Visual convex time preferences: Lab, field and high schools

Pablo Branas-Garza (branasgarza@gmail.com)
Universidad Loyola Andalucia https://orcid.org/0000-0001-8456-6009
Benjamin Prissé
Universidad Loyola Andalucia

Research Article

Keywords: Time Preferences, MPL, Experiments

Posted Date: May 18th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-519888/v1

Visual convex time preferences
Lab, field and high schools*

Pablo Brañas-Garza[†], Benjamin Prissé[‡]

May 11, 2021

Abstract

We introduce the Visual Convex Time Preferences (VCTP) task, a new tool for measuring time preferences that synthesizes the simplicity of Multiple Price List (MPL) and the precision of Convex Time Budget (CTB) tasks. We evaluate VCTP in three environments: lab, field and high school. The lab experiment suggests that VCTP improves the precision of time preferences without increasing task time or decreasing subject consistency. The field experiment partially replicates the results, since subjects make little use of the additional precision. The high school experiment shows that younger populations find it difficult to perform both the MPL and the VCTP, but older teenagers use the advantage of VCTP at no cost. Overall, the results indicate that the task successfully measures time preferences but should be adapted to specific populations.

Keywords: Time Preferences, MPL, Experiments.

JEL-codes: C91, C93, D15.

---

*Special thanks to Glenn Harrison for his comments and support and Charlie Sprenger for sharing code and time. We also thank Antonio Cabrales, Diego Jorrat, Antonio Espín, John Smith and the audience at ESA 2019, SABE 2020 and ESA 2020 for their helpful comments. This paper uses data from several studies: Almudena Amoros, Ernesto Mesa, Alvaro Nuñez and Juan F. Ferrero ran the lab experiments; The Honduras field data were gathered by Diego Jorrat, Lorenzo Estepa and Gladis Gonzales (PILARH) and the data from schools by Antonio Alfonso and María José Vázquez. Financial support from MINECO-FEDER (PGC2018-093506-B-I00) and Excelencia–Andalucia (PY18-FR-0007) is gratefully acknowledged.

[†]*LoyolaBehLab*, Universidad Loyola Andalucía, Spain.

[‡]*LoyolaBehLab*, Universidad Loyola Andalucía, Spain.

"It should be noted that the seeds of wisdom that are to bear fruit in the intellect are sown less by critical studies and learned monographs than by insights, broad impressions and flashes of intuition."
– Von Clausewitz, *On War* (1873)

1 INTRODUCTION

Humans commonly make decisions implicating inter-temporal outcomes. From daily choices such as consuming or saving to decisions with long-term consequences such as buying a house or rising a family through periodic investments in cultural, economic or social capital, time preferences of individuals are at the core of decision-making. Understanding the formation of time preferences and estimating them is therefore of great importance for economists and policy makers. The economic literature links patience to better decision-making with several papers. Ashraf, Karlan and Yin (2006) found that more patient women are more likely to open a bank account and increase their savings, Meier and Sprenger (2010) showed that impatience is correlated with more frequent credit card debt and higher level of borrowing, Burks, Carpenter, Goette and Rustichini (2009) linked higher cognitive skills with higher patience and better decision-making in economic problems, Meier and Sprenger (2013) showed that most patient individuals are also the ones likely to take financial literacy lessons. The literature also suggests that time preferences is a domain-general characteristic. Chabris, Laibson, Morris, Schuldt and Taubinsky (2008) found that time preferences predicts field behaviors better than individual characteristics, Reuben, Sapienza and Zingales (2010) showed that people impatient with primary rewards are similar with monetary rewards, Meier and Sprenger (2015) found that 60 to 70% of the variance in time preferences can be explained with individual fixed effect and that the aggregate distribution of time preferences is indistinguishable across the two parts of their experiment despite a two years gap between them. It suggests that time preferences are stable over time.

The recent state of the literature on time preferences is an evolution from the use of the Multiple Price List (MPL) proposed by Coller and Williams (1999) to the use of the Convex Time Budget (CTB) by Andreoni and Sprenger (2012). Historically, the literature dedicated to the measurement of individuals discount rates started with Thaler and Shefrin (1981) obtaining discount rates from 1% to more than 1000%. Comparable psychological experiment of Kirby and Marakovic (1996) or Kirby, Petry and Bickel (1999) similarly found unrealistic Annual Interest Rate (AIR) between 1000% and several billions, questioning how to plausibly estimate individuals discount rates. Coller et al. (1999) proposed a 15-choices MPL task as the first measurement tool for time preferences, thanks to a more controlled environment likely to elicit more realistic preferences[1]. Subjects choose between an amount of money in the present

[1]MPL were previously used in other economic measurements, such as risk attitudes with real payoffs (Binswanger, 1980, 1981) or valuation of a commodity like in Kahneman, Knetsch

and an amount of money in the future, with the standard version proposing a fixed amount of money in the present and an increasing amount of money in the future. Subjects should initially take the money in the present, then the choice at which they switch to future preferences gives an interval of potential values for their discount rate. They found estimates of discounted interest rates between 15% and 22%, lower than previous studies and consistent with market borrowing rates. This study established MPL as the gold standard for measuring time preferences at that time.

Harrison, Lau and Williams (2002) and Coller, Harrison and Rutström (2003) built on MPL and obtained estimates on a similar range, proving robustness of the tool. Harrison et al. (2002) estimated individual discount rates using a 20-choices MPL with a field survey population. They measured subjects preferences over four time intervals: 6 months, 12 months, 24 months and 36 months. They obtained AIR between 2.5% and 50% for individuals and estimated an overall discount rate of 28.1%, with variations according to sociodemographic characteristics of subjects. These estimates were in a plausible range, showing it was possible to elicit discount rates in a Field environment. Coller et al. (2003) investigated the effect of a Front End Delay (FED) in a 15-choices MPL task with interest rates from 2% to 100%. FED is the time delay between the date of the experiment and the date of payment. It can be of 0, 7 or 30 days, with a FED of 0 day meaning that subjects are paid between the moment they finish the experiment and the moment they leave the lab. Results demonstrated that subjects need to be paid a premium of approximately $10 in the 0 day FED condition to compensate them for waiting if they choose the future allocation, premium that does not exists in the 7 and 30 days conditions. It is consistent with constant discount rates with a FED of 7 or 30 days and with high discount rates for short time horizons before they decline for longer time horizons with a FED of 0 day.

Andersen, Harrison, Lau and Rutström (2006) pointed three limitations of MPL: First, subjects only indicate at which interest rate they switch for the larger amount of money in the future. It only estimates an interval for the interest rate, which is not as precise as a "point" valuation. Second, subjects could make multiple switches between present and future instead of switching to the future at one choice and sticking to it with larger future amounts. Third, framing effect can occurs if subjects are attracted by the middle choice. They proposed the two parts mechanism of *iterative MPL* (iMPL) to tackle these issues. The first part is named the *switching MPL* (sMPL) and enforces consistency by only allowing subjects to indicate at which interest rate they switch to future allocations. In the second part, subjects play MPL again for interest rates between their last present choice and their first future choice. Results suggest that the elicited discount rate is not influenced by whether MPL, sMPL or iMPL is the measurement task and that the framing effect can be controlled with minimal

and Thaler (1990)

adjustments. It also shows that the second step of iMPL increases precision by largely decreasing the interval size of estimated risk preference parameters, implying the same improvement for estimated time parameters. Unfortunately, results also show that the iMPL is not robust to minor procedural or subject pool differences. Ultimately, Andersen, Harrison, Lau and Rutström (2008) reset the investigation of new elicitation tools by showing that estimating the curvature of utility function improved the estimation of time discount parameter. They estimated risk and time preferences of subjects with Double Multiple Price List (DMPL) in which MPL identifies the discounting and Holt-Laury task (HL) the curvature, finding that using risk preferences to estimate the concavity of the utility function leads to lower and more plausible discount rates than with a linear utility function. This result is robust to alternative specifications of the utility function (hyperbolic, prospect theory) and a mixture model supports exponential over hyperbolic discounting. Andersen, Harrison, Lau and Rutström (2014) extended these results by studying several potential discounting functions. Results show no evidence of hyperbolic models (quasi, fixed cost, simple, general) or Weibull discounting but support exponential discounting since the population of subjects behave according to two types of this discounting.

Andreoni and Sprenger (2012) (AS thereafter) attracted a great deal of attention by creating the Convex Time Budget (CTB), an instrument identifying discounting and curvature of the utility function alone. CTB allows subjects to differentially allocate in the present or the future for each available unit of reward. Subjects are given 100 tokens and chose to allocate each of them in the future at $0.20 or in the present at a lower value between $0.13 and $0.19. Subjects are answering for three sooner payment dates and three payment delay lengths, making 45 budget choices corresponding to AIR between 20.5% and 1300%. Using a time separable Constant Relative Risk Aversion (CRRA) utility function discounted via the quasi-hyperbolic $\beta - \delta$ discounting function authors estimated a 30% AIR, rejected linear utility for slight concavity, found no evidence of present bias and substantial sensitivity of parameters to the value of background consumption. Authors also elicited time preferences with DMPL with results suggesting that CTB better estimates time discount by not being prone to the downward bias of price lists, and that using risk preferences to estimate curvature may be problematic since each method lead to uncorrelated estimates despite measuring the same construct. Andreoni and Sprenger (2012) and Andreoni, Kuhn and Sprenger (2015) further argued for CTB being a better elicitation method than DMPL. Andreoni et al (2012) proposed that measuring curvature with risk-preferences is inappropriate because time-preferences are different from risk-preferences. They associated the present and future payoffs of CTB with different levels of risk and documented substantial violations of predictions from standard model, since identical payoffs are treated differently according to risk level. Indeed, when both payoffs are uncertain subjects equalize payments between time periods more than when both are certain, and subjects show a preference for certainty when one of the payoff is certain and the other uncertain. Behavior is consistent with predictions of standard models only when all payments are uncertain. Andreoni et al (2015) compared CTB and DMPL by evaluating the predictive validity of both methods, using a simplified version of CTB where subjects answered for six interest rates with six possible allocation choices. Authors found that both methods perform well when predicting their own results (DMPL predicts 89% and CTB predicts 75% of their own results), but CTB outperforms DMPL when predicting results of the other method (DMPL predicts 16% and CTB predicts 86% of the other task results). Additionally, CTB outperforms DMPL when predicting results at individual level of three Becker-DeGroot-Marshak elicited willingness to accept.

Others researchers widely adopted CTB despite Andersen et al (2008) opinion that interacting time and risk preferences was a confound and that the specific identification of non-standard models of risk preferences could not be included in a broad task like CTB. Harrison, Hofmeyr, Ross and Swarthout (2018) and Andersen, Harrison, Lau and Rutström (2018) continued to separately estimate time and risk preferences with different methodologies. Harrison et al (2018) found that smokers discount more the future and that heavy smokers are more likely to show hyperbolic discounting. Andersen et al (2018) showed evidences of intertemporal risk aversion suggesting that the additive specification of the intertemporal utility function is not an appropriate representation of intertemporal preferences. Recently, Imai, Rutter and Camerer (2020) made a meta-analysis of CTB focused on present bias and found that studies with monetary reward had significant but small in magnitude present bias.

A significant limitation of CTB is that it requires a laboratory equipped with computers and a large amount of time to obtain the measurement[2]. A second important limitation is that consistency in CTB is fairly smaller than in MPL[3]. Therefore, a valuable development for this literature would be to propose a task integrating the small amount of time and high consistency of MPL with the precision of CTB. We created this new task combining the advantages of both methods and added the key methodological development of a simple visual representation of the task. It inspired us to name our task the *Visual Convex Time Preferences* (VCTP).

The aim of this paper is to validate this tool in the lab, the field and high schools. The rest of the paper is organized as follows. Section 2 explains the experimental design. Section 3 summarizes the main questions of this research. The results of the lab experiment, the field experiment and the high school experiment are respectively shown in sections 4, 5 and 6. Section 7 concludes.

[2]We ran the Andreoni-Sprenger experiment in the lab in London, finding that participants needed on average 18.5 minutes for the CTB and 2.3 minutes for the MPL.

[3]Using the datas of AS we found that only 30% of the participants made consistent choices, with most subjects only using corner solutions.

2 EXPERIMENTAL DESIGN

The key feature of our design is the use of a graphical representation with ten coins instead of a list. We use two experimental treatments that are identical with the exception of one key difference:

- MPL: subjects are forced to allocate all the coins either in the present or in the future.
- VCTP: subjects are allowed to allocate each coin in the present or in the future.

Since the only difference between MPL and VCTP is that *subjects can only choose present or future in the MPL while they can use interior solutions in the VCTP*, comparing both allow us to judge which mechanism is superior.

In what follows we explain the experimental design in detail.

Graphical representation: In AS subjects have 100 units and allocate each one to the present or the future for each decision. In our experiment we reduced the task to a 10-choices space for concision. Each choice space comprises a solid line circle representing one euro in the present, and a dotted circle representing the monetary bonus associated with waiting one week at the current interest rate. The cross inside the circle is used by subjects to indicate their choice by marking it with a two-colored pencil. They mark the cross blue if they want the payment in the present or red if they want the payment in the future. Figure 1 displays the decision space[4]. As can be seen, the circles are arranged in three horizontal rows: the middle row is composed of four circles, while the first and third rows are composed of three circles positioned at the middle of the spaces in the center line. This allows the circles to form the most homogeneous arrangement possible in order to avoid any potential effect related to their arrangement. This graphical representation explains why we think that VCTP is an improvement over the simplified CTB of Andreoni et al (2015), since it helps subjects to understand and compare the monetary amounts at stake as well as improving the overall readability of the task[5]. The set of interior solutions is also larger in VCTP.

Number of scenarios: Subjects are invited to perform the decision task six times with increasing interest rates: 0%, 20%, 40%, 60%, 80% and 100%. We chose these values as a compromise between obtaining precise data and not requiring too much time or effort from subjects. The order of the decisions remains the same for all participants.

[4]English translation of the instructions: "Interest rate=60%. Tomorrow: €1, In one week and one day: €1.60".

[5]Additionally, the detailed numerical monetary amounts of simplified CTB can be added to the design if necessary.

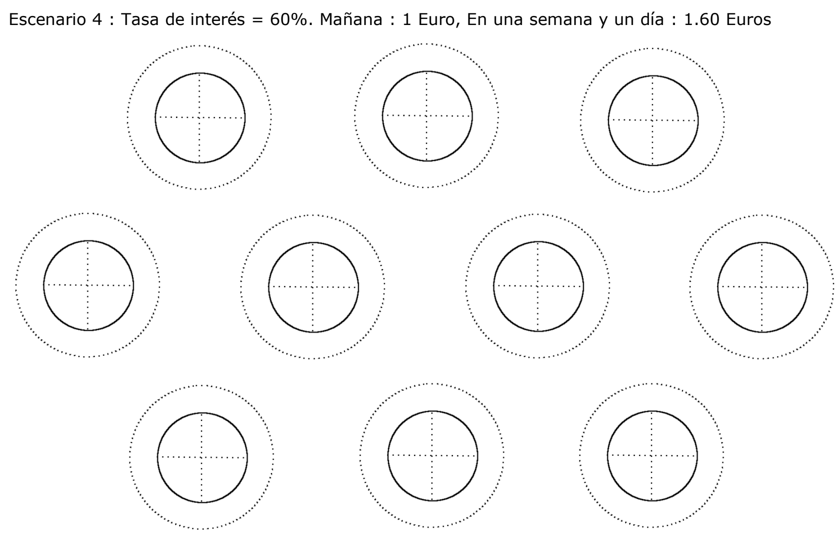


Figure 1: One decision task in the experiment.

Time Frame: The early payment date is tomorrow. This FED avoids the bias associated with obtaining the reward immediately. The later payment date is in one week and one day. We chose a one-week delay because it seemed a minimum amount of time to be perceived as a future payment for subjects, but not long enough to be perceived as the far future.

Measurements: We use three basic measurements to compare MPL and VCTP (and an additional one for the VCTP only):

(i) *Time* is the number of seconds subjects take to complete the task. In order to compute the total experimental time, we recorded the starting and finishing time for each task.

(ii) *Consistency* identifies subjects correctly performing the task. Subjects are consistent if their number of future choices always remains the same or increases when the interest rate increases. It would not be logical for a subject to allocate X€ in the present at one interest rate and then allocate more than X€ in the present at the next (and higher) interest rate. It should be noted that the consistency criteria is slightly different (and possibly harder) in VCTP than in MPL since X may take any value from 0 to 10 and not only 0 or 10.

(iii) *Allocations to the future* refers to the number of euros taken in the future at each interest rate during each task.

Subjects choose from within an extensive set of possible allocations when answering VCTP, with most of these allocations allowing to put coins both in the present and the future. These additional measures allow us to study the precision of subject's time preferences and we will refer to them as *interior solutions* from now on.

Ethical committee: The experiment was approved by the Ethical Committee of Universidad Loyola Andalucia on February 12th, 2019 (Teenagers) and April 28th, 2019 (Honduras and Spain).

3 QUESTIONS TO BE ADDRESSED

This research intends to analyze to what extent the VCTP mechanism outperforms the MPL mechanism in this design. Since MPL is the basic tool for measuring time preferences, it gives us the opportunity to show that our VCTP mechanism is better than a benchmark. To this end, we cover three different settings: the lab, the field and high schools. First, this allows us to compare three radically different subject populations: western university students, rural inhabitants of developing countries and western teenagers. Second, it gives us the opportunity to address several questions by using each setting appropriately.

The *lab* experiment aims to answer the following basic question: Do MPL and VCTP provide similar outcomes? The study was pre-registered in AsPredicted before being conducted and can be consulted at https://aspredicted.org/yd6pt.pdf.

Q1: Do MPL and VCTP provide similar outcomes in the lab?

The original idea of the CTB created by AS was to propose a more precise task than MPL. The VCTP task is an intermediate case between the binary MPL and the continuous CTB. The VCTP provides a more flexible version of the MPL. However, it is an open question whether subjects will take advantage of the flexibility of the VCTP and provide more precise answers than MPL or, conversely, if they will continue to make corner choices. Our paper will investigate if MPL and VCTP differ in three dimensions: (*i*) whether subjects take more or the same amount of time to use VCTP versus MPL; (*ii*) if subjects find the task similar in terms of complexity by counting the proportion of consistent subjects in both tasks; and (*iii*) whether there are differences in future allocations because subjects take advantage of the flexibility of VCTP by using interior choices.

To answer these questions, we asked our subjects to perform both the MPL and the VCTP experimental tasks (in random order). Summarizing, VCTP outperforms MPL if we prove that it does not require more time or diminish consistency, but that subjects use the additional precision of interior solutions.

We ran the $field$ experiment in order to answer two different questions besides replicating Q1. The first question refers to the number of circles to investigate if the interval length makes any difference. We then aimed to test whether using enumerators has an impact on outcomes. Finally, we wanted to test Q1 by comparing the performance of VCTP and MPL in the field. The study was pre-registered in AsPredicted before being conducted and the documentation can be consulted at https://aspredicted.org/dx52q.pdf.

Q2: Testing Q1 outside the lab (field)

We conducted the same analysis as Q1 and tested in the field whether subjects spend more time using VCTP. Recall that time in the field is expensive. We also tested whether the proportion of consistent subjects is similar. Inconsistent choices are costly since data are not considered valid and the sample size decreases. We feared that participants from rural areas of developing countries would have a higher level of inconsistency. Finally we checked whether subjects use the flexibility of VCTP.

In Appendix A.4 we extend the analysis in two dimensions. The first refers to the use of a shorter version of the experiment (five circles). Our original experimental design in the lab used ten balls, but time is money outside the lab. We need to test whether a shorter version, by definition less time-consuming, has a negative impact on results. That is, we need to verify that simplifying the experimental design is detrimental to the precision of results. To answer this question, we implemented two treatments: i) the original *10-balls* version and ii) the shorter *5-balls* version. Second, we study the use of enumerators. Experiments in the field are typically conducted by enumerators that help the responder. Lupu and Michelitch (2018) noted that the use of enumerators not only has a relevant impact on the casting, but that an enumerator effect could occur over a large range of technicalities that still has to be more extensively understood. Relevant to our study is that attributes of the enumerator influence replies to questions linked to the same attribute. It is also worth investigating if VCTP can be self-managed, meaning that it can be run independently as part of larger questionnaires. We investigated if enumerators have an effect by means of two conditions: *i*) *Self-managed* where subjects perform the task without any help from the enumerator and *ii*) *Externally-managed* where subjects perform the task with the help of the enumerator.

Finally, we repeated the experiments in *high schools*. We ran the same experiment in several high schools of Andalusia, Spain, to test whether teenagers are able to perform both MPL and VCTP with a minimum level of consistency. The study was pre-registered in AsPredicted before being conducted and the documentation can be consulted at https://aspredicted.org/pk7gd.pdf. Our question of interest is the following:

Q3: Testing Q1 outside the lab (schools)

We repeat questions for the lab by studying time, consistency and allocations to the future (and the use of the precision of the VCTP) across a large sample of teenagers ranging from 12 to 17 years old. Our study in high schools was only exploratory and there were no additional treatments.

4 THE LABORATORY EXPERIMENT

4.1 Procedures

The lab experiment was conducted with 151 subjects at the School of Economics and Business of the Universidad de Sevilla[6] in late May 2019. Regardless of the treatment, all subjects completed both VCTP and MPL. The order was randomized: 74 subjects played the VCTP first and then the MPL, 77 subjects played the reverse order. We also used three types of monetary incentives (paying all, hypothetical payment and paying one out of ten) but the distributions were identical.[7]

A sample of 151 subjects allows us to detect an effect size of 0.44 standard deviations (SD) with power 0.8 and a significance level of 0.1 for *time* (seconds). For *consistency*, the minimum detectable effect is equal to 0.12. We assume an SD of 100 seconds and a consistency rate of 0.85. In Appendix A.3 we check ex-post whether the estimates are sufficiently powered.

4.2 Is VCTP an improvement over MPL?

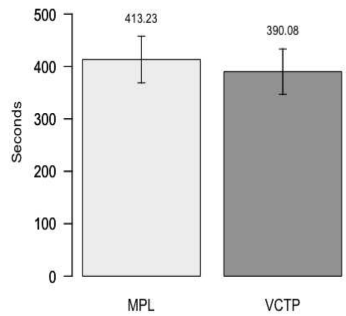
While VCTP is an improvement over MPL in terms of precision, it comes at the expense of increased complexity. VCTP could be considered an improvement over MPL if the time needed to collect data and the share of consistent subjects giving exploitable data remain similar. We also need to verify that there is a significant improvement in the precision of the results.

Is time higher in VCTP?

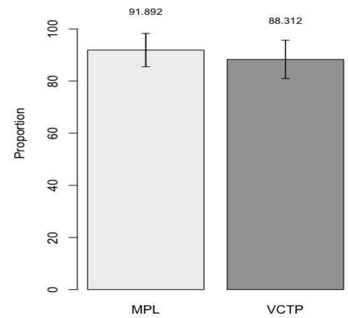
Figure 2a displays the average amount of time the subjects need to perform a task when played first (between-subjects). The t-test of equality of means between MPL and VCTP time with H0: the average time is equal between tasks (H1: MPL requires less time than VCTP) does not reject that MPL and VCTP need similar amounts of time to be performed in the lab (t = -0.305, $p = 0.760$). We therefore conclude that the amount of time needed to perform MPL and VCTP is similar.

[6] The Universidad de Sevilla is the largest public university in southern Spain.

[7] In Appendix 1 we show that monetary incentives do not matter. This result is consistent with a previous paper on time preferences in the lab, the field and online MPL experiments (see Brañas-Garza et al., 2020)



(a) Time



(b) Consistency

Figure 2: Time and consistency by task in the lab

Result 1a: There is no difference in time between MPL and VCTP in the lab.

Is consistency lower in VCTP?

Figure 2b displays consistency when the task is performed first. The t-test of equality of means between MPL and VCTP consistency does not reject H0: average consistency is equal between tasks (t = -0.731, p = 0.465). We therefore conclude that consistency in MPL and VCTP is similar. Additionally, we see that VCTP consistency is much higher than CTB consistency (see footnote 3).

Result 1b: There is no difference in consistency between MPL and VCTP in the lab.

Appendix A.2. revises results 1a and 1b using regression analysis (with and without controls). We find additional support for the above results.

Are VCTP choices more precise?

To show that VCTP is more precise, we need to prove that the additional precision of VCTP is used by subjects, and used *meaningfully*. Figure 3 displays multi-histograms of allocations to the future in MPL and VCTP by interest rate, allowing us to study if subjects use interior allocations.
According to interest rates by increasing order,[8] interior solutions are respectively used 11.92%, 60.93%, 58.28%, 52.32%, 48.34% and 26.49% of the times in VCTP. We see a considerable use of interior allocations in all scenarios except for the salient extremes of 0% and 100% interest rates where subjects are

[8] 0%, 20%, 40%, 60%, 80%, 100%.

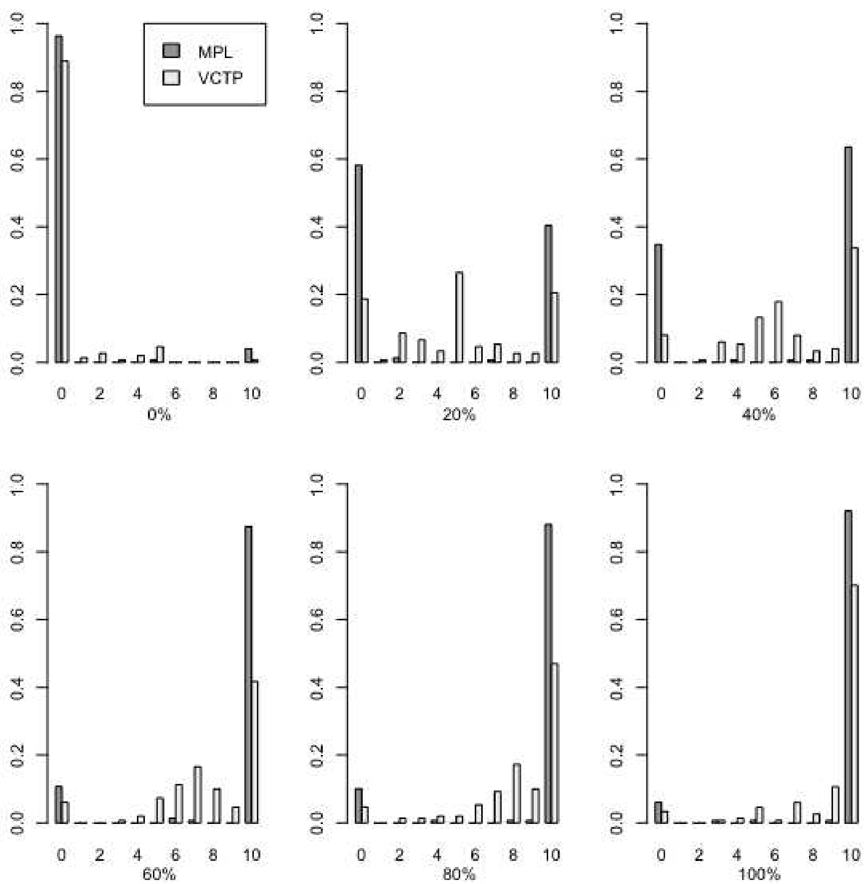


Figure 3: Multi-histograms of allocations to the future by task in the lab

expected to take everything in the present or future, yet they still use interior allocations at these interest rates. We conclude that the additional precision of VCTP is used by subjects.

To judge whether subjects use the additional precision of VCTP *meaningfully*, we took advantage of having MPL and VCTP choices for all subjects. We think that subjects switching from present allocations to future allocations in MPL at interest rates of 20%, 40% or 60% have precise time preferences they cannot express with the rigid structure of MPL. Figure 4 displays box plots showing these subjects' allocations to the future in VCTP compared with their allocations to the future in MPL (represented by the red line). The subjects should therefore take advantage of the VCTP's flexibility to allocate closer to their preferences. As we expected, subjects switching at 20% in MPL allocated more to the present in VCTP, since allocating 50 balls to the future is the 75^{th} quartile (and almost the median) allocation in VCTP. Subjects switching at 40% in MPL made slight adjustments to the present or the future in VCTP. This is coherent with the fact that this interest rate is close to the indifference point between the subjects' present and future. In VCTP, the median allocation to the future was 40 balls, with the 25^{th} quartile slightly below and the 75^{th} quartile slightly above. Finally, subjects switching at 60% in MPL allocated more to the future in VCTP (30 balls) as the 25^{th} quartile in VCTP. Therefore, we conclude that subjects use VCTP meaningfully.

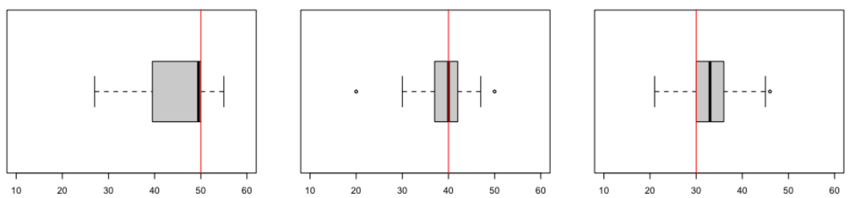


Figure 4: Allocations to the future in VCTP according to MPL switch

Additionally, we ran a bootstrap estimation to investigate to what extent VCTP correctly predicts choices in MPL. The results show that VCTP makes correct predictions for MPL 85.6% of the time, but with a large standard deviation of 35.1%. We also estimated the error size when MPL predicted VCTP with given α=0.9 (since MPL does not estimate the curvature parameter). We found an average error of 1.02 units with a large standard deviation of 2.34 coming from the inability of MPL to predict interior solutions, thus creating large errors when subjects use them. This suggests that VCTP is better at predicting choices of subjects because more information is available.

Result 1c: Subjects use the additional precision of VCTP to allocate closer to their preferences.

The last result is the main result of our paper because it allows us to conclude that the VCTP mechanism is superior to the MPL mechanism, since the precision of the time preference measurement increased without any associated cost in time or consistency of the task.

4.3 Parametric Estimation of VCTP data

To evaluate whether or not our instrument provides meaningful and useful estimations of delay discounting, we need to estimate them using a parametric model. The goal of this paper is not to propose a new estimation model, but to verify that the tool we propose provides meaningful estimates of time preferences by comparing the obtained estimations with those of AS. Since our experiment is a simplification of their extensive one, our goal is to obtain comparable estimates in terms of value with sizable diversity. We adapted to our environment their time separable CRRA utility function discounted via the quasi-hyperbolic β-δ discounting by removing the present bias β from the original utility function, since we defined tomorrow and not today as our early period[9].

$$U(c_t, c_{t+k}) = \frac{1}{\alpha}(c_t - \omega_1)^\alpha + \delta^k \frac{1}{\alpha}(c_{t+k} - \omega_2)^\alpha, \tag{1}$$

Where α is the CRRA curvature parameter, δ the one period discount parameter, t the present period, $t{+}k$ the future period, c_t earnings in the present, c_{t+k} earnings in the future and ω_1 and ω_2 the Stone-Geary consumption minima or background consumption. We set $k = 1$ since our experiment has only two periods. Being m the experimental budget, subjects have the following budget constraint:

$$(1 + r)c_t + c_{t+k} = m \tag{2}$$

Maximizing (1) subject to the future value budget (2) gives the tangency condition:

$$\frac{c_t - \omega_1}{c_{t+k} - \omega_2} = (\delta^k(1 + r))^{\frac{1}{\alpha-1}} \tag{3}$$

Giving us the Stone-Geary linear demand for c_t:

$$c_t = \frac{1}{1 + (1 + r)(\delta^k(1 + r))^{\frac{1}{\alpha-1}}} \cdot \omega_1 + \frac{(\delta^k(1 + r))^{\frac{1}{\alpha-1}}}{1 + (1 + r)(\delta^k(1 + r))^{\frac{1}{\alpha-1}}} \cdot (m - \omega_2) \tag{4}$$

Additionally, we estimate the constant absolute risk aversion (CARA) utility with $u(c_t) = -exp(-\rho c_t)$ to check for robustness to alternate forms of utility. The tangency condition is:

$$c_t - c_{t+k} = \frac{ln(\delta)}{-\rho} \cdot k + \frac{1}{-\rho} \cdot ln(1 + r). \tag{5}$$

And the solution function is:

$$c_t = \frac{ln(\delta)}{-\rho} \cdot \frac{k}{2 + r} + \frac{1}{-\rho} \cdot \frac{ln(1 + r)}{2 + r} + \frac{m}{2 + r} \tag{6}$$

Table 1 gives estimates using the different techniques and specifications (only pooled estimations for concision are presented). *Weekly time discount* $\hat{\delta}$ refers to the estimated weekly discount rate, *CRRA curvature parameter* $\hat{\alpha}$ refers to the estimated CRRA curvature and *CARA curvature parameter* $\hat{\rho}$ refers to the estimated CARA parameter.

As can be seen, the estimated values for $\hat{\alpha}$ and $\hat{\delta}$ in the regressions of Table 1 are slightly lower than the AS estimates. The estimated $\hat{\delta}$ corresponds to daily discount rates between 0.970 and 0.996, while the AS estimated daily discount

[9]The interested reader can find the original mathematical model in AS.

Table 1: Discounting and curvature parameter estimates

	NLS (1)	NLS (2)	Tobit (3)	Tobit (4)	Tobit (5)
Weekly time discount $\hat{\delta}$	0.814	0.807	0.929	0.954	0.972
	(0.010)	(0.010)	(0.008)	(0.004)	(0.002)
CRRA curvature parameter $\hat{\alpha}$	0.901	0.859	0.926		
	(0.008)	(0.007)	(0.009)		
CARA curvature parameter $\hat{\rho}$				0.046	0.029
				(0.005)	(0.002)
$\hat{\omega}_1 = \hat{\omega}_2$	0.938	0.01	-	-	
	(0.206)	(-)	(-)	(-)	(-)
R^2 / LL	0.746	0.744	-1773.8	-1896.2	-1691.8
# Observations	894	894	880	880	894
# Uncensored			368	368	382
# Cluster	149	149	149	149	149

NLS and two-limit Tobit ML estimators. Column (1): CRRA regression of equation (3) with restriction $\omega_1 = \omega_2$. Columns (2) and (3): CRRA regressions of equations (3) and (2), with restriction $\omega_1 = \omega_2 = 0$. Columns (4) and (5): CARA regressions of equations (4) and (5).

rates are between 0.996 and 0.997. A small difference in the daily discount rate makes a large difference at annual level, where discount rates below 0.990 translate into an almost complete discounting of the monetary amount, thus suggesting that our estimates are too low. We think these differences are due to the different experimental designs. The rationality check scenario with a 0% interest rate represents $\frac{1}{6}$ of our experiment and $\frac{1}{45}$ of the AS experiment. Consequently, his impact on decreasing $\hat{\delta}$ is much larger in our design. We also see that regressions (4) and (5) verify the robustness of the tool by providing more coherent $\hat{\delta}$ estimates. However, the regression with ω_1 and ω_2 provided implausible parameters when estimated separately. Hence, we choose to not include it in the paper and attributed this result to the erroneous methodology of asking to estimate several values of background consumption although other results suggest this has no influence. Figure A1 and A2 display histograms of the values of α and δ individuals according to regression (2), since we judged this estimation to be the most plausible one. The interested reader can find the details of individual estimations according to regression (2) in the Supplementary Material. Since we obtain coherent estimations in line with the AS estimations and develop the findings by suggesting that background consumption does not influence subjects' choices, we conclude that the parametric estimation validates the pertinence of our tool for estimating time preferences.

Result 1d: The estimated values for α and δ are lower than in AS but remain coherent. Subjects do not consider background consumption in their choices.

5 THE FIELD EXPERIMENT

5.1 Procedures

The field experiment was conducted at Santa Rosa de Copán, Copán Region, Honduras. After running an ex-ante power analysis with the results obtained in the lab experiment and Brañas-Garza et al. (2020)[10], we determined that a sample of 360 subjects would allow us to detect effects size of 0.36 SD with power 0.8 and a significance level of 0.1 for time (in seconds). The minimum detectable effect for consistency is 0.08. We assumed a SD of 100 seconds and a consistency rate of 0.85. Appendix A.3 revises the ex-post power of the estimates.

A consultancy firm (PILARH) was hired to run the experiment as part of a larger World Bank project. We chose eight primary schools and randomly selected households from their listing to recruit our experimental population. All participants signed an informed consent form and were paid for real, thus we had no treatment related to the payment scheme. The subjects received 25 lempiras as a show-up fee and earnings in the task ranged from 50 to 100 lempiras (L100 $\equiv$ \$4.1).

We were interested in testing whether our task could be used in the context of rural inhabitants of developing countries with poor and possibly illiterate farmers who would likely find it difficult to understand the mechanism. We considered this experiment as a robustness test for our mechanism that also provides the context for answering two additional questions: the use of enumerators and the number of balls.

As in the lab experiment, all subjects completed both tasks (VCTP and MPL) regardless of the treatment. The order was randomized, with 157 subjects first playing MPL then VCTP and 172 subjects doing the experiment in reverse order. The subjects were randomly allocated to 2x2 treatments. From the randomization we got $n_{5,S} = 69$, $n_{10,S} = 77$, $n_{5,E} = 91$ and $n_{10,E} = 92$ with S (E) referring to self-managed (externally-managed) and 5(10) to the number of balls.[11] For brevity we will focus on the main question of the paper, the analysis of treatment effects is relegated to Appendix A.4.

We slightly modified our experimental design since we expected subjects to have difficulties understanding the representation of interest rates with circles (see Figure 1). Instead, we used the universally known symbol of piggy banks to represent interest rate level. Figure 5 shows an example of a piggy bank for

[10] There is no evidence in the literature of the effect of different designs and administration methods.

[11] The precise numbers by order, number of balls and management type are: MPL first: $n_{5,S} = 35$, $n_{10,S} = 33$, $n_{5,E} = 47$ and $n_{10,E} = 42$ and VCTP first: $n_{5,S} = 33$, $n_{10,S} = 44$, $n_{5,E} = 44$ and $n_{10,E} = 50$.

Figure 5: Piggy bank example with a 60% interest rate

a 60% interest rate and Figure A3 displays the complete decision task.

Due to the population of subjects and the difficulty of the task, we expected that some subjects would not be able to perform the task without the help of an enumerator. A total of 18 subjects fully reported NA in the task, 15 of which were self-managed and 3 of which were externally-managed, respectively representing 10.3% of self-managed subjects and 1.6% of externally-managed subjects. Of these subjects, 13 stated they were unable to read and 5 did not want to perform the task, with all subjects lost due to illiteracy being in the self-managed condition. This suggests that enumerators prevented the loss of 7.5% of the sample due to illiteracy and clearly shows that enumerators increase sample size.

5.2 Replication of Laboratory Results

Table 2 studies the differences between MPL and VCTP for the Honduras field data in terms of time (seconds) and consistency. To control for treatment effects, we ran a regression including dummies for VCTP, self-managed, 5-balls and interactions terms for these variables. The reference group is the externally-managed 10-balls MPL.

We found that the average amount of time needed to complete the VCTP task[12] is similar to MPL. Moreover, none of the treatments had a significant impact. Figure A4 in the Appendix shows that this result holds regardless of the number of balls. We conclude that MPL and VCTP require similar amounts of time when performed in the field.

Result 2a: Time is similar between tasks in the field.

The right side of Table 2 focuses on consistency. Compared to the control group (externally-managed 10-balls MPL), the VCTP itself has no effect. However, this is true for sessions with enumerators only, since we observe that self-managed experiments with VCPT (the interaction term VCTP×Self-managed) report a significantly lower level of consistency ($p < 0.01$) than MPL. Figure 8b decomposes consistency in each task by type of management and Figure A5 by number of balls.

[12]When played first, like all results in this section.

Table 2: Regression analysis for the field experiment

	Time (1)	Time (2)	Consistency (3)	Consistency (4)
VCTP	-32.835	-32.385	-0.010	-0.009
	(47.860)	(47.977)	(0.084)	(0.084)
Self-managed	18.585	16.432	0.130*	0.127*
	(58.714)	(58.772)	(0.074)	(0.075)
5-balls	-75.215	-76.697	0.105	0.113
	(49.484)	(50.081)	(0.074)	(0.074)
VCTP × Self-managed	77.612	78.289	-0.316***	-0.312***
	(74.039)	(73.943)	(0.120)	(0.120)
VCTP × 5-balls	30.301	30.197	-0.133	-0.141
	(53.678)	(53.758)	(0.113)	(0.113)
Self-managed × 5-balls	-62.914	-61.687	-0.159	-0.157
	(104.158)	(103.285)	(0.101)	(0.101)
VCTP × Self-man. × 5-b	23.836	29.301	0.220	0.209
	(118.808)	(115.853)	(0.173)	(0.174)
Constant	269.415***	195.090***	0.810***	0.719***
	(45.628)	(74.629)	(0.061)	(0.158)
Observations	302	302	329	329
R-squared	0.051	0.054	0.073	0.080
Controls	No	Yes	No	Yes

Robust standard errors in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Result 2b: MPL and VCTP have similar consistency with field enumerators. MPL is superior without them.

Hence the use of enumerators does not only help to avoid losing 7.5% of the sample because of illiteracy, it also maintains higher levels of consistency.

We were interested in testing whether interior solutions are used in the field. Figure 6 displays the multi-histograms of allocations by task and interest rate for consistent subjects in the 10-balls task[13], regardless of the use of enumerators. The figure shows that both tasks follow a similar trend and that interior solutions are hardly used. Figure A6 displays the cumulative multi-histograms of allocation type by tasks, management types and number of balls for all subjects, showing that field subjects use interior solutions but are not consistent if they do so.

[13]Interior solutions are rarely used by consistent subjects in 5-balls VTCP.

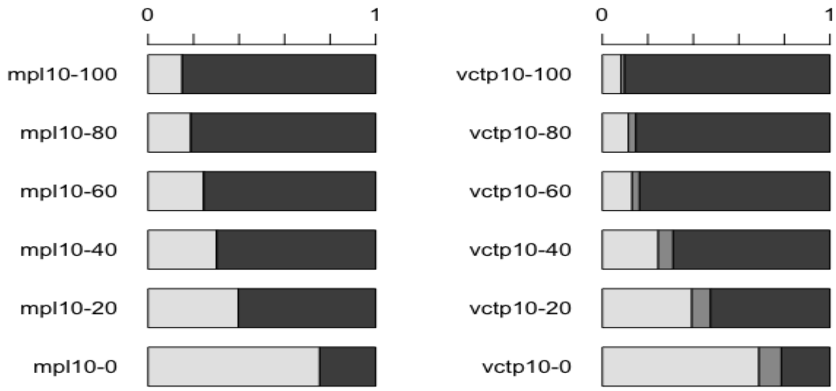


Figure 6: Frequency of allocation types by interest rate and task

Result 2c: Consistent subjects make little use of interior solutions.

In sum, the (10-balls) VCTP is similar to the MPL in the field when the experiment is conducted with enumerators. There are no costs in time or consistency level but the VCTP precision is barely used.

6 THE HIGH SCHOOL EXPERIMENT

6.1 Procedures

The high school experiment was conducted online using LimeSurvey with high school students from southern Spain completing standard courses. The experiment was incentivized by randomly choosing one student over twenty to be paid[14]. The maximum payoff was 30 euros and the money was paid by bank transfer. The average earnings for those selected was €10.95.

Based on previous results, we targeted 400 participants (200 young teenagers and 200 older teenagers randomly assigned to MPL or VCTP). This sample size allowed us to detect an effect size of 0.4 SD with power 0.8 and a significance level of 0.1 for time (in seconds). For consistency the minimum detectable effect is 0.1. We assumed a SD of 100 seconds and a consistency rate of 0.85. Appendix A.3 reviews the ex-post power of the estimates.

As in Honduras, we used piggy banks since we did not expect teenagers to understand the concept of interest rate. Participants were randomly assigned

[14]Subjects gave a number between 1 and 20. If that number matched the number between 1 and 20 randomly determined by the computer, they were paid.

to either MPL or VCTP but not both. Selecting subjects between twelve and seventeen years old who completed the task, we were left with a sample of 446 subjects (n_{MPL}=212 or n_{VCTP}=234).

For briefness, we only present the main results regarding time, consistency and the use of interior solutions in each task. We separated the sample according to the median age (14.97). If subjects were younger or equal to the median age they were considered Young ($n = 222$, mean age=14.07), otherwise they were considered Old ($n = 224$, mean age=15.60). It is worth emphasizing that these participants completed the experiment online and this may partially explain some differences with experiments conducted in the lab and in Honduras.

6.2 Replication of Laboratory Results

Table 3 studies the differences between MPL and VCTP for time (in seconds) and consistency in high schools. We use dummies for VCTP and Old, the reference group is the sample of young teenagers performing MPL.

We found no differences between both mechanisms regarding time (in seconds). Interestingly, teenagers who completed the task online needed a similar amount of time as the university students completing the task in the lab and both groups completed the task in more than 400 seconds. We therefore conclude:

Result 3a: Time is similar between tasks among teenagers.

The right side of Table 3 studies differences in consistency. Recall that the reference group is the sample of young teenagers using MPL. We observe that VCTP has a negative and significant impact on consistency ($p < 0.01$) and that the interaction of VCTP×Old is positive but not significant. In order to test whether there are no differences between MPL and VCTP for older teenagers, we ran one LM test (Table 3 bottom) confirming that there are no differences.[15]

Result 3b: MPL and VCTP are similar in consistency for older teenagers, MPL is more consistent with younger teenagers.

Figure 7 displays the allocation type in each task and age group for consistent subjects only ($n = 239$). As can be seen, young teenagers allocate more to the future than older teenagers. We also see that both groups allocate more to the future in VCTP than in MPL. More importantly, older subjects use interior solutions more frequently than the younger ones.

We interpret these results as suggesting that younger teenagers tend to choose the largest income, while older teenagers possibly have a more sophisticated

[15]We ran a separate regression for older teenagers only ($n = 222$) and found that the estimated coefficient for $VCTP = -0.07$ ($p = 0.286$).

Table 3: Regression analysis for high school

	Time (1)	Time (2)	Consistency (3)	Consistency (4)
VCTP	78.861	75.159	-0.197***	-0.192***
	(48.416)	(48.093)	(0.065)	(0.065)
Old	-12.086	-17.915	-0.137**	-0.130*
	(48.009)	(48.523)	(0.067)	(0.067)
VCTP × Old	-72.679	-61.359	0.125	0.111
	(66.957)	(66.932)	(0.094)	(0.093)
Constant	439.051***	391.039***	0.677***	0.735***
	(36.875)	(37.874)	(0.047)	(0.053)
LM Test: VCTP × Old			-0.072	-0.081
			(0.067)	(0.066)
Observations	446	446	446	446
R-squared	0.012	0.030	0.026	0.039
Controls	No	Yes	No	Yes

Robust standard errors in parentheses. LM stands for Lagrange-Multiplier test. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

conception of their preferences and discriminate more between the present and the future.

Result 3c: Older teenagers use the precision of VCTP.

We conclude that teenagers face difficulties when answering MPL and even more difficulties when answering VCTP in terms of both time and consistency in the task. Older teenagers perform slightly better than younger ones and seem to start understanding the complexity of the task as they use interior solutions in a similar way to the lab subjects. This suggests that VCTP can be used with older teenagers, but needs to be adapted for younger populations.[16].

[16]In a parallel paper we study inconsistencies in time and risk preferences among teenagers (for more details, see Alfonso et al., 2020).

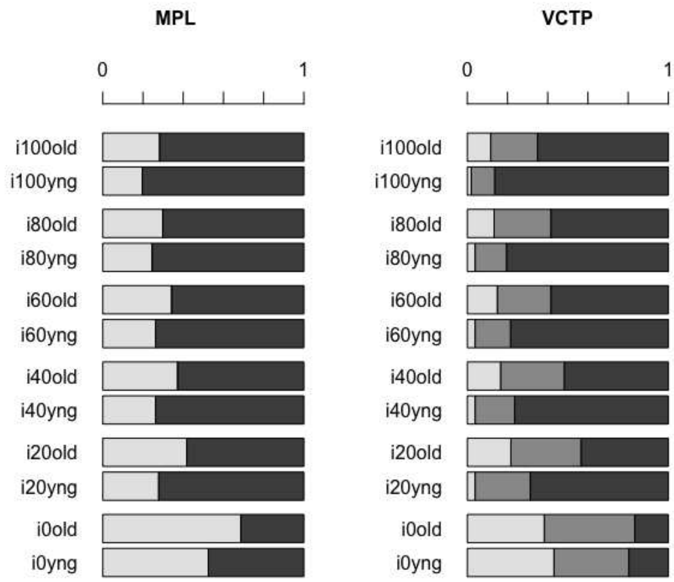


Figure 7: Frequency of allocation type by interest rate, task and age group for teenagers

7 DISCUSSION

Our paper aimed to develop a new tool for measuring time preferences. We developed a simplified version of the CTB with a visual aspect since we wanted to create a time measurement task that overcomes differences in educational level. We kept from CTB the idea of allocating to both the present and the future, and we kept from MPL the idea of minimizing the complexity and number of trials to be performed.

We named our task the Visual Convex Time Preferences. We first evaluated VCTP in the lab and found that the time and consistency of subjects are similar in VCTP and MPL, but that subjects use the additional precision of VCTP. We also estimated plausible weekly discount rates and curvature parameters using Andreoni-Sprenger's methodology. Additionally, we obtained a wider panel of daily interest rates because the experiment makes each choice meaningful and captures a greater diversity of estimated interest rates.

The task was then performed in Honduras to validate our design in the field with poor rural farmers, as well as testing whether the experimental design could be simplified to five balls and whether subjects could perform the task without enumerators. We partially replicated lab results. Specifically, we found that subjects do not spend more time using the VCTP than the MPL and consistency is maintained. However, we also showed that enumerators improve the quality of results by helping subjects to avoid inconsistencies and correctly use interior solutions, as well as increasing the sample size since letting subjects manage themselves leads to a 10% loss of the sample due to illiteracy. The main limitation is that this population does not frequently use the precision of the VCTP.

Finally, we carried out our experiment in high schools to study if teenagers are able to perform a task designed for adults. The results suggest that high school students find it difficult to complete both MPL and VCTP in terms of time and consistency, while older students perform better and more similarly to adults. The pattern of allocations to the future of consistent subjects suggests that older teenagers use the precision of VCTP appropriately.

We conclude that VCTP has potential for measuring time preferences both in the lab and in the field, but the results from young teenagers are problematic. A potential development of our task is to launch a visual version on electronic tablets. This development is interesting because an electronic application would be more accessible, faster to perform and could be adapted to younger teenagers and children by designing it like a game.

REFERENCES

Alfonso, A., Brañas-Garza, P., Prissé, B. & Vazquez, M. J. (2020). Testing consistency among teenagers: Experiments from high schools, *mimeo.*

Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2006). Elicitation using multiple price list formats. *Experimental Economics, 9*(4), 383-405.

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

Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2014). Discounting behavior: A reconsideration. *European Economic Review, 71*, 15-33.

Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2018). Multiattribute utility theory, intertemporal utility, and correlation aversion. *International Economic Review, 59*(2), 537-555.

Andreoni, J., Kuhn, M. A., & Sprenger, C. (2015). Measuring time preferences: A comparison of experimental methods. *Journal of Economic Behavior & Organization, 116*, 451-464.

Andreoni, J., & Sprenger, C. (2012). Estimating time preferences from convex budgets. *The American Economic Review, 102*(7), 3333-56.

Andreoni, J., & Sprenger, C. (2012). Risk preferences are not time preferences. *The American Economic Review, 102*(7), 3357-76.

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

Brañas-Garza, P., Jorrat, D., Espín, A.M. & Sánchez, A. (2020). Paid and hypothetical time preferences are the same: Lab, field and online evidence. *mimeo*

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

Binswanger, H. P. (1981). Attitudes toward risk: Theoretical implications of an experiment in rural India. *The Economic Journal, 91*(364), 867-890.

Burks, S. V., Carpenter, J. P., Goette, L., & Rustichini, A. (2009). Cognitive skills affect economic preferences, strategic behavior, and job attachment. *Proceedings of the National Academy of Sciences, 106*(19), 7745-7750.

Chabris, C. F., Laibson, D., Morris, C. L., Schuldt, J. P., & Taubinsky, D.

(2008). Individual laboratory-measured discount rates predict field behavior. *Journal of Risk and Uncertainty, 37*(2-3), 237.

Coller, M., Harrison, G. W., & Rutström, E. E. (2003). *Are discount rates constant? Reconciling theory and observation*. Working Paper 3.

Coller, M., & Williams, M. B. (1999). Eliciting individual discount rates. *Experimental Economics, 2*(2), 107-127.

Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1990). Experimental tests of the endowment effect and the Coase theorem. *Journal of Political Economy, 98*(6), 1325-1348.

Kirby, K. N., & Maraković, N. N. (1996). Delay-discounting probabilistic rewards: Rates decrease as amounts increase. *Psychonomic Bulletin & Review, 3*(1), 100-104.

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

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

Harrison, G. W., Lau, M. I., & Williams, M. B. (2002). Estimating individual discount rates in Denmark: A field experiment. *The American Economic Review, 92*(5), 1606-1617.

Imai, T., Rutter, T. A., & Camerer, C. F. (2021). Meta-Analysis of Present-Bias Estimation Using Convex Time Budgets. *The Economic Journal, 131*(636), 1788-1814.

Lupu, N., & Michelitch, K. (2018). Advances in survey methods for the developing world. *Annual Review of Political Science, 21* 195-214.

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

Meier, S., & Sprenger, C. (2015). Temporal stability of time preferences. *Review of Economics and Statistics, 97*(2), 273-286.

Meier, S., & Sprenger, C. D. (2013). Discounting financial literacy: Time preferences and participation in financial education programs. *Journal of Economic Behavior & Organization, 95*, 159-174.

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

Thaler, R. (1981). Some empirical evidence on dynamic inconsistency. *Economics letters, 8*(3), 201-207.

Von Clausewitz, C. (1873). *On war (Vol.1).* London, N. Trübner & Company.

A APPENDIX

A.1 Monetary vs. hypothetical incentives

The 151 participants were randomly allocated to treatments with $n_R = 50$ in the *Real* condition, $n_H = 52$ in the *Hypothetical* and $n_{1/10} = 49$ in the *One Tenth* treatment.

Regardless of the treatment, all subjects completed both VCTP and MPL. The order was randomized: 74 played the VCTP first and then the MPL, and 77 played in the reverse order. The resulting numbers from the randomization are:

- VCTP first: real ($n = 24$), hypothetical ($n = 25$) and 1/10 ($n = 25$).
- MPL first: real ($n = 26$), hypothetical ($n = 27$) and 1/10 ($n = 24$).

Table A1: VCTP allocations to the future in the first round

	All Subjects	All Subjects	VCTP Consistent	VCTP Consistent	Consistent	Consistent
Hypothetical	0.582	1.289*	0.861	1.473	0.905	1.512
	(1.07)	(2.16)	(1.31)	(2.00)	(1.35)	(2.01)
Onetenth	0.796	0.946	0.799	0.798	0.842	0.828
	(1.46)	(1.68)	(1.29)	(1.24)	(1.33)	(1.26)
Age		0.0819		0.167		0.171
		(0.81)		(1.18)		(1.19)
Female		1.481**		1.153		1.133
		(2.97)		(1.92)		(1.86)
CRT		-0.295		-0.381		-0.400
		(-0.46)		(-0.45)		(-0.47)
Constant	5.771***	3.137	5.926***	1.715	5.882***	1.610
	(14.82)	(1.39)	(12.76)	(0.56)	(12.22)	(0.51)
N	74	67	59	54	58	53
R^2	0.0310	0.1789	0.0380	0.1486	0.0404	0.1501

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

We wanted to know whether the payment scheme had an effect on allocations of subjects. Table A1 shows the estimates of the impact of payment schemes on individual choices in VCTP played first and Table A2 displays the same regressions for MPL played first. We ran each regression with and without controls (Age, Female, Cognitive Reflection Test (CRT)). We considered 6*2 potential cases in each table and found that only one is weakly significant (at 10%).

We concluded that *the payment scheme had no impact* and therefore merged the data. Hence, we had 74 observations for VCTP and 77 for MPL in the between-subjects analysis.

Table A2: MPL allocations to the future in the first round

	All Subjects	All Subjects	MPL Consistent	MPL Consistent	Consistent	Consistent
Hypothetical	-0.774	-0.366	-1.325	-0.832	-1.133	-0.622
	(-1.22)	(-0.50)	(-1.97)	(-1.06)	(-1.64)	(-0.78)
Onetenth	-0.182	-0.0603	-0.674	-0.462	-0.280	-0.0614
	(-0.28)	(-0.09)	(-0.98)	(-0.67)	(-0.39)	(-0.09)
Age		0.0656		0.0355		0.0398
		(0.97)		(0.51)		(0.58)
Female		1.309*		1.230*		1.094
		(2.39)		(2.13)		(1.84)
Crt		-0.632		-0.890		-1.001
		(-0.73)		(-1.00)		(-1.09)
Constant	6.731***	4.732**	7.222***	5.910**	7.105***	5.745**
	(14.87)	(2.88)	(14.43)	(3.36)	(13.76)	(3.21)
N	77	66	71	60	64	54
R^2	0.0215	0.1329	0.0541	0.1517	0.0473	0.1505

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

A.2 Regression analysis for the lab experiment

Regressions (2) and (4) include controls: Age, Female, CRT.

Table A3: Regression analysis for the lab

	Time (1)	Time (2)	Consistency (3)	Consistency (4)
VCTP	-24.159	-11.825	-0.036	-0.054
	(30.912)	(34.620)	(0.049)	(0.053)
Constant	413.229***	535.468***	0.919***	0.638***
	(22.308)	(77.822)	(0.032)	(0.134)
Observations	142	124	151	133
R-squared	0.004	0.011	0.004	0.031
Controls	No	Yes	No	Yes

Robust standard errors in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

A.3 Power calculations

We ran an ex-post power analysis to test whether the lack of significant results are due to lack of power. Table A4 explores the relevant cases. Below the table we comment on the problem of power associated to each case, if any.

Table A4: Ex-post power analysis

	#Subjects per arm	SD MPL	Effect (MDE)	Estimate	Power	Comments
Panel A: lab						
Time	75	173.74	69.50	-11.81	No, but irrelevant	C1
Consistency	75	0.28	0.09	-0.036	Yes	C2
Panel B: field						
Time	180	86.15	24.98	8.00	No, but irrelevant	C3
Consistency	180	0.36	0.08	0.17	Yes	C4
Panel C: high school - old						
Time	110	326.77	111.10	8.00	No, but irrelevant	C5
Consistency	110	0.50	0.16	-0.08	No, but irrelevant	C6
Panel D: high school - young						
Time	110	367.11	124.82	78.00	Problematic	C7
Consistency	110	0.47	0.15	-0.21	Yes	C8

C1: There is no power. Given our sample size, we find larger effects than 0.4 SD (which is equal to a difference of 70 seconds or more). However, this is not relevant since the difference is very small (7 seconds).

C2: There is sufficient power.

C3: There is no power, but the effect is irrelevant (0.09 SD) according to Cohen's effect size (<0.2 SD).

C4: There is sufficient power.

C5: There is no power, but the effect is irrelevant (0.02 SD) according to Cohen's effect size (<0.2 SD).

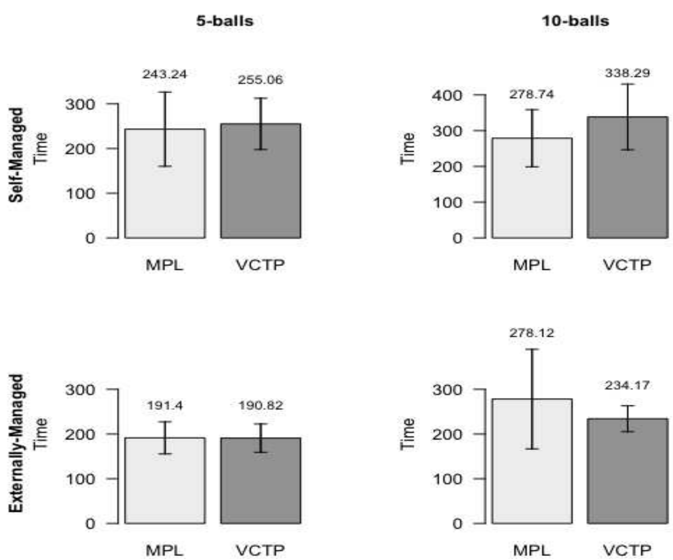
C6: There is no power, but the effect is irrelevant (0.15 SD) according to Cohen's effect size (<0.2 SD).

C7: There is no power and the effect is moderately relevant (0.21 SD) according to Cohen's effect size (<0.2 SD).

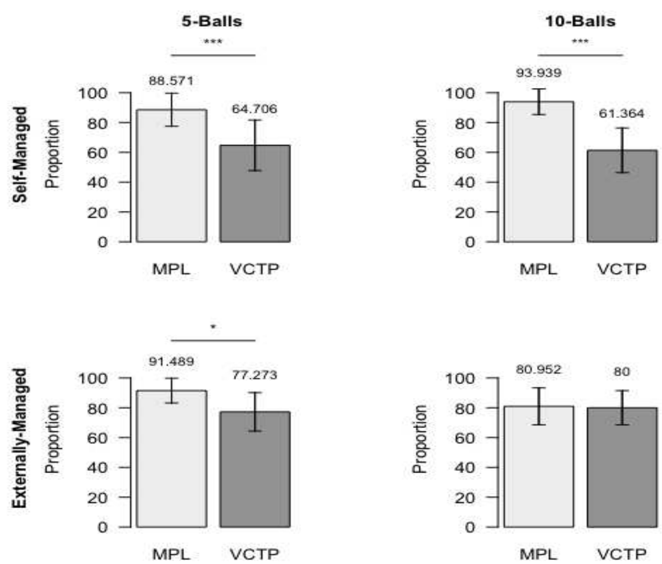
C8: There is sufficient power.

A.4 Effects of number of balls and management type

Figure 8 displays the time (8a) and consistency (8b) for each task according to number of balls and management type.



(a) Time



(b) Consistency

Figure 8: Time and consistency by number of balls and type of management

We see that consistency in VCTP is lower than in MPL for self-managed. However, with the help of enumerators these differences almost disappear for the externally-managed 10-balls task where consistency is high and similar between tasks, thus suggesting that enumerators increase consistency. To identify any effect of management on consistency, Figure A6 displays consistency for management type according to task and number of balls. As the figure shows, consistency is higher with an enumerator in all conditions. We conclude that enumerators seem to increase consistency in the field, especially with the complex 10-balls VCTP task.

We now investigate potential effects of number of balls or management type on precision. Figure 9 displays the cumulative multi-histograms of allocation type in each task and interest rate according to number of balls and management type for consistent subjects. As can be seen, the subjects barely use interior allocations at low interest rates in the externally-managed 10-balls task. This suggests that subjects are able to correctly use the additional precision of the VCTP task with the help of an enumerator [17]. Figure A7 shows the same cumulative multi-histograms as Figure 9 for all subjects. As shown, self-managed subjects use interior solutions more, while externally-managed subjects use interior solutions less but more meaningfully. The fact that most of these subjects are inconsistent suggests that they do not understand how to use interior solutions, but enumerators decrease this lack of understanding.

The field results suggest that enumerators slightly improve the size and quality of the outcomes. They also suggest that the 10-balls version is a better experimental design because it provides slightly more precision but not more errors than the 5-balls version. Overall, the field results suggest that the externally-managed 10-balls VCTP task is able to measure time preferences in a rural context if subjects receive the active help of an enumerator. In contrast to the lab experiment, *the field results support the VCTP less.*

[17] Percentage of interior solutions by interest rate: 12.8%(0%), 10.2%(20%), 7.7%(40%), 2.6%(60%), 5.1%(80%) and 2.6%(100%).

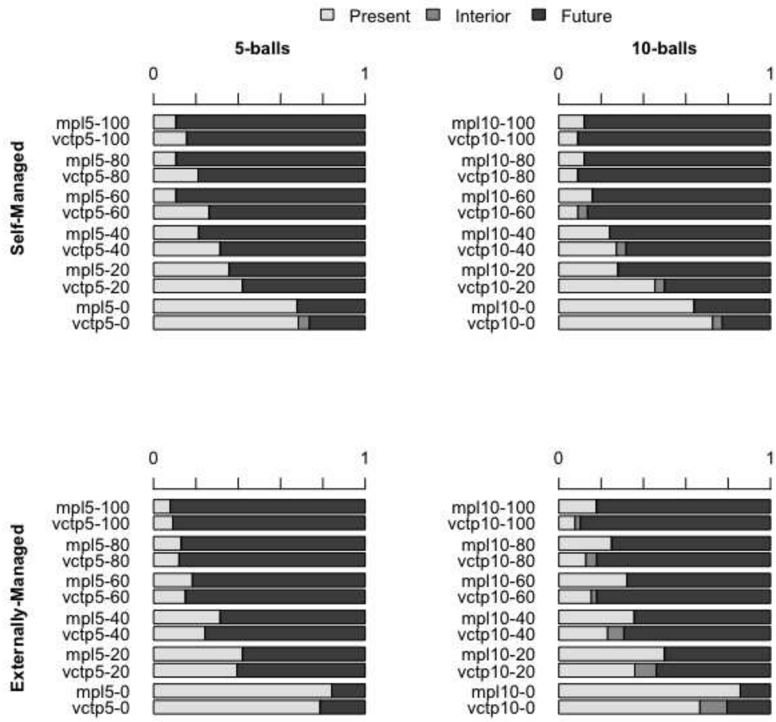


Figure 9: Frequency of allocation type by interest rate, task, management type and number of balls in the field

A.5 Additional figures

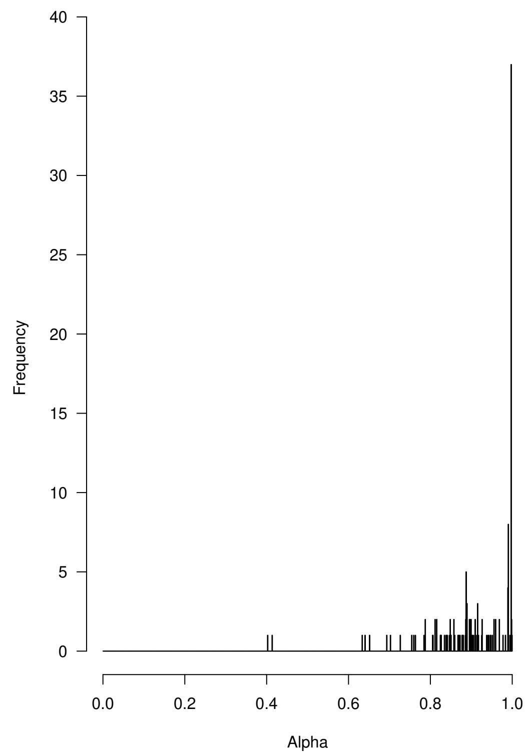


Figure A1: Histogram of estimated values for alpha of individuals

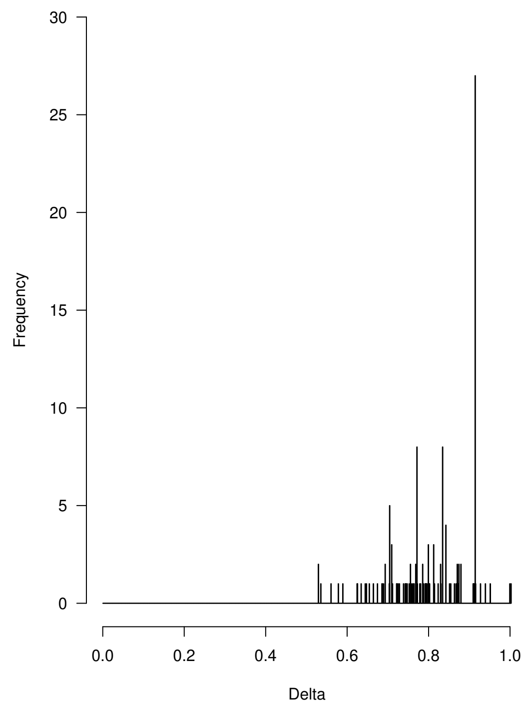


Figure A2: Histogram of estimated values for delta of individuals

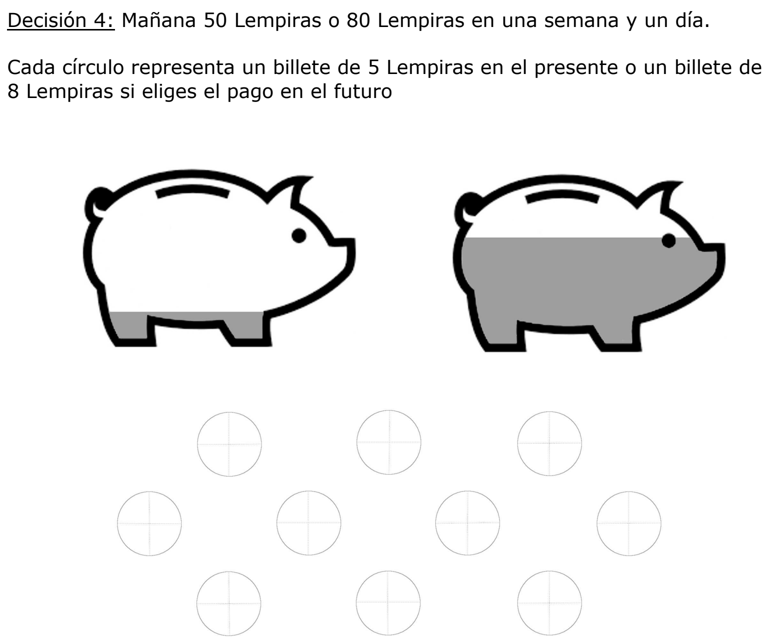


Figure A3: Example of a decision task in Honduras

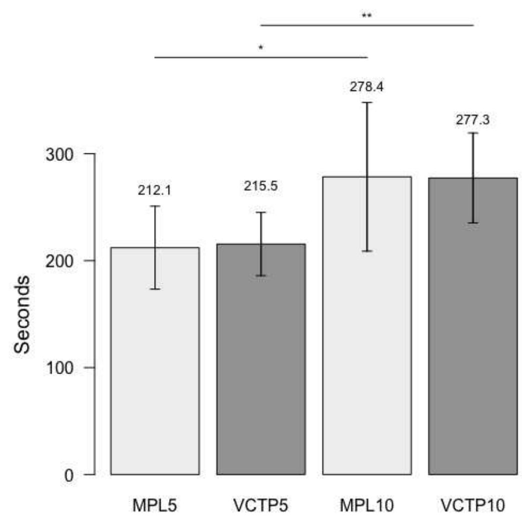


Figure A4: Time for task in the field according to number of balls

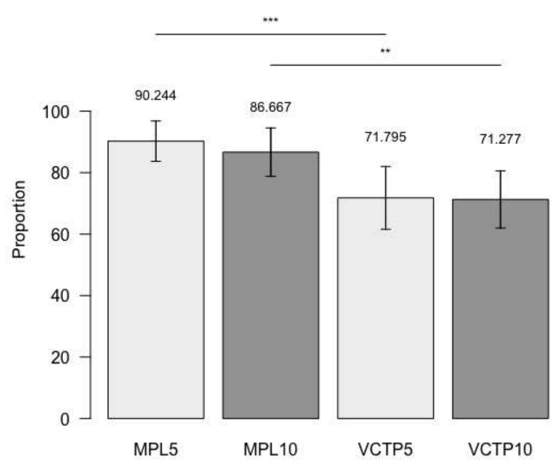


Figure A5: Consistency for task in the field according to number of balls

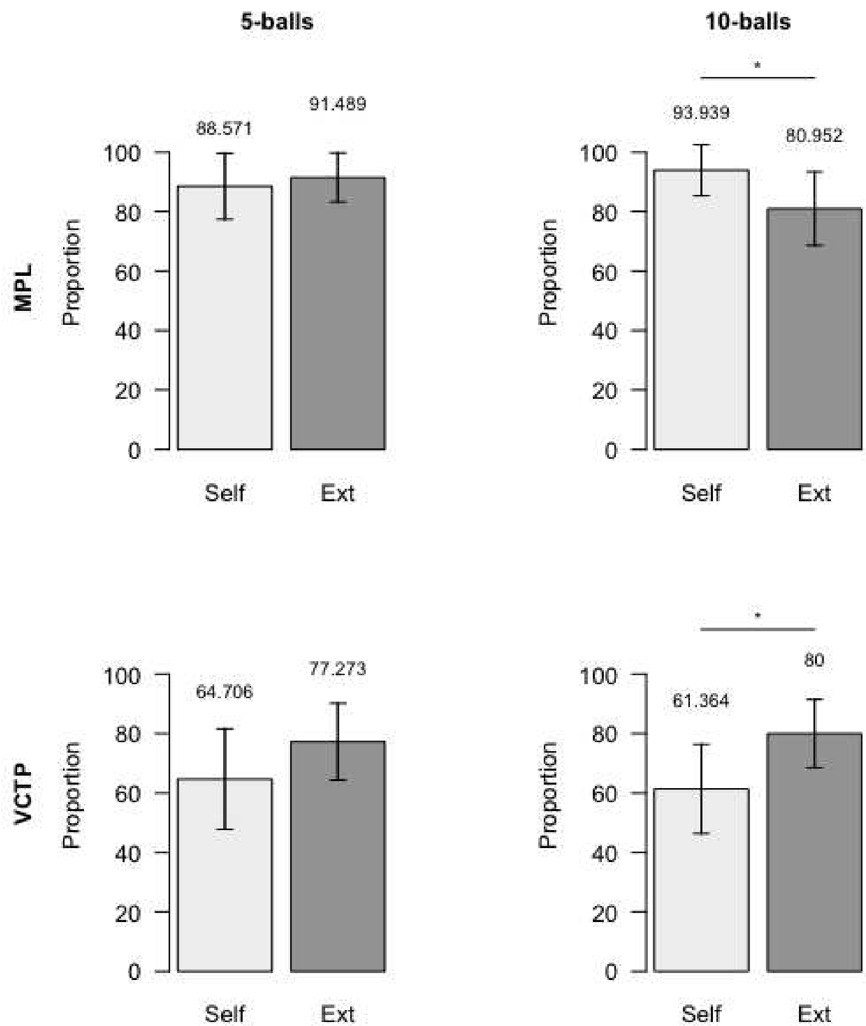


Figure A6: Consistency by type of management in the field according to task and number of balls

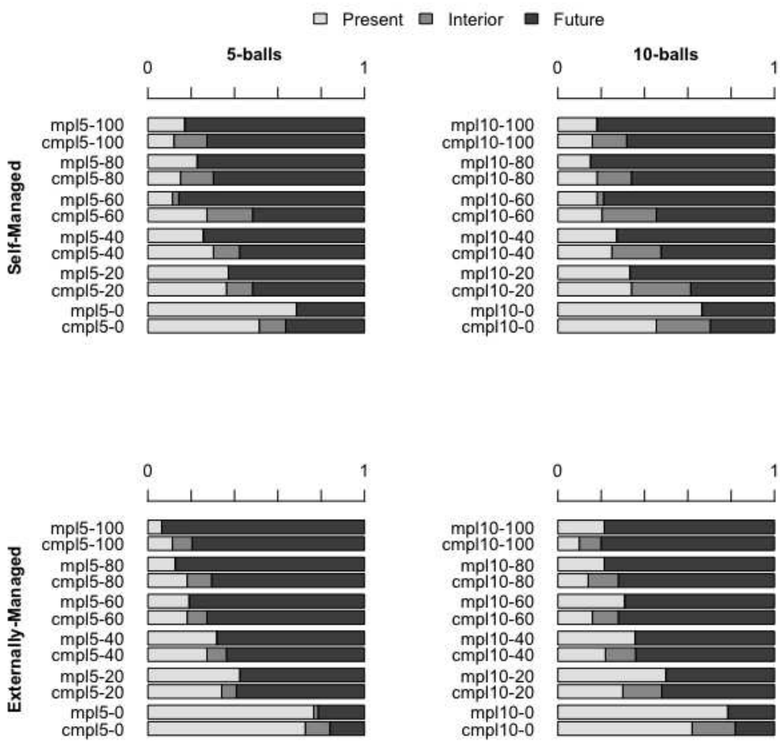


Figure A7: Frequency of allocation type by interest rate, task, management type and number of balls in the field for all subjects