of age on *risk* inconsistencies is weaker ($p < 0.05$) when we control for *ProbAb*.

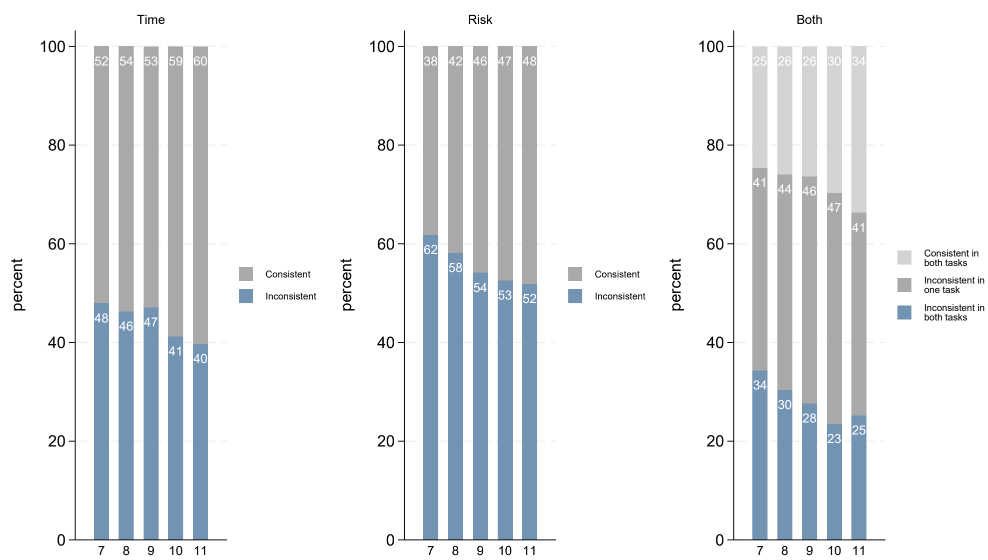
Table 3: *Determinants of inconsistency*

Vars.	(1) *Time*	(2) *Risk*	(3) *Both*	(4) *Time*	(5) *Risk*	(6) *Both*	(7) *Time*	(8) *Risk*	(9) *Both*
Female	-0.006	0.057***	0.026*	-0.008	0.047**	0.020	-0.011	0.043**	0.016
	(0.021)	(0.021)	(0.015)	(0.021)	(0.021)	(0.014)	(0.021)	(0.020)	(0.014)
Year	-0.011	-0.030***	-0.020***	-0.010	-0.024***	-0.017**	-0.008	-0.020**	-0.014**
	(0.009)	(0.009)	(0.007)	(0.009)	(0.009)	(0.007)	(0.009)	(0.009)	(0.006)
CogAb				-0.008	-0.032***	-0.020***			
				(0.010)	(0.010)	(0.007)			
ProbAb							-0.020**	-0.052***	-0.036***
							(0.008)	(0.008)	(0.006)
School 1	0.017	0.029	0.023	0.015	0.021	0.018	0.008	0.005	0.007
	(0.053)	(0.049)	(0.037)	(0.053)	(0.048)	(0.037)	(0.053)	(0.049)	(0.037)
School 2	0.059	0.021	0.040	0.053	-0.003	0.025	0.042	-0.024	0.009
	(0.082)	(0.052)	(0.049)	(0.083)	(0.052)	(0.049)	(0.085)	(0.049)	(0.051)
School 3	0.041	0.099*	0.070*	0.037	0.083	0.060	0.028	0.065	0.047
	(0.043)	(0.057)	(0.038)	(0.045)	(0.058)	(0.039)	(0.044)	(0.059)	(0.040)
School 4	-0.061	0.030	-0.015	-0.061	0.027	-0.017	-0.065	0.018	-0.023
	(0.043)	(0.043)	(0.028)	(0.043)	(0.042)	(0.028)	(0.042)	(0.044)	(0.028)
School 5	-0.098***	0.182***	0.042	-0.105***	0.156***	0.026	-0.116***	0.134**	0.009
	(0.036)	(0.057)	(0.032)	(0.038)	(0.056)	(0.032)	(0.037)	(0.052)	(0.028)
School 6	0.029	0.123***	0.076**	0.025	0.107***	0.066**	0.016	0.088**	0.052
	(0.047)	(0.041)	(0.033)	(0.048)	(0.040)	(0.033)	(0.048)	(0.042)	(0.034)
School 7	-0.043	0.136***	0.046**	-0.046	0.123***	0.038*	-0.058	0.095***	0.019
	(0.046)	(0.028)	(0.022)	(0.047)	(0.028)	(0.023)	(0.049)	(0.027)	(0.025)
School 8	-0.040	-0.031	-0.035	-0.042	-0.037	-0.039	-0.047	-0.048	-0.047
	(0.045)	(0.044)	(0.037)	(0.045)	(0.044)	(0.037)	(0.045)	(0.045)	(0.037)
School 9	0.026	0.071	0.048	0.022	0.057	0.040	0.014	0.040	0.027
	(0.027)	(0.053)	(0.033)	(0.028)	(0.052)	(0.033)	(0.027)	(0.050)	(0.031)
School 10	0.000	0.074**	0.037	-0.004	0.059	0.028	-0.008	0.053	0.022
	(0.037)	(0.036)	(0.022)	(0.038)	(0.037)	(0.024)	(0.038)	(0.036)	(0.023)
School 12	0.029	0.079***	0.054**	0.028	0.077***	0.053**	0.026	0.071**	0.049**
	(0.034)	(0.028)	(0.021)	(0.034)	(0.028)	(0.021)	(0.034)	(0.028)	(0.021)
Constant	0.428***	0.617***	0.522***	0.440***	0.663***	0.551***	0.478***	0.749***	0.614***
	(0.065)	(0.069)	(0.050)	(0.067)	(0.071)	(0.051)	(0.067)	(0.074)	(0.053)
Observations	2,248	2,248	2,248	2,248	2,248	2,248	2,248	2,248	2,248
R-squared	0.009	0.023	0.016	0.010	0.028	0.019	0.012	0.044	0.033

Robust standard errors in parentheses (clustering at the class level)

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Figure 5: *Inconsistency across school years*



Models 3, 6 and 9 (Table 3) analyze inconsistencies in *both* tasks and we observe that the impact of age is again negative and significant ($p < 0.01$ without controls; $p < 0.05$ when we control for individual abilities *CogAb* or *ProbAb*). However this effect seems to be entirely driven by the improvement in risk choices with age – since time decisions do not improve.

Result 4: *As subjects develop they are less likely to make inconsistent choice under uncertainty but do not learn to make inter-temporal choices consistently.*

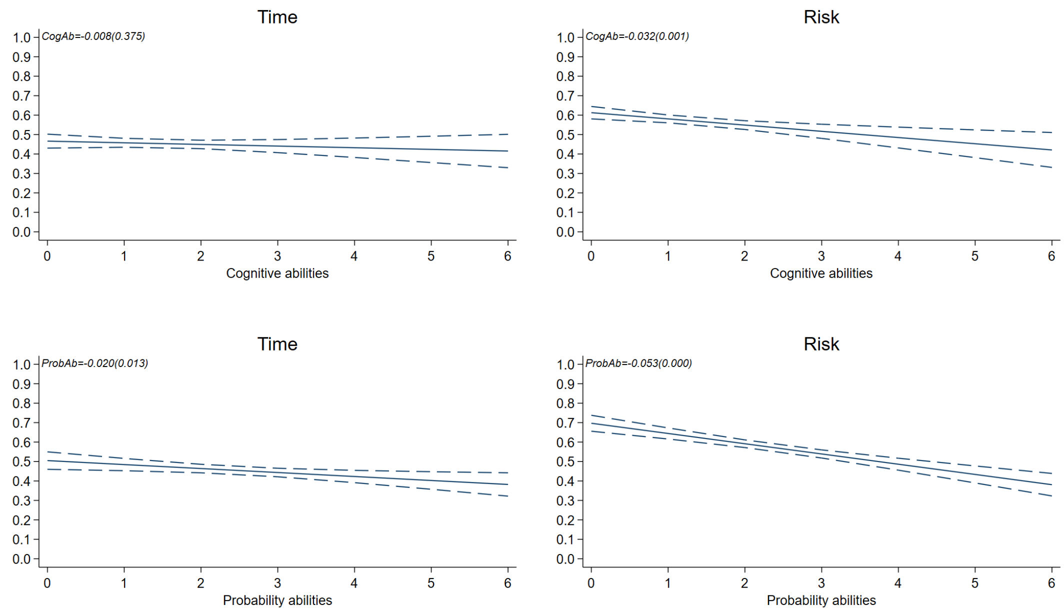
3.5. *The role of cognitive skills*

Models 4, 5 and 6 in Table 3 analyze the impact of *CogAb* on inconsistent choices. While these abilities do not improve (worse) intertemporal decision-making, the impact on uncertain decisions is clear: subjects endowed with these abilities are less likely to make errors ($p < 0.01$, see model 5). By extension, these abilities also decrease the probability of making both tasks inconsistently ($p < 0.01$, see model 6). Models 7 and 8 show that *ProbAb* has a negative and significant impact on both *time* ($p < 0.05$) and *risk* inconsistent decision-making ($p < 0.01$).

Figure 6 presents the estimated effects of *CogAb* and *ProbAb* in the likelihood of making inconsistent choices in *time* and *risk* preferences. Individuals with cognitive and probability abilities are less likely to make inconsistent choices under uncertainty. While subjects with the minimum *CogAb* score (zero) have a 61% likelihood of being inconsistent in the gumball task, this probability drops 20 percentage points for those with the maximum *CogAb* score (six). The effect of *ProbAb* on inconsistent *risk* choices is stronger, dropping from a 70% likelihood of inconsistent choices in subjects with the minimum *ProbAb* score to 38% with the maximum *ProbAb* score.[21]

[21] Given that our measures of cognitive skills $CogAb$ and $ProbAb$ are both composed of different dimensions (CRT and $FinAb$ in the case of $CogAb$; $accuracy$ and $noviolations$ in the case of $ProbAb$), we explored whether these findings were driven by a specific dimension. In tables C.2 and C.3 of Section C in the Appendix, we compare the overall effect of $CogAb$ and $ProbAb$, respectively, with the individual exploration of the dimensions that these measures are composed of. Our findings suggest that both $CogAb$ measures have similar effects in $time$ and $risk$ inconsistencies, whereas for $ProbAb$ we do find that $accuracy$ is the main

Figure 6: Inconsistency and cognitive abilities



Result 5: *Individuals with cognitive and probability abilities are less likely to make inconsistent choices under uncertainty. Consistency in inter-temporal choices is affected by probability (not cognitive) abilities.*

3.6. Task repetition

A relevant question is to determine whether repeating the task improves the results. For example, Breig and Feldman (2024) found that inexperience with the task is a significant driver of inconsistent choices. In other words, we want to understand whether a student who repeats the task for the second time performs better. To do this, we take advantage of having two waves (with a 9-month gap) and study the behavior of 1,322 students[22] who participated in both the first and second waves.

Figure 7 examines this case. On the horizontal axis, we show how their behavior was in the first wave (in terms of *time*, *risk*, *both*), and on the vertical axis, we have the results of the second wave. The figure shows that 55% of the subjects who were inconsistent in *time* preferences in wave 1 become consistent in the second wave. For *risk*, this percentage is only 36%. The results may seem modest, but positive.

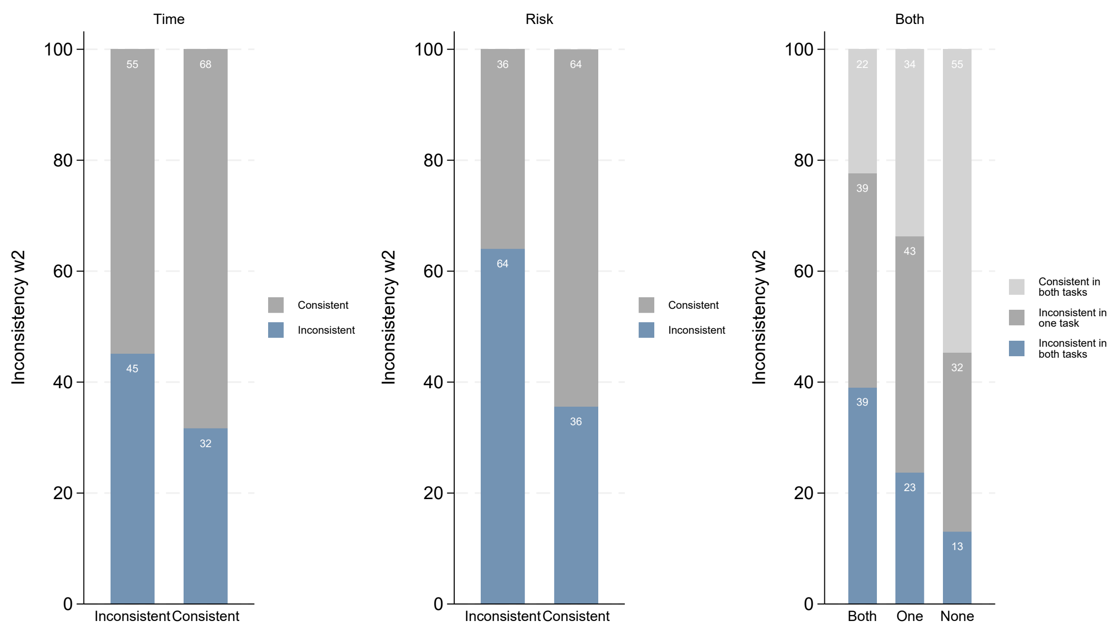
However, if we look at the subjects who were consistent in the first wave, many of them become inconsistent: 32% in the case of *time* and 36% in the case of *risk*. In other words, students who used to perform well are now performing poorly.

The graph on the right shows the behavior in both tasks simultaneously. The results are not positive. Out of the subjects who did not commit any inconsistency in wave 1 (none), only half of them, 55%, remained consistent in the second wave.

driver of inconsistencies in *time* preferences.

[22] A total of 1,960 students from wave 1 repeated the experiment of which 1,322 met the same selection criteria as wave 1.

Figure 7: *Inconsistency across waves*



Result 6: *Task repetition does not improve overall consistency.*

3.7. Time and risk preferences under alternative definitions of consistency

In previous sections, we have seen that using standard definitions of inconsistency creates a huge loss of data. Along this section, we will use alternative – less demanding – definitions of consistency.

Replacement: One inconsistent choice is replaced by the consistent one for all subjects making only one inconsistency.

Training: First choice is considered as training for all subjects. Therefore both time and risk preferences are measured with 5 decisions.

First, we defined eligible subjects for replacement. Among inconsistent subjects, we selected those who would become consistent by changing only one decision: 78% of inconsistent subjects in *time* preferences and 58% in *risk* preferences. Then, we replaced one decision that would lead to a re-classification of those subjects into consistent ones. For *time* preferences, we replaced 611 cases of switch back decisions. For *risk* preferences, we replaced 415 cases of strictly dominated strategies and 296 cases of switch back decisions. [23]

The use of alternative definitions of consistency changes results substantially. In *time* preferences we see that, compared to the standard procedure with only 55% of consistent subjects, the replacement of a single error increases this fraction to 82%. Likewise, using the first choice as a trial and eliciting *time* preferences with the subsequent five decisions also increases the consistent fraction to 59%.

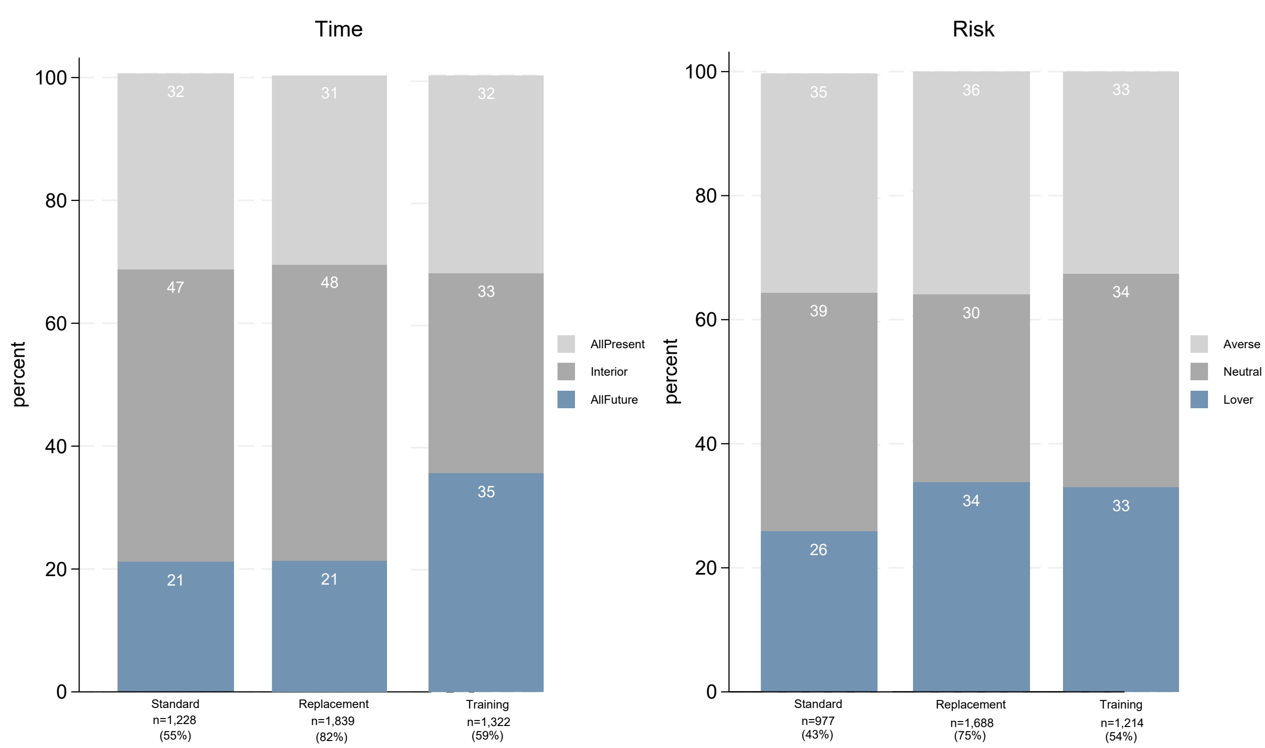
[23]For both tasks, we replaced subject's switching back decision to the one she had made immediately before. Take the case of a subject with the decision pattern ABABBB. In this case, the switching back took place in the third decision A and the replacement consisted in changing it to the one she had made immediately before (second decision) B, leading to the corrected pattern ABBBBB. We acknowledge that this replacement strategy involves non-trivial assumptions that could distort the elicitation of subjects' preferences. For a more elaborate discussion see Appendix E.

Similar results are found for *risk* preferences. While the standard procedure provided 43% of consistent subjects, the replacement increases this fraction to 75% and the use of one trial increases consistency to 54%.

Figure 8 shows the classification of subjects as *present*-oriented ($AllPresent$), *interior* and *future*-oriented ($AllFuture$) for time preferences (Figure 8–left side) and *averse*, *neutral* and *lover* for risk preferences (Figure 8–right side) according to the standard and alternative definitions.

For *time* preferences, replacing one inconsistency does not affect the distribution of preferences: the fraction of *present*-oriented shifts from 32% under the standard definition to 31% (replacement), while the fraction of future-oriented remains in 21% under both definitions. However, eliciting *time* preferences with 5 decisions (training) increases the proportion of *future*-oriented from 21% to 35% and reduces *interior* allocations from 47% to 33%.

Figure 8: *Alternative definitions of consistency*



For risk preferences, the proportion of *averse* subjects under the standard definition (35%) changes slightly when replacing one inconsistency (36%) and when the first decision is taken as trial (33%). Under the alternative definitions, *neutral* decreases from 39% to 30% (replacement) and 34% (training), and *lover* increases from 26% to 34% (replacement) and 33% (training).

To summarize, using alternative – less demanding – definitions of consistency leads to a substantial increase in the fraction of consistent subjects in the sample and slightly modifies the distribution of preferences.

4. Implications

Along this section we explore to different implications of the problem shown across this paper. First refers to the implication to academics or policy-makers aimed to run experiments in the field. Second explore how inconsistencies may have an impact on expectations, that is, whether the expectations of inconsistent individuals differs from those consistent.

4.1. Implications for running experiments

There are two important implications from these results. First, these data bring bad news for multi-cultural experiments with kids and adolescents. Besides the efforts we made to adapt the language (not the graphics) to local students, our findings suggest that these students did not understand the task as Spanish teens did.

Most likely, this problem is directly related to differences in the *quality* of education. While participants are of the same age, the level of educational achievement varies significantly across countries. Another potential explanation could be differences in educational *motivation*. If Salvadoran students are less motivated by school-related activities compared to their Spanish counterparts, this lack of engagement might explain why they did not take their choices as seriously and performed worse on the tasks. Although our data does not allow us to directly test the extent to which *motivation* is driving our results, in the next section we explore the relationship between inconsistency and educational expectations in our sample.

We could also argue that the salience of monetary incentives may play a more significant role in shaping behavior in El Salvador compared to Spain. Given the different socioeconomic contexts, the perceived value of financial incentives could vary, influencing participants' engagement and decision-making processes. However, in adult populations this transportation has worked quite successfully in previous international studies, regardless of potential context differences in the salience of monetary incentives. For example, Brañas-Garza et al. (2021) elicited risk preferences in three countries—Honduras, Nigeria, and Spain—using similar instructions and found negligible differences in inconsistencies across the samples. Furthermore, the authors concluded that eliciting risk preferences with monetary incentives in the field yields results comparable to those obtained in hypothetical scenarios. Brañas-Garza et al. (2023) replicate the analysis for time preferences finding similar results. Future research should assess whether Brañas-Garza et al. (2021, 2023)'s findings hold in adolescent subject pools.

The second problem refers to the quality of the data and the amount of observations we lost due to the lack of understanding or/and inattention. When a participant makes inconsistent choices her data are simply not usable since we cannot elicit her preferences from her choices. We simply do not know the implications of her choices. For instance, her choices might be simply random. The implications for our study are overwhelming: about half of the data are useless for time and more than half are lost for risk preferences. Most of the sample (73%) is lost for both the elicitation of time and risk simultaneously.

4.2. Implications for educational outcomes

A relevant question is whether inconsistency is more than just a nuisance for experimentalists and represents a deeper problem. In other words, if the inconsistencies are showing us a problem in the decision-making fundamentals. If subjects who cannot complete a task satisfactorily are revealing other, more profound issues with long-term implications.

In this section, we will assume that the results obtained are not due to a failure in the transfer of tools from one context to another, but that the results accurately reflect the situation of adolescents in El Salvador. In other words, we assume that in another format, the results would have been similar. Under this assumption, we will examine issues that reflect the immediate and long-term implications of inconsistency.

To simplify, we will consider the extreme cases: those without inconsistency problems (performing both tasks well, 27% of the sample) and those with serious problems (performing both tasks poorly, 29% of the sample), and we will show the Cumulative Distribution Functions (CDFs). Note that these comparisons are based on the standard consistency definitions described in 2.5.

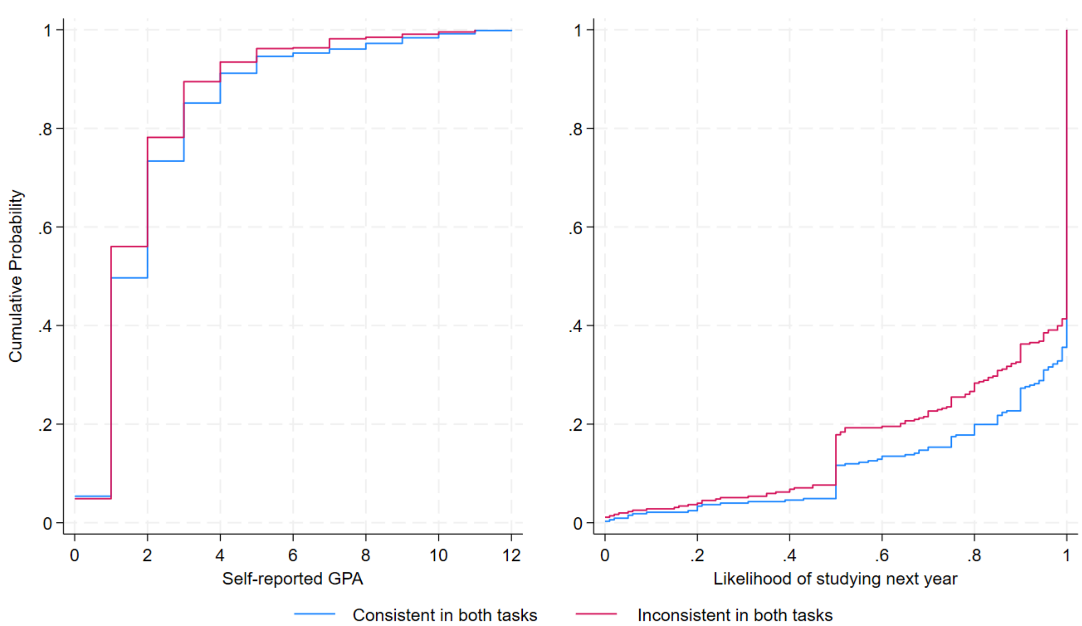
Immediate implications

As immediate measurements of the educational implications of inconsistency, we analyzed:

- Their performance at school (self-reported GPA).
- Their expectations of continuing studying in school the following year.

The red (blue) line in Figure 9 displays inconsistent (consistent) students' CDFs for these two outcomes. As we can see, inconsistent students have lower educational performance and expectations. For instance, while for the 20th percentile of inconsistent students the likelihood of continuing studying the next year is 50% or lower, for consistent students this likelihood increases to 80%.

Figure 9: *Inconsistency and immediate implications*



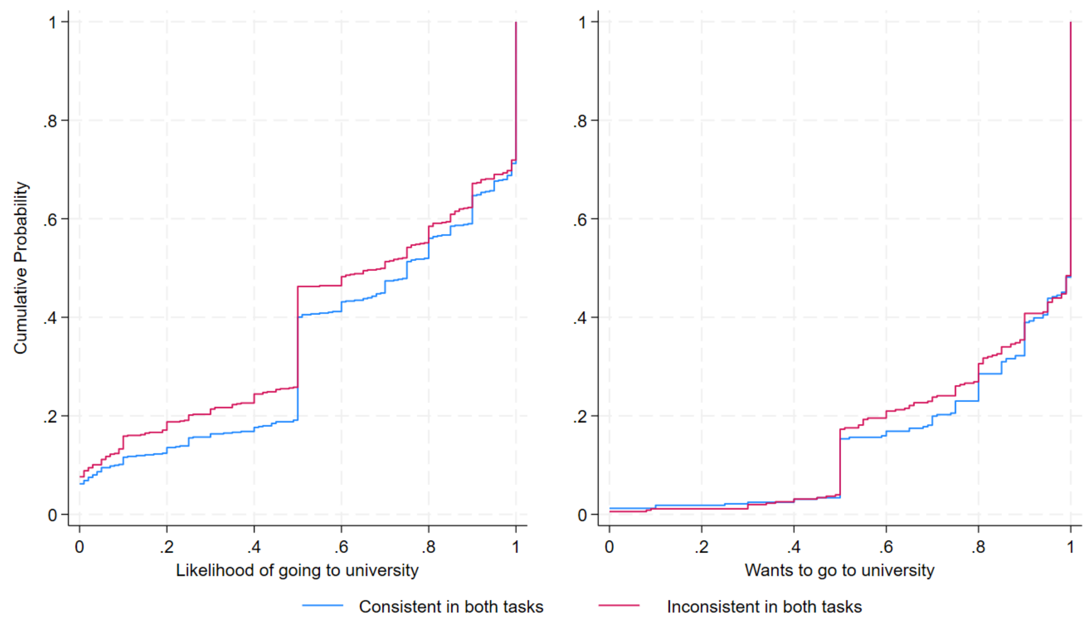
Long-term implications of inconsistency

We analyzed students' long-term educational aspirations in terms of:

- Their expectations of going to university.
- Whether they would like to attend university.

Figure 10 shows that inconsistent students have lower expectations and aspirations of going to university. For the 20th percentile of inconsistent students the likelihood of attending university is 20% or lower, while for consistent students this likelihood increases to 50%.

Figure 10: *Inconsistency and long-term implications*



We extend this analysis to explore how risk preferences and the propensity to make mistakes influence educational expectations. [24] We found that neither riskiness nor the probability of making mistakes is significantly correlated with any of our immediate or long-term educational expectations measures. However, the interaction between riskiness and the propensity to make mistakes is statistically significant for three of our four measures. Conditional on a higher likelihood of making mistakes, risky individuals are less likely to believe that they will study next year, attend university, or even desire to attend university. Nevertheless, their school performance, as captured by their self-reported GPA, is not significantly correlated with riskiness, the tendency to make mistakes, or the interaction between the two. More details are provided in Section G in the Appendix.

To summarize, this section showed that students with problems in the fundamentals have clear educational implications, both immediately and in the long term.

[24] We followed Jacobson and Petrie (2009)'s approach to examine how mistakes affect financial decisions. We employed an error choice model to estimate the error rate in our sample using maximum likelihood estimation. Using this error rate, we then calculated the probability of making at least one mistake for each observed choice pattern in our sample. Finally, we analyzed whether the probability of making mistakes is correlated with subjects' educational aspirations and expectations. The estimates were obtained by sampling with replacement 10,000 times to generate bootstrapped estimates and standard errors.

5. Conclusion

The experiments involving young people and adolescents are scarce, and they are even scarcer in developing countries, where paradoxically, they could be much more relevant in terms of policy implications. This work analyzes how a large sample of adolescents ($n > 2000$) makes decisions regarding time – choosing between a nearby prize and a distant one – and risk – choosing between two lotteries. Our primary concern is to see if students in El Salvador are capable of performing the tasks satisfactorily, and the answer is negative. In fact, just over a quarter of the sample completes both tasks consistently. While we do find some differences between girls and boys – for example, in risk consistency – the reality is that, in the end, the number of girls and boys making mistakes is almost the same.

In the study, we address whether inconsistencies are different among the twelve schools (quite diverse among them), and we find that, for the most part, students from different schools do not exhibit very different behaviors. This implies that the problem could be general, and not a specific case of a particular school. Although the twelve schools are not representative of the country, these findings lead us to believe that they may represent a good number of schools of various natures.

We also analyze whether inconsistency evolves with age and find that it does, as they get better with years. This result is consistent with previous research, such as that of Brocas and Carrillo (2020a) and Alfonso et al. (2024, 2023a), which shows that the ability to choose rationally improves with age. However, the improvement in reasoning of our participants is only modestly positive because, on the one hand, the improvements are not very significant – in fact, in the highest grade (with mean age of 17.7 years), the percentage of consistent boys remains small and concerning (60% for *time*, 48% for *risk* and 34% for *both*) – and, on the other hand, because the previously mentioned works show that consistency appears at much younger ages. For example, in Alfonso et al. (2023b), half of the 6-7-year-olds play games with private information consistently.

In parallel, we analyze the role of cognitive abilities. As expected, we find that subjects who perform better on the CRT (Frederick, 2005; Brañas-Garza et al., 2019) are also less likely to exhibit inconsistencies. Similarly, we find that the grade on a finance test (of three questions) as well as the ability to assign probabilities to events with true value also predict consistency in the tasks. All of this is not surprising and is perfectly consistent with many previous works (see, for example, Vasco and Vazquez (2023); Amador-Hidalgo et al. (2021)).

Taking advantage of the fact that many participants had taken part in a second wave, we tested whether repeating the tasks a few months later would help alleviate the problem. The results are not particularly encouraging. While many subjects improve with repetition (become consistent), there is also a percentage who were consistent in the first wave but cease to be in the second. In summary, the aggregate result is practically null. Therefore, as opposed to what other studies have found (Charness and Chemaya, 2023; Ert and Haruvy, 2017), repeated exposure to risk elicitation tasks, such as the Holt-Laury task, does not always result in individuals making more consistent decisions.

Only about 25% of the sample completes the tasks consistently, and this result is not specific to a particular school but seems quite general, and repeating the tasks doesn't help much either. This number is overwhelming for two very different reasons. First, because it is much lower than the work conducted in Spain with very similar instructions – identical graphic illustrations and minor text adaptations – and identical procedures. Second, because these results indicate that we do not know what or how 75% of the sample has decided on two canonical tasks. In other words, we run the risk of the collected data being useless.

We must ask ourselves what the lack of consistency in decisions means. After all, we are saying that a student fails to choose, for example, between a and A (where $A = d{\cdot}a, d > 1$), which means they are presenting a serious problem with the basics. An alternative explanation is that their problem is not with the basics but with their attention during the experiment, i.e., whether they took it seriously or not. It seems credible that 75% of the students in the entire sample made random decisions. But even accepting that possibility, we have analyzed whether this consistency problem is only a theoretical concern that is not so relevant in the empirical field. For example, Jacobson and Petrie (2009) found that inconsistent choices in the Holt and Laury task lead to sub-optimal financial decisions in Rwanda. To do this, we have taken four measures: their performance at school, their expectations to return to school next year, to attend university, and whether they would like to go to university. The answer is straightforward: in all four cases, the CDF of the consistent group stochastically dominates the CDF of the inconsistent group. This implies that the inconsistent students have worse expectations and even less interest in attending university.

This result no longer seems like just a theoretical concern; rather, it reflects a fairly real problem: a problem with the development of fundamentals at a young age can have long-term consequences on people's lives. It is well known that expectations are a good predictor of outcomes (Samara et al., 2021; Bandura et al., 2001; Khattab, 2015). In this sense, this finding opens up significant avenues for further research. Future studies could investigate the mechanisms behind educational expectations and assess whether targeted interventions can address students' decision-making processes and mitigate the long-term effects of inconsistency. Longitudinal research tracking the life outcomes of inconsistent decision-makers compared to their consistent peers could provide deeper insights into how early decision-making patterns influence long-term success, particularly in educational and career trajectories. Exploring these aspects further could offer valuable contributions to both educational policy and behavioral economics.

Acknowledgements

We thank Pablo Brañas-Garza for his guidance, contributions and feedback. Thanks to Mónica Vasco, Lara Ezquerra, Diego Jorrat and Antonio Alfonso for their advice. We also thank the field team Pablo Montero and Eduardo Gutierrez, and other $TeensLAB$ members –Antonio Cabrales, José Cuesta, Teresa García, Jaromir Kovarik, and Anxo Sánchez– for their comments and ideas. This paper was presented at LAER Workshop and the Workshop on Market, Cooperation and Votes held at the Carlos III de Madrid University. Thanks to Charles Sprenger, Isabel Trevino and Hernán Bejarano for their comments; and Paul Feldman for his feedback. We thank the anonymous referees for their valuable feedback and constructive suggestions, which greatly contributed to improving the quality of this article. This research was supported by Agencia Andaluza de Cooperación Internacional para el Desarrollo (2021UI004) and the Spanish Ministry of Economy and Competitiveness (PID2021-126892NB-100). Adriana Gaviria acknowledges financial support from EnCon EULAC-2022 (PCI 2023-143420) and Gladis Gonzales from SustainWell (HORIZON-101095175).

References

Adelman, Melissa A and Miguel Szekely (2016), "School dropout in Central America: An overview of trends, causes, consequences, and promising interventions." *World Bank Policy Research Working Paper*.

Alan, Sule, Ceren Baysan, Mert Gumren, and Elif Kubilay (2021), "Building social cohesion in ethnically mixed schools: An intervention on perspective taking." *The Quarterly Journal of Economics*, 136, 2147–2194.

Alfonso, A, P Brañas-Garza, D Jorrat, B Prissé, and MJ Vazquez (2024), "The baking of preferences throughout high school." *mimeo*.

Alfonso, Antonio, Pablo Brañas-Garza, Isabelle Brocas, Juan D Carrillo, Juan B González, Maria Jose Vazquez, et al. (2023b), "The development of rationality in games with hidden information."

Alfonso, Antonio, Pablo Brañas-Garza, Diego Jorrat, Pablo Lomas, Benjamin Prissé, Mónica Vasco, and María J. Vázquez-De Francisco (2023a), "The adventure of running experiments with teenagers." *Journal of Behavioral and Experimental Economics*, 106, 102048, URL https://www.sciencedirect.com/science/article/pii/S2214804323000745.

Amador-Hidalgo, Luis, Pablo Brañas-Garza, Antonio M Espín, Teresa García-Muñoz, and Ana Hernández-Román (2021), "Cognitive abilities and risk-taking: Errors, not preferences." *European Economic Review*, 134, 103694.

Bandura, Albert, Claudio Barbaranelli, Gian Vittorio Caprara, and Concetta Pastorelli (2001), "Self-efficacy beliefs as shapers of children's aspirations and career trajectories." *Child development*, 72, 187–206.

Bettinger, Eric and Robert Slonim (2007), "Patience among children." *Journal of Public Economics*, 91, 343–363.

Bonan, Jacopo, Sergiu Burlacu, and Arianna Galliera (2023), "Prosociality in variants of the dictator game: Evidence from children in El Salvador." *Journal of Behavioral and Experimental Economics*, 107, 102084.

Brañas-Garza, Pablo, Lorenzo Estepa, Diego Jorrat, Victor Orozco, and Ericka Gabriela Rascon Ramirez (2021), "To pay or not to pay: Measuring risk preferences in lab and field." *Judgment and Decision Making*, 16, 1290 – 1313.

Brañas-Garza, Pablo, Diego Jorrat, Antonio M Espín, and Angel Sánchez (2023), "Paid and brocasthetical time preferences are the same: Lab, field and online evidence." *Experimental Economics*, 26, 412–434.

Brañas-Garza, Pablo, Praveen Kujal, and Balint Lenkei (2019), "Cognitive reflection test: Whom, how, when." *Journal of Behavioral and Experimental Economics*, 82, 101455.

Breig, Zachary and Paul Feldman (2024), "Revealing risky mistakes through revisions." *Journal of Risk and Uncertainty*, 1–28.

Brocas, Isabelle and Juan D Carrillo (2020a), "The evolution of choice and learning in the two-person beauty contest game from kindergarten to adulthood." *Games and Economic Behavior*, 120, 132–143.

Brocas, Isabelle and Juan D. Carrillo (2020b), "Introduction to special issue "Understanding Cognition and Decision Making by Children." Studying decision-making in children: Challenges and opportunities." *Journal of Economic Behavior & Organization*, 179, 777–783.

Cárdenas, Juan-Camilo, Anna Dreber, Emma Von Essen, and Eva Ranehill (2012), "Gender differences in competitiveness and risk taking: Comparing children in Colombia and Sweden." *Journal of Economic Behavior & Organization*, 83, 11–23.

Castillo, Marco (2020), "Negative childhood experiences and risk aversion: evidence from children exposed to domestic violence."

Castillo, Marco, Paul J Ferraro, Jeffrey L Jordan, and Ragan Petrie (2011), "The today and tomorrow of kids: Time preferences and educational outcomes of children." *Journal of Public Economics*, 95, 1377–1385.

Charness, Gary and Nir Chemaya (2023), "Repeated experience and consistent risk preferences." *Economics Letters*, 233, 111375.

Chew, Soo Hong, Bin Miao, Qiang Shen, and Songfa Zhong (2022), "Multiple-switching behavior in choice-list elicitation of risk preference." *Journal of Economic Theory*, 204, 105510, URL https://www.sciencedirect.com/science/article/pii/S0022053122001004.

Coller, Maribeth and Melonie B Williams (1999), "Eliciting individual discount rates." *Experimental Economics*, 2, 107–127.

Delavande, Adeline and Hans-Peter Kohler (2009), "Subjective expectations in the context of HIV/AIDS in Malawi." *Demographic Research*, 20, 817.

Dinarte-Diaz, Lelys and Pablo Egana-delSol (2023), "Preventing Violence in the Most Violent Contexts: Behavioral and Neurophysiological Evidence from El Salvador." *Journal of the European Economic Association*, jvad068.

Ert, Eyal and Ernan Haruvy (2017), "Revisiting risk aversion: Can risk preferences change with experience?" *Economics Letters*, 151, 91–95.

Espinosa, María Paz and Lara Ezquerra (2022), "The effect of domain and framing on elicited risk aversion." *Plos one*, 17, e0267696.

Exadaktylos, Filippos, Antonio M Espín, and Pablo Branas-Garza (2013), "Experimental subjects are not different." *Scientific Reports*, 3, 1213.

Fan, Chinn-Ping (2000), "Teaching children cooperation—An application of experimental game theory." *Journal of Economic Behavior & Organization*, 41, 191–209.

Fehr, Ernst, Daniela Glätzle-Rützler, and Matthias Sutter (2013), "The development of egalitarianism, altruism, spite and parochialism in childhood and adolescence." *European Economic Review*, 64, 369–383.

Filippin, Antonio and Paolo Crosetto (2016), "A reconsideration of gender differences in risk attitudes." *Management Science*, 62, 3138–3160.

Frederick, Shane (2005), "Cognitive reflection and decision making." *Journal of Economic Perspectives*, 19, 25–42.

Harrison, Glenn W., Steven J. Humphrey, and Arjan Verschoor (2009), "Choice under Uncertainty: Evidence from Ethiopia, India and Uganda." *The Economic Journal*, 120, 80–104, URL https://doi.org/10.1111/j.1468-0297.2009.02303.x.

Henrich, Joseph, Steven J Heine, and Ara Norenzayan (2010), "The weirdest people in the world?" *Behavioral and brain sciences*, 33, 61–83.

Holt, Charles A and Susan K Laury (2002), "Risk aversion and incentive effects." *The American Economic Review*, 92, 1644–1655.

Instituto de Medicina Legal (2021), "Datos de homicidios en 2021." URL https://transparencia.oj.gob.sv/descargar/3/20791/HOMICIDIOS%20A%25C3%2591O%25202021/11-01-2023. Accessed: 2023-06-11.

Jacobson, Sarah and Ragan Petrie (2009), "Learning from mistakes: What do inconsistent choices over risk tell us?" *Journal of risk and uncertainty*, 38, 143–158.

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

Khattab, Nabil (2015), "Students' aspirations, expectations and school achievement: What really matters?" *British educational research journal*, 41, 731–748.

La Prensa Gráfica (2021), "Cuántos homicidios han ocurrido en tu municipio año por año en la última década." URL https://www.laprensagrafica.com/elsalvador/html. Accessed: 2023-06-11.

Lührmann, Melanie, Marta Serra-Garcia, and Joachim Winter (2018), "The impact of financial education on adolescents' intertemporal choices." *American Economic Journal: Economic Policy*, 10, 309–32.

Moreira, Bruno, Raul Matsushita, and Sergio Da Silva (2010), "Risk seeking behavior of preschool children in a gambling task." *Journal of Economic Psychology*, 31, 794–801.

Munro, Alistair, Yuki Tanaka, et al. (2014), "Risky rotten kids: an experiment on risk attitudes amongst adolescents in rural Uganda." Technical report, National Graduate Institute for Policy Studies.

Paluck, Elizabeth Levy, Hana Shepherd, and Peter M Aronow (2016), "Changing climates of conflict: A social network experiment in 56 schools." *Proceedings of the National Academy of Sciences*, 113, 566–571.

Samara, Muthanna, Bruna Da Silva Nascimento, Aiman El-Asam, Sara Hammuda, and Nabil Khattab (2021), "How can bullying victimisation lead to lower academic achievement? A systematic review and meta-analysis of the mediating role of cognitive-motivational factors." *International journal of environmental research and public health*, 18, 2209.

Starkey, Fenella, Suzanne Audrey, Jo Holliday, Laurence Moore, and Rona Campbell (2009), "Identifying influential young people to undertake effective peer-led health promotion: the example of A Stop Smoking In Schools Trial (ASSIST)." *Health education research*, 24, 977–988.

Sutter, Matthias, Martin G Kocher, Daniela Glätzle-Rützler, and Stefan T Trautmann (2013), "Impatience and uncertainty: Experimental decisions predict adolescents' field behavior." *The American Economic Review*, 103, 510–31.

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

Tymula, Agnieszka (2019), "An experimental study of adolescent behavior under peer observation: Adolescents are more impatient and inconsistent, not more risk-taking, when observed by peers." *Journal of Economic Behavior & Organization*, 166, 735–750.

Ugur, Zeynep B (2021), "Does self-control foster generosity? Evidence from ego depleted children." *Journal of Behavioral and Experimental Economics*, 90, 101652.

Valente, Thomas W, Anamara Ritt-Olson, Alan Stacy, Jennifer B Unger, Janet Okamoto, and Steve Sussman (2007), "Peer acceleration: effects of a social network tailored substance abuse prevention program among high-risk adolescents." *Addiction*, 102, 1804–1815.

Vasco, M and MJ Vazquez (2023), "Holt Laury as a gumball machine." *mimeo*.

Appendix

A. Experimental tasks

This section describes the tasks as they were presented to subjects. In schools 1 to 8, they completed the tasks in the following the order: truck + CRT + Fin + gumball + Delavande. In schools 9 to 12, the order of the truck and gumball tasks was inverted.

A.1. Truck

Screen 1.
Throughout the next 6 screens, you will have to make 6 decisions (one per screen) about **how you want to receive a hypothetical amount of money.** Your task is to choose which option you prefer, knowing that if you choose one you will receive the money **tomorrow**, and if you choose the other you will receive it **next week**. As you progress through the decisions, the amount of money you will receive for waiting will get **larger and larger**.

Figure A.1: *Screen 2*

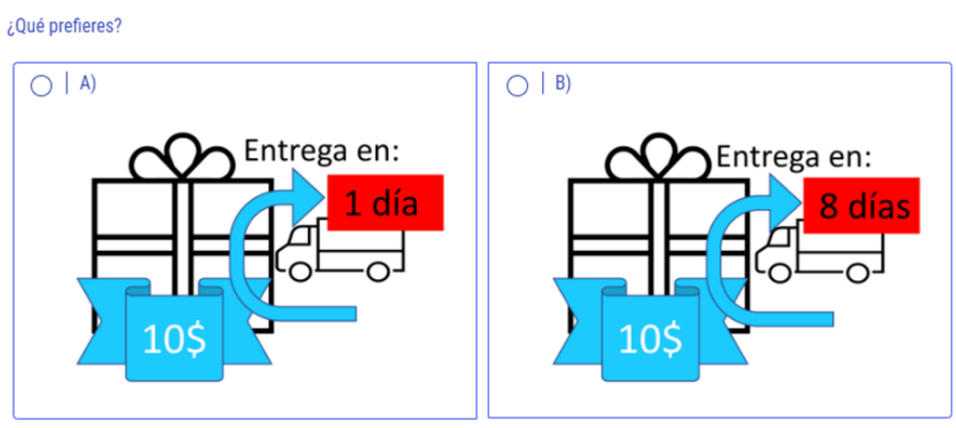


Figure A.2: *Screen 3*

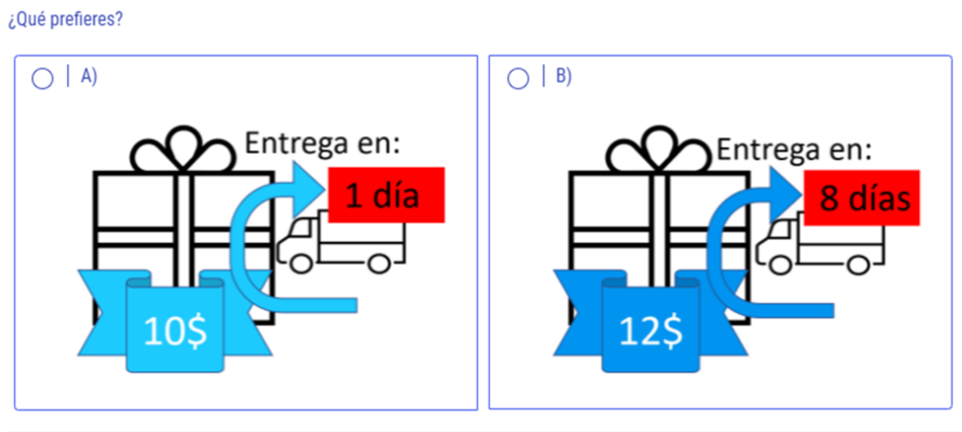


Figure A.3: *Screen 4*

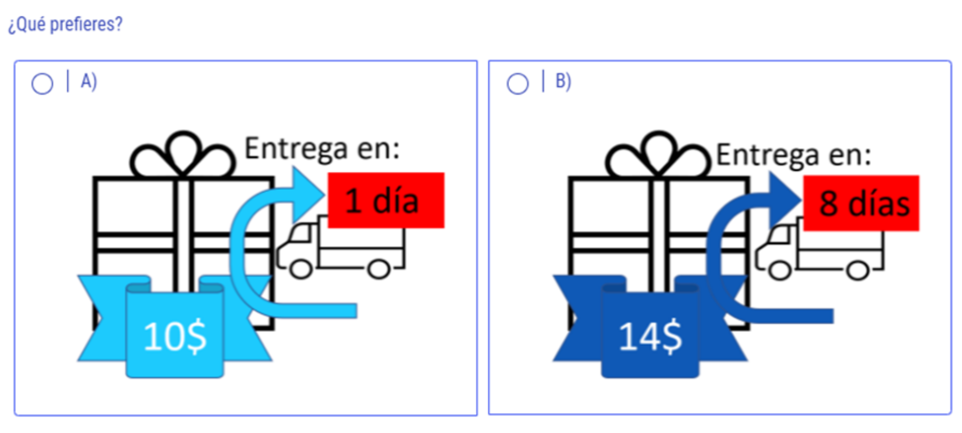


Figure A.4: *Screen 5*

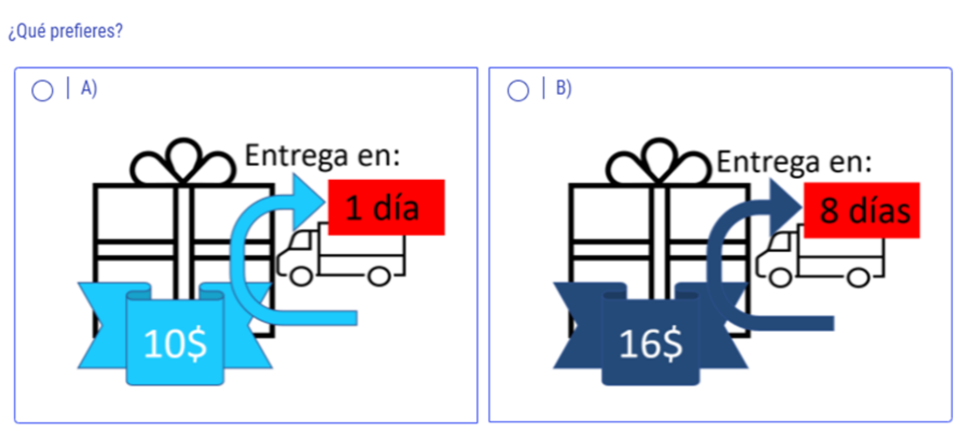


Figure A.5: *Screen 6*

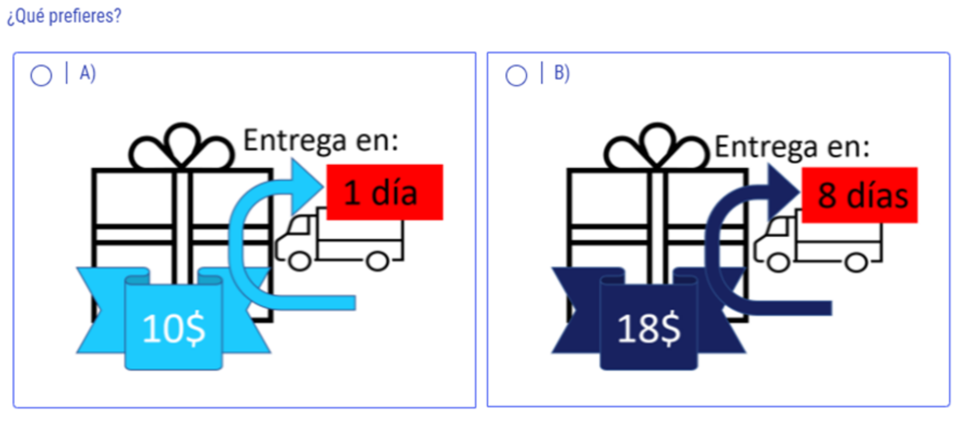
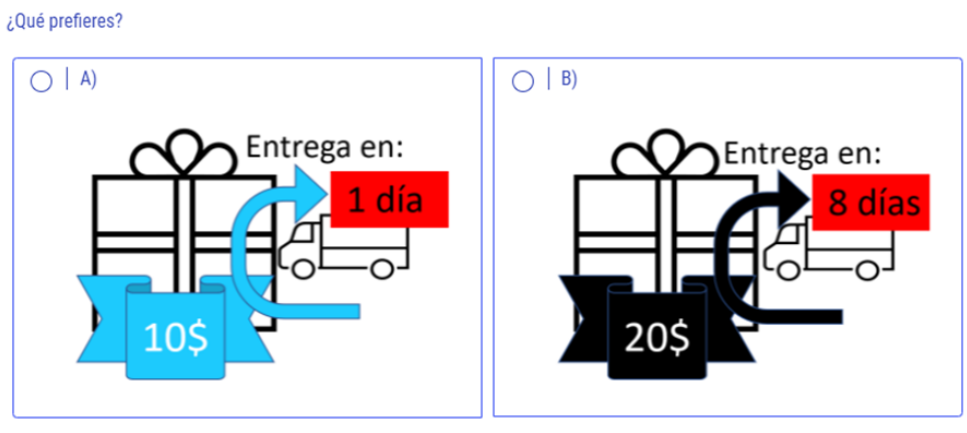


Figure A.6: *Screen 7*



A.2. Gumball task

Screen 1.
In this task you are going to make decisions about **probabilities**. Probability is the possibility of something happening, that is, how someone is sure what is going to happen.
In each decision **you will have to choose between two options** where you can win fictitious money: in both you will win, but you can be luckier and win more, or you can be less lucky and win less. **The probabilities of winning change** from one decision to another.
Below, you are presented with **6 different questions**. Your task is to choose option **A)** or option **B)** in all questions.
To complete the task, proceed to the **next screen**.

Figure A.7: Screen 2

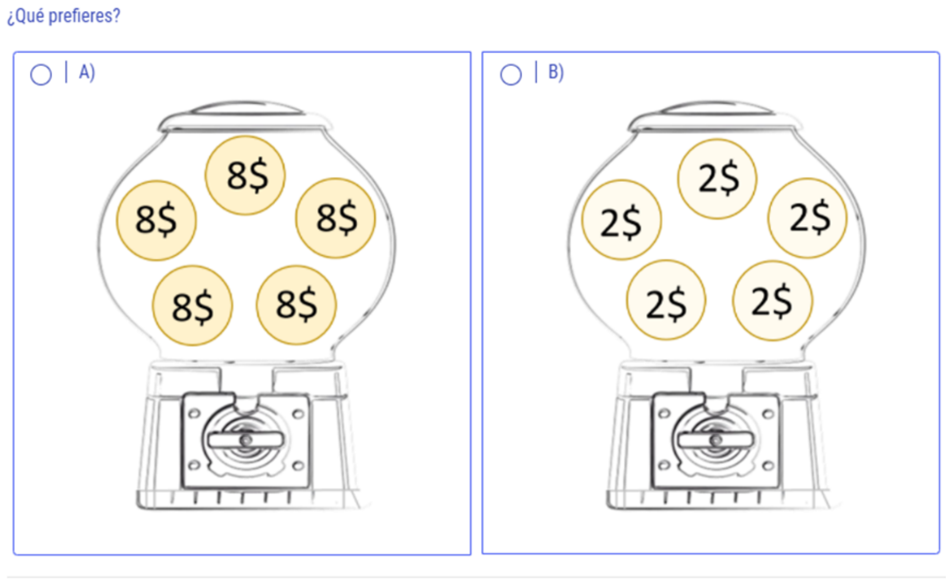


Figure A.8: Screen 3

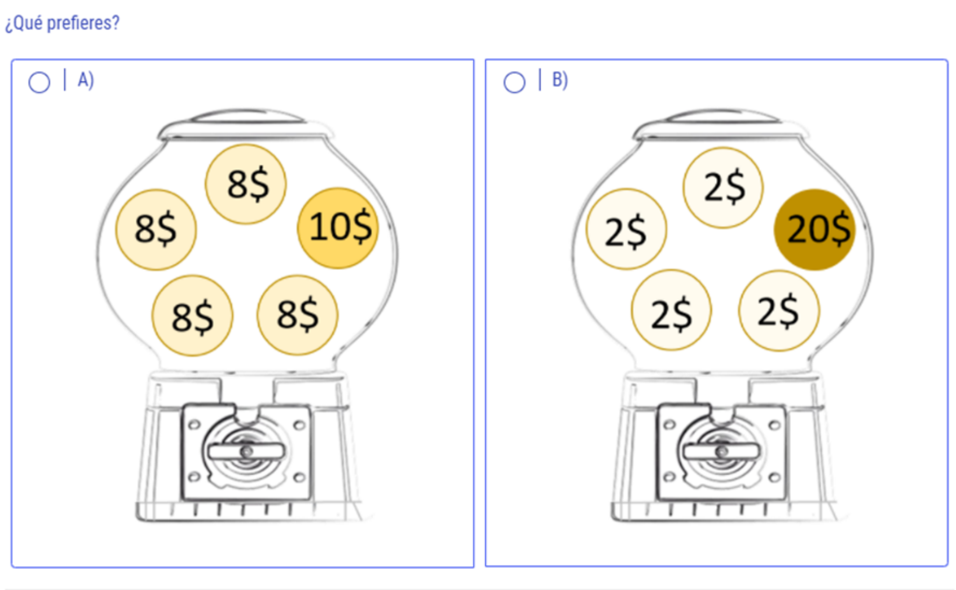


Figure A.9: Screen 4

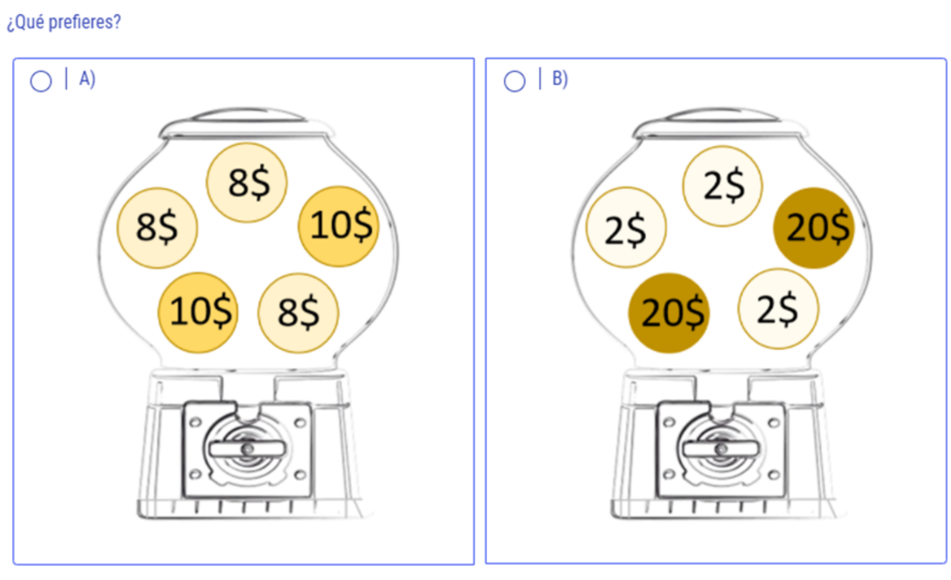


Figure A.10: Screen 5

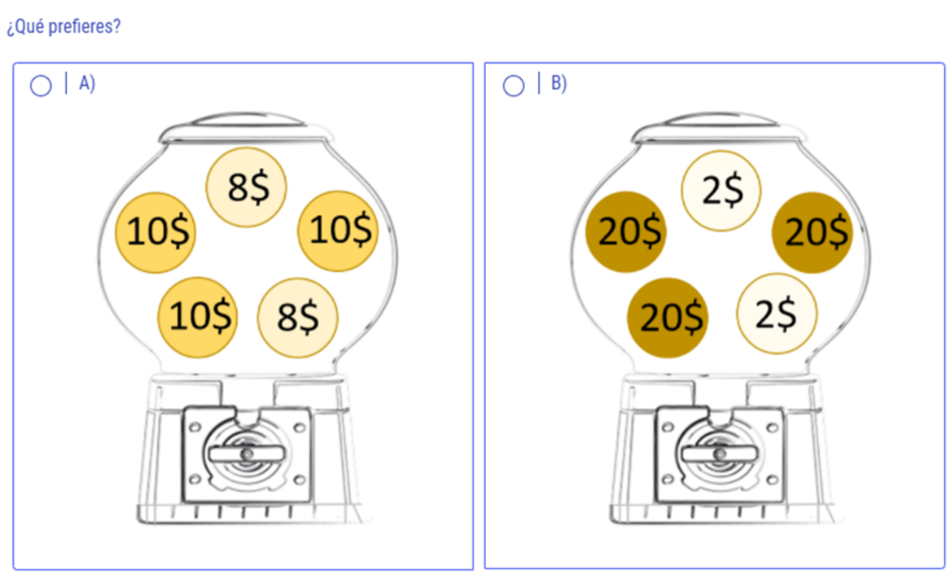


Figure A.11: Screen 6

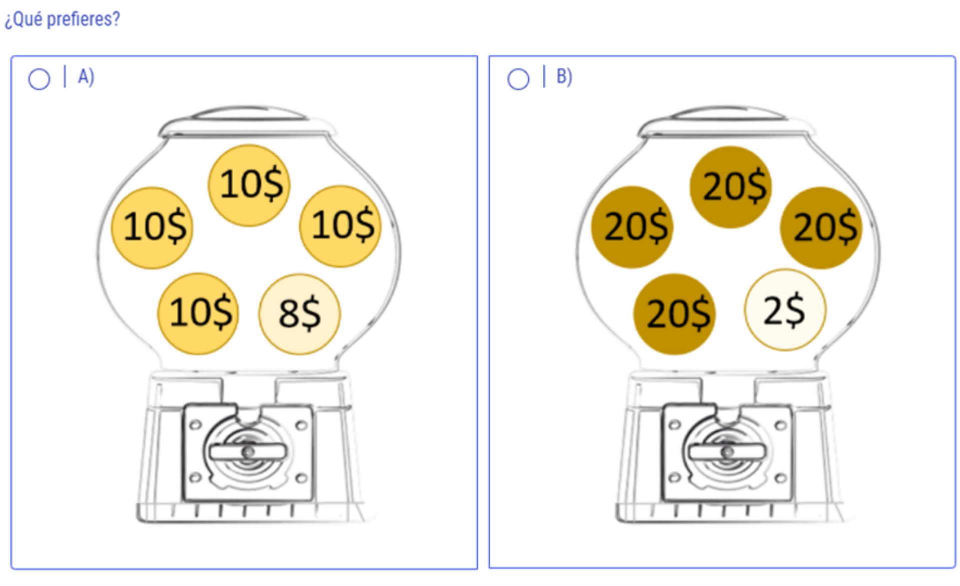
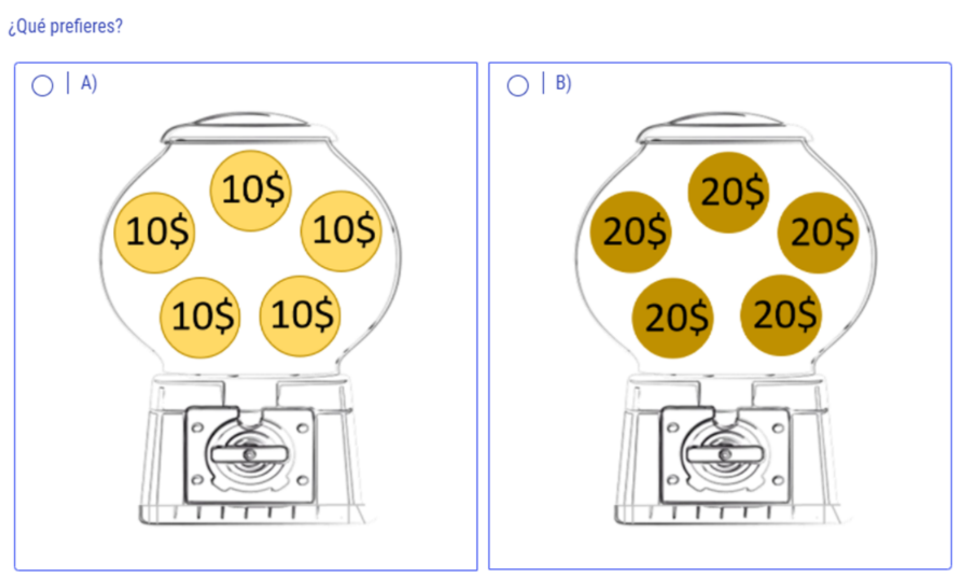


Figure A.12: Screen 7



A.3. Cognitive Reflection Test

Screen 1.
Please answer the following questions. **There are** correct and incorrect answers to these questions.

Screen 2.
Randomly ordered questions:

1. In a library, the number of books doubles every month. If the library takes 48 months to fill, how long will it take to fill it halfway? Indicate with a number. (reflective: 47; intuitive: 24).
2. If you are running a race and you pass the person in second place, what place are you in? Indicate with a number. For example: 1 (first), 2 (second), etc. (reflective: second; intuitive: first).
3. The father of Emilia has 3 daughters. The first two are named April and May. What is the name of the third daughter? (reflective: Emilia; intuitive: June).

A.4. Financial abilities

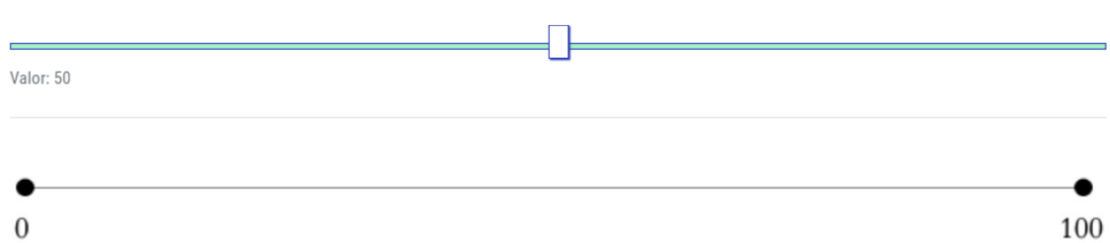
Screen 1.
NOTE: In case you **do NOT know how** to answer or you think this question is not for you, please answer with this question is not for you, answer with: **00**

1. If there are 5 people who possess the winning ticket of a lottery and the prize to be shared is 2 million dollars, how much money will each person receive? Indicate with numbers without periods or commas. (correct: 400000).
2. Imagine that you have $100 in a savings account. The account accumulates an interest rate of 10% by year. How much money will you have in the account after 2 years? (correct: 121).
3. Imagine you have $100 in a savings account and the interest rate you earn on the savings is 2% by year. If you keep the money in the account for 5 years, how much money will you have at the end of 5 years? (This question had options to select as an answer: Less than $102, Exactly $102, More than $102 or 00. I don't know) (correct: More than 102).

A.5. Delavande

Note that participants in schools 1 to 8 responded to the task by entering a number, while in schools 9 to 12 half of the sample entered numbers and the other half moved a slider (see Figure A.13). The instructions and questions were maintained in both versions.

Figure A.13: *Slider*



Screen 1.
In the next screen, we **will ask you** about the **probability** that certain things could happen. In each question, you should **write down** how **likely you think** it is that these things will happen in reality.
If you think something will happen for **sure**, you should write **100**. If you are fairly but not completely sure, write a number close to 100 (but less than 100). If you think it is something that will **never** happen, write **0**.

If you think it is unlikely, write a number close to 0 (but greater than 0). If you think it is **equally likely** to happen or not to happen, write **50**.

Screen 2.
EXAMPLES:
- If it is a sunny day and I ask you how likely it is that it will be sunny within 1 hour, you would typically say 100 or numbers very close to 100.
- If it is 12 o'clock at night, the most reasonable thing to say is that the probability that it will be sunny in 1 hour is 0.
- If I flip a coin, the probability that it will come up heads is 50.

Screen 3.
(First question format schools 1 to 8: how likely from 0 to 100 do you think it is possible to GET the GREEN apple?)

1. Imagine I have a basket with 5 APPLES: 1 GREEN and 4 RED. If I ask you to pick ONE of the apples WITHOUT LOOKING inside the basket, how likely is it that you will get the GREEN apple from 0 to 100?
2. Imagine I have a basket with 10 APPLEs: 1 GREEN and 9 RED. If I ask you to pick ONE of the apples WITHOUT LOOKING inside the basket, how likely is it that you will get the GREEN apple from 0 to 100?
3. How likely are you to EAT RICE in THE NEXT WEEK from 0 to 100 (including today)?
4. How likely are you to EAT RICE in THE NEXT MONTH from 0 to 100 (including today)?

Screen 4.

5. How likely is it that you will NOT ATTEND your school any time in the NEXT MONTH from 0 to 100 (including today)?
6. How likely is it that you will take a bath at least once during the NEXT MONTH from 0 to 100 (including today)?

Screen 5.[25]

(First question format center 1 to 8: How likely from 0 to 100 is it that you will still be STUDYING in 1 YEAR?)

7. How likely are you to CONTINUE your STUDIES next year from 0 to 100?
9. How likely are you to go to UNIVERSITY from 0 to 100?

[25] Question 8 is not of interest for this study.

B. Distribution of individual abilities

Table B.1: *Cognitive and Probability abilities*

Index	CogAb						ProbAb					
	CRT + FinAb		CRT		FinAb		Accuracy + No violations		Accuracy		No violations	
	n	Cumul%	n	Cumul%	n	Cumul%	n	Cumul%	n	Cumul%	n	Cumul%
0	450	20.02%	698	31.05%	1,287	57.25%	139	6.18%	713	31.72%	196	8.72%
1	761	53.87%	858	69.22%	835	94.40%	440	25.76%	791	66.90%	1,035	54.76%
2	613	81.14%	674	99.20%	124	99.91%	612	52.98%	533	90.61%	1,017	100%
3	341	96.31%	18	100%	2	100%	520	76.11%	139	36.80%		
4	76	99.69%					350	91.68%	72	100%		
5	7	100%					115	96.8%				
6	0	100%					72	100%				
Total	2,248						2,248					

Table B.2: *School performance*

Self-reported school score	n	Cum. %
0	117	5.2%
1	1,061	52.4%
2	501	74.7%
3	266	86.5%
4	109	91.4%
5	86	95.2%
6	14	95.8%
7	38	97.5%
8	14	98.1%
9	13	98.8%
10	14	99.3%
11	12	99.9%
12	3	100.00
Total	2,248	

C. Robustness checks

In this section we present results from the following robustness checks[26]:

- *Adjustment in the standard errors*. In Table C.1 we replicate Table 3 with clustering at the individual level to account for potentially correlated decisions across tasks. Note that in Table 3 we included clustering at the class level. We found that the results described in the paper are substantially the same with both kinds of clustering.
- *Numerical and non-numerical dimensions of $CogAb$*. Given that: (i) our measure of cognitive abilities $CogAb$ is constructed by adding subjects' scores in two different tests: the CRT, which reflects their ability to think deliberately and carefully, and the and financial abilities test ($FinAb$), which measures analytical skills, particularly in financial maths; and (ii) provided that reflection (CRT) differs from numeracy ($FinAb$), a valid question is whether math literacy is driving the inconsistencies described in the paper. In Table C.2 we compare the overall effect of $CogAb$ (Models 1-3) with its numerical (Models 7-9) and non-numerical dimensions (Models 4-6) explored individually. Our findings suggest that both measures have similar effects in terms of the magnitude of the coefficient and statistical significance. Therefore, we do not find evidence that math literacy is the only driver of the inconsistencies described in the paper, as reflection seems to play a similar role.
- *Accuracy and consistency dimensions of $ProbAb$*. Given that our measure of Probability abilities $ProbAb$ is composed of the number of correct probability estimations ($accuracy$) and the consistent decisions between sets ($noviolations$), in Table C.3 we compare the overall effect of $ProbAb$ (Models 1-3) with its $accuracy$ (Models 4-6) and $noviolations$ dimensions (Models 7-9) explored individually. Regarding inconsistencies in $time$ preferences, we found that the effect of $ProbAb$ is mainly driven by $accuracy$, as the $noviolations$ coefficient is not statistically significant for this task (Model 7). For $risk$, both dimensions are statistically significant and effect sizes are similar (Models 5 and 8).

[26]We thank our reviewers for suggesting these checks.

Table C.1: *Determinants of inconsistency (clustering at the individual level)*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Vars.	*Time*	*Risk*	*Both*	*Time*	*Risk*	*Both*	*Time*	*Risk*	*Both*
Female	-0.006	0.057***	0.026	-0.008	0.047**	0.020	-0.011	0.043**	0.016
	(0.021)	(0.021)	(0.016)	(0.021)	(0.021)	(0.016)	(0.021)	(0.021)	(0.016)
Year	-0.011	-0.030***	-0.020***	-0.010	-0.024**	-0.017**	-0.008	-0.020**	-0.014*
	(0.010)	(0.010)	(0.007)	(0.010)	(0.010)	(0.007)	(0.010)	(0.010)	(0.007)
CogAb				-0.008	-0.032***	-0.020***			
				(0.010)	(0.010)	(0.008)			
ProbAb							-0.020**	-0.052***	-0.036***
							(0.008)	(0.008)	(0.006)
School 1	0.017	0.029	0.023	0.015	0.021	0.018	0.008	0.005	0.007
	(0.050)	(0.049)	(0.036)	(0.050)	(0.049)	(0.036)	(0.050)	(0.049)	(0.035)
School 2	0.059	0.021	0.040	0.053	-0.003	0.025	0.042	-0.024	0.009
	(0.060)	(0.060)	(0.047)	(0.061)	(0.061)	(0.047)	(0.060)	(0.059)	(0.047)
School 3	0.041	0.099*	0.070*	0.037	0.083	0.060	0.028	0.065	0.047
	(0.053)	(0.052)	(0.039)	(0.053)	(0.052)	(0.039)	(0.053)	(0.052)	(0.039)
School 4	-0.061	0.030	-0.015	-0.061	0.027	-0.017	-0.065	0.018	-0.023
	(0.048)	(0.049)	(0.036)	(0.048)	(0.049)	(0.036)	(0.048)	(0.049)	(0.036)
School 5	-0.098*	0.182***	0.042	-0.105*	0.156***	0.026	-0.116*	0.134**	0.009
	(0.059)	(0.059)	(0.044)	(0.060)	(0.060)	(0.045)	(0.061)	(0.061)	(0.046)
School 6	0.029	0.123*	0.076	0.025	0.107	0.066	0.016	0.088	0.052
	(0.077)	(0.073)	(0.057)	(0.077)	(0.073)	(0.057)	(0.077)	(0.071)	(0.057)
School 7	-0.043	0.136***	0.046	-0.046	0.123***	0.038	-0.058	0.095**	0.019
	(0.039)	(0.038)	(0.029)	(0.039)	(0.038)	(0.029)	(0.039)	(0.038)	(0.029)
School 8	-0.040	-0.031	-0.035	-0.042	-0.037	-0.039	-0.047	-0.048	-0.047
	(0.046)	(0.046)	(0.035)	(0.046)	(0.046)	(0.035)	(0.046)	(0.046)	(0.035)
School 9	0.026	0.071	0.048	0.022	0.057	0.040	0.014	0.040	0.027
	(0.045)	(0.044)	(0.033)	(0.045)	(0.044)	(0.034)	(0.044)	(0.043)	(0.033)
School 10	0.000	0.074	0.037	-0.004	0.059	0.028	-0.008	0.053	0.022
	(0.047)	(0.046)	(0.035)	(0.047)	(0.046)	(0.035)	(0.047)	(0.046)	(0.035)
School 12	0.029	0.079**	0.054*	0.028	0.077**	0.053*	0.026	0.071*	0.049*
	(0.037)	(0.037)	(0.029)	(0.037)	(0.037)	(0.029)	(0.037)	(0.037)	(0.028)
Constant	0.428***	0.617***	0.522***	0.440***	0.663***	0.551***	0.478***	0.749***	0.614***
	(0.079)	(0.079)	(0.059)	(0.081)	(0.081)	(0.060)	(0.081)	(0.081)	(0.061)
Observations	2,248	2,248	2,248	2,248	2,248	2,248	2,248	2,248	2,248
R-squared	0.009	0.023	0.016	0.010	0.028	0.019	0.012	0.044	0.033

Robust standard errors in parentheses (clustering at the individual level)
*** p<0.01, ** p<0.05, * p<0.1

Table C.2: *Determinants of inconsistency ($CogAb$, CRT and $FinAb$)*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Vars.	*Time*	*Risk*	*Both*	*Time*	*Risk*	*Both*	*Time*	*Risk*	*Both*
Female	-0.008	0.047**	0.020	-0.007	0.052**	0.023	-0.008	0.051**	0.022
	(0.021)	(0.021)	(0.014)	(0.021)	(0.021)	(0.015)	(0.021)	(0.020)	(0.014)
Year	-0.010	-0.024***	-0.017**	-0.011	-0.026***	-0.019***	-0.010	-0.026***	-0.018***
	(0.009)	(0.009)	(0.007)	(0.009)	(0.009)	(0.007)	(0.008)	(0.009)	(0.007)
CogAb	-0.008	-0.032***	-0.020***						
	(0.010)	(0.010)	(0.007)						
CRT				-0.006	-0.035**	-0.021**			
				(0.013)	(0.014)	(0.009)			
FinAb							-0.014	-0.038**	-0.026**
							(0.016)	(0.016)	(0.012)
School 1	0.015	0.021	0.018	0.016	0.022	0.019	0.016	0.027	0.022
	(0.053)	(0.048)	(0.037)	(0.053)	(0.050)	(0.037)	(0.053)	(0.048)	(0.036)
School 2	0.053	-0.003	0.025	0.056	0.003	0.029	0.056	0.012	0.034
	(0.083)	(0.052)	(0.049)	(0.083)	(0.052)	(0.049)	(0.083)	(0.053)	(0.049)
School 3	0.037	0.083	0.060	0.039	0.084	0.061	0.040	0.096	0.068*
	(0.045)	(0.058)	(0.039)	(0.044)	(0.057)	(0.038)	(0.044)	(0.058)	(0.038)
School 4	-0.061	0.027	-0.017	-0.061	0.026	-0.018	-0.060	0.031	-0.014
	(0.043)	(0.042)	(0.028)	(0.043)	(0.043)	(0.028)	(0.042)	(0.042)	(0.027)
School 5	-0.105***	0.156***	0.026	-0.102***	0.160***	0.029	-0.101***	0.175***	0.037
	(0.038)	(0.056)	(0.032)	(0.038)	(0.057)	(0.033)	(0.037)	(0.056)	(0.031)
School 6	0.025	0.107***	0.066**	0.028	0.113***	0.071**	0.026	0.114***	0.070**
	(0.048)	(0.040)	(0.033)	(0.047)	(0.039)	(0.032)	(0.047)	(0.042)	(0.033)
School 7	-0.046	0.123***	0.038*	-0.044	0.126***	0.041*	-0.044	0.131***	0.043*
	(0.047)	(0.028)	(0.023)	(0.046)	(0.027)	(0.022)	(0.047)	(0.028)	(0.022)
School 8	-0.042	-0.037	-0.039	-0.042	-0.039	-0.040	-0.040	-0.029	-0.035
	(0.045)	(0.044)	(0.037)	(0.046)	(0.044)	(0.037)	(0.045)	(0.044)	(0.037)
School 9	0.022	0.057	0.040	0.024	0.062	0.043	0.023	0.065	0.044
	(0.028)	(0.052)	(0.033)	(0.027)	(0.052)	(0.033)	(0.027)	(0.053)	(0.033)
School 10	-0.004	0.059	0.028	-0.002	0.063*	0.030	-0.002	0.068*	0.033
	(0.038)	(0.037)	(0.024)	(0.038)	(0.036)	(0.023)	(0.038)	(0.037)	(0.023)
School 12	0.028	0.077***	0.053**	0.028	0.075***	0.052**	0.029	0.080***	0.055**
	(0.034)	(0.028)	(0.021)	(0.034)	(0.028)	(0.021)	(0.034)	(0.028)	(0.021)
Constant	0.440***	0.663***	0.551***	0.435***	0.653***	0.544***	0.434***	0.632***	0.533***
	(0.067)	(0.071)	(0.051)	(0.067)	(0.072)	(0.051)	(0.066)	(0.070)	(0.050)
Observations	2,248	2,248	2,248	2,248	2,248	2,248	2,248	2,248	2,248
R-squared	0.010	0.028	0.019	0.009	0.026	0.018	0.010	0.025	0.018

Robust standard errors in parentheses (clustering at the class level)

*** p<0.01, ** p<0.05, * p<0.1

Table C.3: *Determinants of inconsistency* ($ProbAb$, $Accuracy$ *and* $Noviolations$)

Vars.	(1) *Time*	(2) *Risk*	(3) *Both*	(4) *Time*	(5) *Risk*	(6) *Both*	(7) *Time*	(8) *Risk*	(9) *Both*
Female	-0.011	0.043**	0.016	-0.012	0.043**	0.016	-0.006	0.052**	0.023
	(0.021)	(0.020)	(0.014)	(0.021)	(0.020)	(0.014)	(0.021)	(0.021)	(0.014)
Year	-0.008	-0.020**	-0.014**	-0.007	-0.020**	-0.014**	-0.011	-0.026***	-0.018***
	(0.009)	(0.009)	(0.006)	(0.009)	(0.009)	(0.006)	(0.009)	(0.009)	(0.006)
ProbAb	-0.020**	-0.052***	-0.036***						
	(0.008)	(0.008)	(0.006)						
Accuracy				-0.033***	-0.073***	-0.053***			
				(0.011)	(0.010)	(0.008)			
No violations							-0.010	-0.064***	-0.037***
							(0.017)	(0.016)	(0.013)
School 1	0.008	0.005	0.007	0.007	0.008	0.007	0.015	0.020	0.017
	(0.053)	(0.049)	(0.037)	(0.053)	(0.050)	(0.037)	(0.053)	(0.049)	(0.037)
School 2	0.042	-0.024	0.009	0.040	-0.021	0.009	0.056	0.004	0.030
	(0.085)	(0.049)	(0.051)	(0.084)	(0.045)	(0.049)	(0.083)	(0.055)	(0.052)
School 3	0.028	0.065	0.047	0.026	0.066	0.046	0.039	0.087	0.063
	(0.044)	(0.059)	(0.040)	(0.044)	(0.057)	(0.038)	(0.044)	(0.059)	(0.039)
School 4	-0.065	0.018	-0.023	-0.064	0.022	-0.021	-0.062	0.023	-0.019
	(0.042)	(0.044)	(0.028)	(0.042)	(0.045)	(0.028)	(0.043)	(0.042)	(0.028)
School 5	-0.116***	0.134**	0.009	-0.117***	0.141**	0.012	-0.102***	0.160***	0.029
	(0.037)	(0.052)	(0.028)	(0.037)	(0.055)	(0.030)	(0.037)	(0.053)	(0.029)
School 6	0.016	0.088**	0.052	0.017	0.095**	0.056*	0.026	0.105**	0.065*
	(0.048)	(0.042)	(0.034)	(0.048)	(0.041)	(0.033)	(0.047)	(0.042)	(0.034)
School 7	-0.058	0.095***	0.019	-0.060	0.096***	0.018	-0.045	0.121***	0.038
	(0.049)	(0.027)	(0.025)	(0.049)	(0.027)	(0.025)	(0.047)	(0.028)	(0.023)
School 8	-0.047	-0.048	-0.047	-0.046	-0.043	-0.045	-0.042	-0.041	-0.041
	(0.045)	(0.045)	(0.037)	(0.045)	(0.046)	(0.038)	(0.045)	(0.044)	(0.037)
School 9	0.014	0.040	0.027	0.012	0.041	0.026	0.024	0.060	0.042
	(0.027)	(0.050)	(0.031)	(0.028)	(0.051)	(0.032)	(0.027)	(0.051)	(0.031)
School 10	-0.008	0.053	0.022	-0.007	0.057	0.025	-0.002	0.063*	0.031
	(0.038)	(0.036)	(0.023)	(0.038)	(0.036)	(0.023)	(0.037)	(0.036)	(0.023)
School 12	0.026	0.071**	0.049**	0.025	0.070**	0.048**	0.028	0.077***	0.053**
	(0.034)	(0.028)	(0.021)	(0.034)	(0.027)	(0.020)	(0.034)	(0.028)	(0.021)
Constant	0.478***	0.749***	0.614***	0.465***	0.697***	0.581***	0.443***	0.707***	0.575***
	(0.067)	(0.074)	(0.053)	(0.066)	(0.072)	(0.051)	(0.068)	(0.075)	(0.053)
Observations	2,248	2,248	2,248	2,248	2,248	2,248	2,248	2,248	2,248
R-squared	0.012	0.044	0.033	0.014	0.044	0.036	0.010	0.030	0.020

Robust standard errors in parentheses (clustering at the class level)
*** p<0.01, ** p<0.05, * p<0.1

D. Inconsistency types and intensity

The measure of inconsistency in our paper is constructed by considering three types of inconsistent behavior: subjects who make wrong choices in *time, risk* and *both* ($time \cup risk$). Within each domain, inconsistencies are defined as follows:

Time: refers to multiple switching along the temporal discounting task.

Risk: refers to choosing dominated strategies and/or multiple switching along the risk task.

For simplicity, in the paper we report inconsistency results by considering individuals who made at least one inconsistent choice in the respective domain. In this section, we extend the analysis by examining the *intensity* of inconsistencies, defined as the number of inconsistencies within the same task. Furthermore, considering previous studies have shown the importance of analyzing the differences between different *types* of inconsistency[27], for *risk* we distinguish between *switching* and *dominated* strategies. These are referred to as lack of understanding (hereafter, LoU) and inattention, respectively (see Amador-Hidalgo et al. (2021)).

In the first row of Table D.1 we report the proportion of subjects who made at least one inconsistent choice in each task (as presented in Figure 2 in the main text). In the lower panels, we break down these results to assess the *intensity* of inconsistency for both tasks and the inconsistency *types* for *risk*.

We first analyze *intensity*. Among the inconsistent subjects in *time* (45% of total sample), most of them made only one inconsistent choice (35% of total sample). The proportion of subjects who made two or three inconsistent choices is much smaller (9.7% and 0.4% of total sample, respectively). In Figure D.1 we explore this result by gender, finding that the *intensity* of inconsistency in *time* is almost identical for boys and girls. Models 1, 4 and 7 in Table D.2 show that there are no statistically significant gender differences.

The *intensity* of inconsistency is higher for *risk*. Table D.1 shows that, among inconsistent subjects (57% of total sample), almost half made only one inconsistent choice (26% of total sample), whereas 20% made two inconsistent choices and 10% made three inconsistencies. Figure D.1 shows that the 5% gender gap in inconsistency in *risk* reported in the main text is explained by girls who made one mistake. Models 2, 5 and 8 in Table D.2 show that the gender difference in the *intensity* of inconsistency in *risk* weakens as we include $CogAb$ to the regression ($p < 0.1$) and disappears when we include $ProbAb$ ($p > 0.1$).

We now turn to inconsistency *types* in *risk*. Table D.1 shows the proportion of subjects exhibiting each *type* of inconsistency, regardless of whether they also exhibited other inconsistencies. As a result, the percentages for all *types* do not sum to 100%. A first important finding is that most inconsistent subjects made multiple types of inconsistencies (30% of total sample). In fact, the proportion of subjects who made only one type of inconsistency is rather low (switching (15%) and dominated[28] (11%)).

In Figure D.2 we see that the only *type* in which we observe gender differences is *switching* (13% of boys versus 18% of girls). Accordingly, Models 3 and 7 in Table D.3 show that girls are more likely to do *switching* ($p < 0.01$). Models 2 and 6 in Table D.3 show that both $CogAb$ and $ProbAb$ have a negative effect on the probability of choosing *dominated* strategies ($p < 0.05$ and $p < 0.01$, respectively). For *switching*, $ProbAb$ has a negative effect ($p < 0.01$) (Model 7), but the effect of $CogAb$ is not significant (Model 3). For multiple types (*Dom+Swb*), both $CogAb$ and $ProbAb$ have a negative and statistically significant effect ($p < 0.01$).

[27]For example, Chew et al. (2022) argue that multiple switching is driven by stochastic preferences or a deliberate preference for randomization.

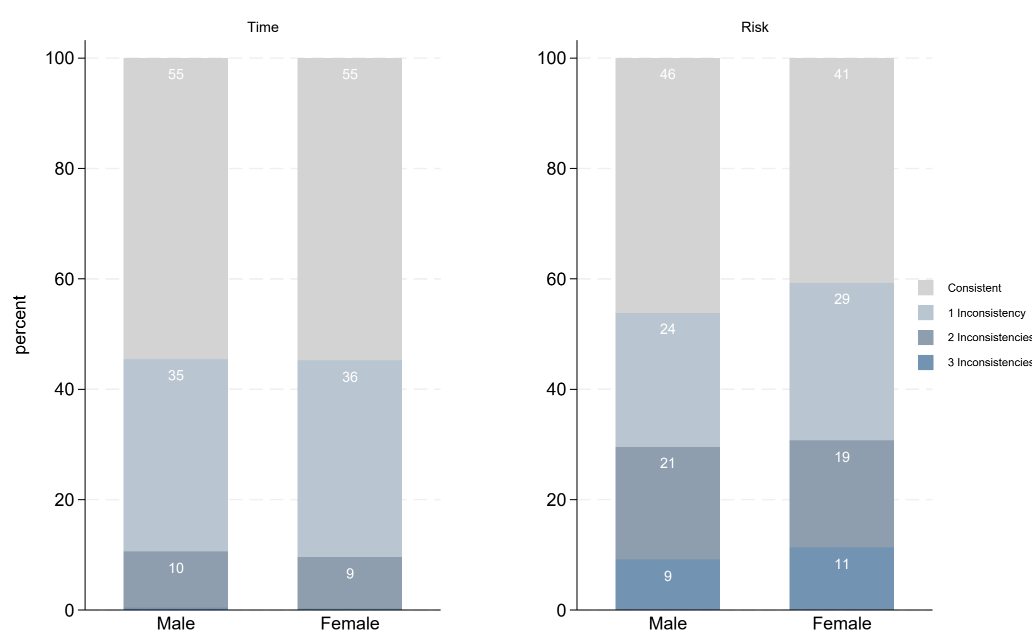
[28]For simplicity, in Figure D.2 we aggregate *LoU* (6.5%) and *Inattention* (4.7%) into the category *Dominated*.

Table D.1: *Inconsistency types and intensity*

	Time	**Risk**	**Both**
At least one inconsistency	45.37%	56.54%	72.78%
Intensity			
1	35.23%	26.38%	27.14%
2	9.74%	19.93%	22.29%
3	0.40%	10.23%	14.59%
4 or more			8.79%
Types			
LoU		27.62%	
Inattention		23.98%	
Switching		45.33%	
Multiple types		30.16%	

Note: All percentages are calculated over total sample. For *types*, each category reflects the proportion of subjects exhibiting that specific type of inconsistency, regardless of them having committed other types as well. Therefore, the percentages reported for types do not sum to 100%. Subjects who exhibited more than one type of inconsistency *Multiple types* are 30.16% over the total sample.

Figure D.1: *Intensity of inconsistency by gender*



Note: The percentage of males and females with 3 inconsistencies in $Time$ is not displayed on the figure as it is very small (0.44% and 0.36%, respectively).

Figure D.2: *Types of inconsistency in risk by gender*

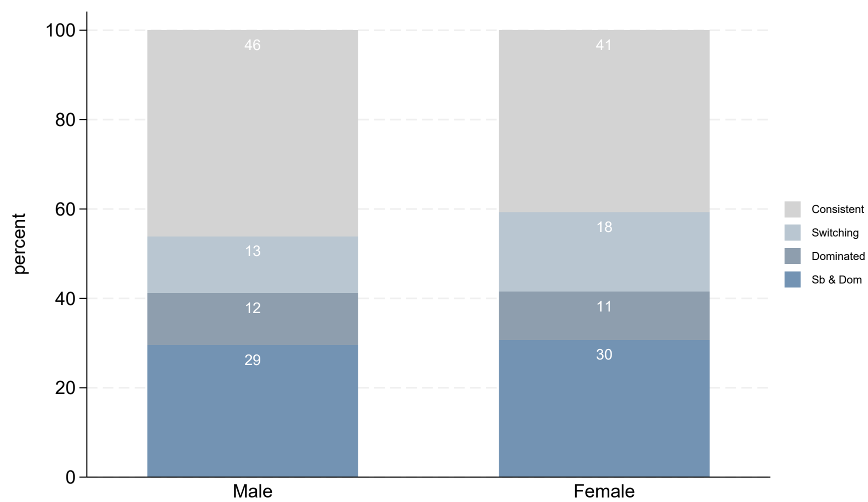


Table D.2: *Intensity of inconsistency*

Vars.	(1) *Time*	(2) *Risk*	(3) *Both*	(4) *Time*	(5) *Risk*	(6) *Both*	(7) *Time*	(8) *Risk*	(9) *Both*
Female	-0.017	0.091**	0.068	-0.018	0.072*	0.047	-0.023	0.067	0.038
	(0.029)	(0.043)	(0.053)	(0.029)	(0.043)	(0.054)	(0.029)	(0.043)	(0.053)
Year	-0.017	-0.061***	-0.076***	-0.016	-0.050**	-0.063**	-0.012	-0.044**	-0.054**
	(0.014)	(0.020)	(0.025)	(0.014)	(0.021)	(0.026)	(0.014)	(0.020)	(0.025)
CogAb				-0.005	-0.060***	-0.068***			
				(0.014)	(0.020)	(0.025)			
ProbAb							-0.025**	-0.089***	-0.112***
							(0.011)	(0.015)	(0.020)
School 1	-0.020	-0.082	-0.095	-0.021	-0.098	-0.113	-0.031	-0.122	-0.146
	(0.065)	(0.096)	(0.119)	(0.066)	(0.096)	(0.119)	(0.065)	(0.096)	(0.117)
School 2	0.043	-0.076	-0.046	0.039	-0.121	-0.096	0.022	-0.152	-0.142
	(0.078)	(0.118)	(0.150)	(0.079)	(0.120)	(0.152)	(0.079)	(0.118)	(0.150)
School 3	0.071	0.073	0.136	0.068	0.042	0.101	0.054	0.015	0.062
	(0.074)	(0.106)	(0.132)	(0.074)	(0.107)	(0.132)	(0.074)	(0.106)	(0.131)
School 4	-0.076	-0.063	-0.148	-0.076	-0.068	-0.155	-0.081	-0.083	-0.173
	(0.067)	(0.098)	(0.120)	(0.067)	(0.099)	(0.120)	(0.067)	(0.097)	(0.120)
School 5	-0.144*	0.226*	0.075	-0.149*	0.177	0.020	-0.167**	0.145	-0.027
	(0.075)	(0.125)	(0.152)	(0.076)	(0.126)	(0.154)	(0.077)	(0.127)	(0.157)
School 6	0.015	0.102	0.102	0.013	0.072	0.067	-0.001	0.043	0.027
	(0.100)	(0.156)	(0.199)	(0.100)	(0.156)	(0.200)	(0.100)	(0.154)	(0.198)
School 7	-0.064	0.158**	0.108	-0.066	0.134*	0.081	-0.083	0.090	0.022
	(0.053)	(0.080)	(0.098)	(0.053)	(0.080)	(0.098)	(0.054)	(0.081)	(0.099)
School 8	-0.035	-0.180**	-0.211*	-0.036	-0.192**	-0.225**	-0.043	-0.209**	-0.248**
	(0.065)	(0.089)	(0.113)	(0.065)	(0.089)	(0.113)	(0.065)	(0.089)	(0.113)
School 9	0.081	0.076	0.120	0.079	0.050	0.091	0.067	0.023	0.054
	(0.065)	(0.092)	(0.115)	(0.065)	(0.093)	(0.115)	(0.065)	(0.091)	(0.113)
School 10	0.032	0.124	0.152	0.029	0.096	0.121	0.022	0.088	0.107
	(0.067)	(0.098)	(0.123)	(0.067)	(0.098)	(0.123)	(0.067)	(0.098)	(0.123)
School 12	0.025	0.187**	0.212**	0.024	0.184**	0.209**	0.021	0.174**	0.196**
	(0.051)	(0.079)	(0.099)	(0.051)	(0.079)	(0.099)	(0.051)	(0.079)	(0.098)
Constant	0.516***	1.204***	1.699***	0.523***	1.291***	1.796***	0.579***	1.428***	1.981***
	(0.113)	(0.164)	(0.204)	(0.115)	(0.166)	(0.207)	(0.116)	(0.168)	(0.210)
Observations	2,248	2,248	2,248	2,248	2,248	2,248	2,248	2,248	2,248
R-squared	0.011	0.019	0.018	0.011	0.023	0.021	0.013	0.033	0.032

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: In models (1), (4) and (7) the dependent variable corresponds to the number of inconsistent choices made in $time$. In models (2), (5) and (8) we add the number of inconsistent choices in $risk$, regardless of the type (dominated or switch-back). In models (3), (6) and (9) we add subjects' inconsistencies in both tasks.

Table D.3: *Determinants of inconsistency types in* $risk$

Vars.	(1) *Any*	(2) *Dom*	(3) *Swb*	(4) *Dom+Swb*	(5) *Any*	(6) *Dom*	(7) *Swb*	(8) *Dom+Swb*
Female	0.047**	0.004	0.088***	0.030	0.043**	0.001	0.077***	0.028
	(0.021)	(0.025)	(0.023)	(0.025)	(0.020)	(0.025)	(0.021)	(0.025)
Year	-0.024***	-0.016*	-0.009	-0.029***	-0.020**	-0.014	-0.003	-0.025***
	(0.009)	(0.009)	(0.013)	(0.009)	(0.009)	(0.009)	(0.013)	(0.009)
CogAb	-0.032***	-0.021**	-0.017	-0.034***				
	(0.010)	(0.010)	(0.011)	(0.012)				
ProbAb					-0.052***	-0.029***	-0.043***	-0.047***
					(0.008)	(0.008)	(0.007)	(0.008)
School 1	0.021	0.101*	0.067	-0.066	0.005	0.091	0.048	-0.075
	(0.048)	(0.057)	(0.050)	(0.046)	(0.049)	(0.056)	(0.051)	(0.047)
School 2	-0.003	0.113*	-0.011	-0.073	-0.024	0.105	-0.035	-0.081
	(0.052)	(0.064)	(0.055)	(0.061)	(0.049)	(0.065)	(0.052)	(0.062)
School 3	0.083	0.130*	0.052	0.048	0.065	0.121*	0.035	0.037
	(0.058)	(0.067)	(0.065)	(0.061)	(0.059)	(0.066)	(0.065)	(0.062)
School 4	0.027	0.110***	0.059	-0.063	0.018	0.103***	0.053	-0.070*
	(0.042)	(0.036)	(0.054)	(0.039)	(0.044)	(0.036)	(0.057)	(0.040)
School 5	0.156***	0.278***	0.069	0.097	0.134**	0.261***	0.031	0.080
	(0.056)	(0.064)	(0.061)	(0.062)	(0.052)	(0.063)	(0.061)	(0.059)
School 6	0.107***	0.213***	0.105	0.037	0.088**	0.205***	0.089	0.026
	(0.040)	(0.050)	(0.071)	(0.056)	(0.042)	(0.055)	(0.074)	(0.055)
School 7	0.123***	0.239***	0.056	0.061*	0.095***	0.224***	0.027	0.042
	(0.028)	(0.043)	(0.047)	(0.032)	(0.027)	(0.043)	(0.047)	(0.031)
School 8	-0.037	0.067***	0.013	-0.121***	-0.048	0.058**	-0.002	-0.128***
	(0.044)	(0.022)	(0.053)	(0.042)	(0.045)	(0.022)	(0.054)	(0.044)
School 9	0.057	0.087*	0.070	0.016	0.040	0.081*	0.053	0.005
	(0.052)	(0.044)	(0.060)	(0.055)	(0.050)	(0.042)	(0.059)	(0.053)
School 10	0.059	0.029	0.035	0.064	0.053	0.021	0.020	0.060
	(0.037)	(0.029)	(0.060)	(0.047)	(0.036)	(0.028)	(0.061)	(0.045)
School 12	0.077***	0.014	0.035	0.095***	0.071**	0.009	0.032	0.091***
	(0.028)	(0.031)	(0.046)	(0.027)	(0.028)	(0.029)	(0.046)	(0.026)
Constant	0.663***	0.226***	0.207**	0.607***	0.749***	0.271***	0.297***	0.672***
	(0.071)	(0.070)	(0.100)	(0.071)	(0.074)	(0.071)	(0.102)	(0.072)
Observations	2,248	1,229	1,318	1,655	2,248	1,229	1,318	1,655
R-squared	0.028	0.059	0.023	0.033	0.044	0.066	0.040	0.045

Robust standard errors in parentheses (clustering at the class level)
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Note: In each model, we compare subjects who committed each type of inconsistency with the consistent group. Therefore, the number of observations varies across models because those committing other types of inconsistencies are not included. For comparison, we include models (1) and (5) in which the dependent variable considers subjects who committed *any* inconsistency, thus using the full sample.

Inconsistency types across waves

For $risk$ we also explored the role of inconsistency *types* across waves. In Table D.4 we report the proportion of subjects who were consistent and who committed each type of inconsistency in the first and second waves. The percentages are calculated over the sub-sample of students who participated in both waves[29] and considering subjects' behavior in wave 1 as denominator. For example, among consistent subjects in wave 1, in wave 2 we find that 64% were consistent, 18% committed LoU, 16% inattention, and 30% switching. Since 30% of subjects committed more than one inconsistency (see Table D.1), percentages in Table D.4 do not sum to 100%.

We focus on the role of inattention and LoU in the transition from consistency to inconsistency and vice versa. In the first case, we could expect that inattention explains why consistent subjects in wave 1 become inconsistent in wave 2. In other words, these subjects understood the task well the first time, but were inattentive the second time. Likewise, LoU in wave 1 could explain why a subject becomes consistent in wave 2, assuming that she understood the task better the second time.

Our data does not suggest that inattention and LoU played a significant role in these transitions. This may be because a high proportion of subjects in our sample exhibited multiple types of inconsistencies (30%, as reported in Table D.1), making it difficult to isolate the role of a specific *type*. Table D.4 shows that, among subjects who transitioned from consistency to inconsistency, the most common *type* of inconsistency was switching (30%), which is almost twice as frequent as LoU and inattention (18% and 16%, respectively). For subjects who transitioned from inconsistency to consistency, the table shows that 30% of those who exhibited LoU became consistent. This proportion is similar for subjects who committed inattention in wave 1 but were consistent in wave 2 (30%), and slightly lower for those who made switching in wave 1 but became consistent in wave 2 (36%).

Table D.4: *Inconsistency types in $risk$ across waves*

		W1			
		Consistent	LoU	Inattention	Switching
W2	Consistent	64.16%	29.72%	30.41%	36.27%
	LoU	17.92%	44.17%	40.64%	35.13%
	Inattention	16.13%	34.72%	36.84%	31.70%
	Switching	29.75%	59.72%	57.31%	54.74%

Note: The percentages do not sum to 100% because each category reflects the proportion of subjects exhibiting that specific type of inconsistency. Subjects who exhibited more than one type of inconsistency are included in multiple categories.

E. Strategies

[29] Recall that a total of 1,960 students from wave 1 repeated the experiment of which 1,322 met the same selection criteria as wave 1.

Replacement strategies and elicitation of preferences

In this section we describe our replacement strategy for switching back decisions and discuss why it involves non-trivial assumptions that could distort the elicitation of subjects' preferences. Recall that the strategy that we adopted replaced subject's switching back decision to the one she had made immediately before. Take the case of a subject whose decision pattern was ABABBB. In this case, the switching back took place in the third decision A and the replacement consisted in changing it to the one she had made immediately before (second decision) B. The underlying assumption of this replacement strategy is that subject's mistake was switching back from B to A in the third decision. However, an alternative interpretation of subject's decision pattern could be that the mistake was choosing B in the second decision. Under this assumption, an alternative replacement strategy would be to replace subject's second decision from a B to an A.

Different assumptions about subject's inconsistent choices have non-trivial implications in the elicitation of their preferences. For *risk* preferences, the implications of the example above are straightforward: the first replacement strategy leads to the new decision pattern ABBBBB and the subject is classified as *risk lover*. Instead, with the alternative replacement, the resulting decision pattern is AAABBB and the subject is classified as *risk averse*. As illustrated in E.1, alternative replacement strategies in *The Gumball* task lead to different classifications that could potentially distort the elicitation of subjects' preferences.

Table E.1: *Replacement strategies and risk preferences*

Decision	Original pattern	Pattern with replacement	Pattern with alternative replacement
1	A	A	A
2	B	B	A
3	A	B	A
4	B	B	B
5	B	B	B
6	B	B	B
Classification	**Inconsistent**	**Lover**	**Averse**

For *time* preferences, the different replacement strategies would not make a difference within the ($Interior$) dummy category. Nonetheless, if we compute the measure of time preferences in terms of the *number* of future allocations ($\#Future$, from 0 to 6 choices of one week later), it obviously matters whether we are changing a B to an A or an A to a B.

F. Background information about schools

In this section, we provide background information on the geographical distribution and performance of the schools included in the sample. Their approximate geographical location is represented on the green dots in the map[30] in Figure F.1. The schools cover 5 of the 14 departments (i.e. states) of the country, three of which are located in the West Zone (Ahuachapán, Santa Ana y Sonsonate), one in the Central Zone (San Salvador), and one in the Eastern Zone (San Miguel).

[30] Accessible at Wikimedia Commons url: https://commons.wikimedia.org/wiki/File:Un-el-salvador.png

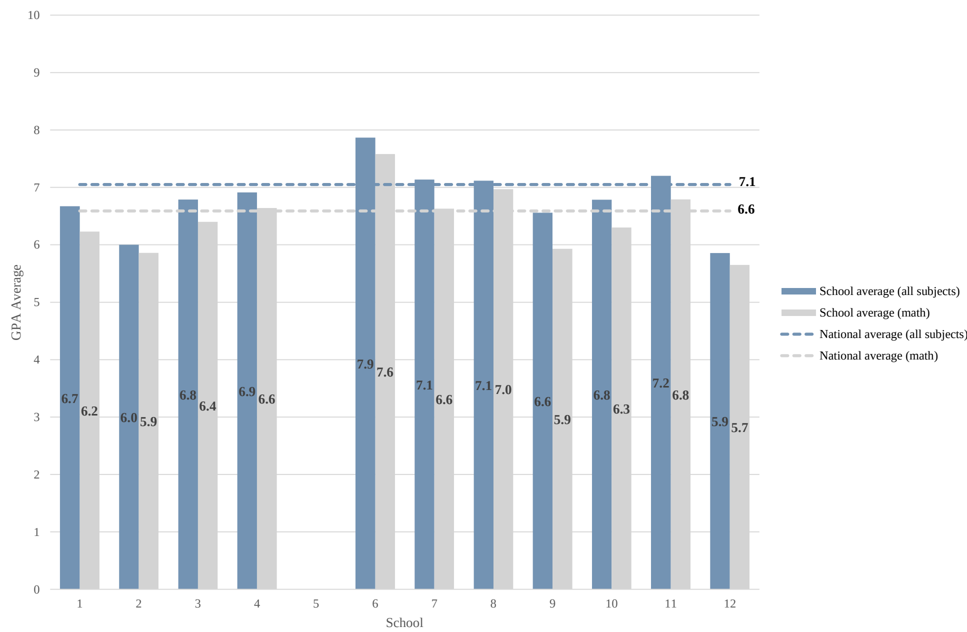
Figure F.1: *Approximate location of schools*

Source: Map derived from a production by the United Nations Cartographic Section (UNCS), in the public domain. Accessible at Wikimedia Commons.

Figure F.2 shows each school's performance in terms of average grades, according to official data from the Ministry of Education, Science, and Technology. The latest available update of this data is February, 2020. The dataset includes grades from the following subjects: Science, Health and Environment, Natural Sciences, Physical Education, Social Studies, Civics, Foreign Language, Computer Science, Educational Computing, English, Language and Literature, Mathematics, and Life Guidance.

Given that our study focuses on students from 7^{th} to 11^{th} grade, we extracted data from 3,699 schools offering these grade levels. For these grades, we computed the overall and the math average for each of the schools of our sample, only excluding school 5, where there was no available data. For comparison, we also computed the overall and math national average. The blue dotted line in Figure F.2 shows the overall national average and the blue bars show each school's average. For all subjects, the national average is 7.1 out of 10. In our sample, half of the schools have an overall average between 6.0 and 7.0, and the lowest and highest scores are 5.9 (school 12) and 7.9 (school 6). For Mathematics, the national score was 6.6 and 7 out of the 12 schools in our sample have an average between 6.0 and 7.0, with the lowest and highest scores of 5.7 (school 12) and 7.9 (school 6). Based on these comparisons, we conclude that, relative to the national average, the performance of the schools in our sample is neither systematically above nor below the national standard.

Figure F.2: School and national average



Source: Authors' calculations based on data from the information system for educational management of El Salvador (SIGES, in Spanish) by the Ministry of Education, Science, and Technology, 2019.

G. Inconsistent Choices and Educational Outcomes

Following the approach of Jacobson and Petrie (2009), we extend our analysis to examine how riskiness and the propensity to make mistakes influence educational expectations.[31] First, we employed an error choice model to estimate the error rate in our sample using maximum likelihood estimation. Our estimated error rate, $\epsilon = 0.227$, closely aligns with the 0.222 error rate for the gain lottery reported by Jacobson and Petrie (2009). Using this error rate, we then calculated the probability of making at least one mistake for each observed choice pattern in our sample.

We begin by analyzing whether our naïve measure of riskiness, *risky*,[32] and the tendency to make mistakes can explain decision-making. Table G.1 presents the results of OLS regressions with bootstrapped standard errors.[33] Models 1, 3, 5, and 7 show that neither riskiness nor the probability of making mistakes is significantly correlated with any of our immediate or long-term educational expectations measures.

[31] We replicated Jacobson and Petrie (2009)'s analysis as it closely aligns with the focus of our paper. We acknowledge that alternative approaches could have been used, such as those by Harrison et al. (2009) and Espinosa and Ezquerra (2022) to incorporate stochastic error using constant relative risk aversion (CRRA).

[32] $Risky$ counts the number of risky allocations (option B vs A) from 0 to 6. The distribution is reported in Table G.2.

[33] As in Jacobson and Petrie (2009), we avoided standard OLS since we are using the estimated probability of making at least one mistake as an independent variable, creating a generated regressor. The estimates were obtained by sampling with replacement 10,000 times to generate bootstrapped estimates and standard errors.

However, the interaction between riskiness and the propensity to make mistakes is statistically significant for three of our four measures. Conditional on a higher likelihood of making mistakes, risky individuals are less likely to believe that they will study next year, attend university, or even desire to attend university. Nevertheless, their school performance, as captured by their self-reported GPA, is not significantly correlated with riskiness, the tendency to make mistakes, or the interaction between the two.

Table G.1: *Educational expectations with inconsistency measure OLS regression with bootstrapped errors*

Vars.	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Self-reported GPA		*Likelihood of studying next year*		*Likelihood of going to university*		*Wants to go to university*	
Risky	0.012	0.021	-0.003	0.010	-0.002	0.008	0.009	0.018**
	(0.035)	(0.038)	(0.008)	(0.009)	(0.007)	(0.008)	(0.007)	(0.008)
Estimated probability of mistake	4.164	8.862	0.832	7.707***	0.635	5.941***	0.076	4.891**
	(3.468)	(12.754)	(0.592)	(2.356)	(0.623)	(2.151)	(0.567)	(1.997)
Risky x Estimated probability of mistake		-1.023		-1.508***		-1.155**		-1.056**
		(2.623)		(0.512)		(0.460)		(0.436)
Female	-0.112	-0.110	0.027*	0.028*	0.059***	0.061***	0.078***	0.078***
	(0.076)	(0.076)	(0.015)	(0.014)	(0.014)	(0.014)	(0.013)	(0.013)
Year	0.099***	0.098***	-0.009	-0.010	-0.004	-0.004	0.005	0.004
	(0.038)	(0.038)	(0.007)	(0.007)	(0.007)	(0.007)	(0.006)	(0.006)
Constant	1.607***	1.577***	0.970***	0.934***	0.522***	0.489***	0.728***	0.703***
	(0.455)	(0.457)	(0.095)	(0.096)	(0.082)	(0.083)	(0.080)	(0.082)
Observations	2,248	2,248	1,160	1,160	2,248	2,248	1,160	1,160
R-squared	0.028	0.028	0.027	0.035	0.077	0.080	0.046	0.051

Standard errors in parentheses (clustering at the subject level)
*** p<0.01, ** p<0.05, * p<0.1

Note: Standard errors reported in parentheses. The estimates are bootstrapped 10,000 times. All regressions include the following control variables: gender, year, class size, device and school-level fixed effects. There are fewer observations for the questions on likelihood of studying next year and wanting to go to university because they were included during data collection.

Table G.2: *Risky distribution*

Risky	Freq.	Percent
0	105	4.67%
1	129	5.74%
2	285	12.68%
3	522	23.22%
4	688	30.60%
5	372	16.55%
6	147	6.54%
Total	2,248	100%