Red Nacional de Investigadores en Economía

Lab vs Online Experiments: no Differences

Benjamin Prissé (LoyolaBehLab, Universidad Loyola Andalucía)

Diego Jorrat (LoyolaBehLab, Universidad Loyola Andalucía)

DOCUMENTO DE TRABAJO N° 137

Abril de 2022

Citar como:
Prissé, Benjamin y Diego Jorrat. (2022). Lab vs Online Experiments: no Differences. *Documento de trabajo RedNIE N°137*.

Lab vs online experiments: no differences*

Benjamin Prissé[†], Diego Jorrat[‡]

March 10, 2022

Abstract

We ran an experiment to study whether lack of control, meaning not controlling the experimental environment, has an effect on experimental results. Subjects were recruited following standard procedures and randomly assigned to complete the experiment online or in the laboratory. The experimental design is otherwise identical between conditions. Results suggest that there are no differences between conditions, except for a larger percentage of online subjects donating nothing in the Dictator Game.

Keywords: Time Preferences, CTB, Experiments.

JEL-codes: C91, C93, D15.

*Thanks to Pablo Brañas-Garza and Praveen Kujal for sharing the data. Karen Khachatryan conducted the experiments and Alex Bueno programmed the software. We gratefully acknowledged comments received from participants at ESA-JILAEE conference and INVECO´s annual meeting (UNT). Financial support from Middlesex Business School and Ministerio de Ciencia y Tecnología (PGC2018-093506-B-100) and Excelencia Junta de Andalucía (PY18-FR-0007) is gratefully acknowledged. Authors declare they have no competing interests.

[†] *LoyolaBehLab*, Universidad Loyola Andalucía. E-mail: benjamin.prisse@gmail.com. Address: Calle Escritor Castilla Aguayo, 4, 14004 Córdoba, Spain.

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

1 INTRODUCTION

There is an increasing number of papers running experiments online, from Amazon Mechanical Turk (AMT) experiments to tailored platforms or coordinated replications among many labs. AMT attracted a lot of attention to run online experiments by gathering the qualities of these environments. AMT is a crowdsourcing platform where "Turkers" (workers) perform short "HITs" (tasks) proposed by "Requesters" (sponsors) against small remuneration. The main advantage of AMT is to give access to a large and diverse panel of subjects, rapidly and at cheap cost. It also simplifies the logistical burden of running experiments by treating them more like computer simulations. Additionally, experiments using these platforms have desirable features that improve data quality: subjects are paid only if requesters approve their answers, questions verifying the attention of subjects can be added, subjects are not disclosed that they participate in an experiment, performing multiple answers is difficult because each ID is linked to an unique credit card number, subjects can be sorted by their qualifications (nationality, gender, age...) or reputation (requiring ≧90% approval rate is common) to obtain the desired sample.

Experimental economists are concerned about the validity of these experiments because they differ from traditional experiments in at least three dimensions: stakes of payoffs, heterogeneous samples and lack of control. Resolving these three issues would help to validate the use of online experiments.

The first issue refers to whether small payoffs of online experiments influence results. Ipeirotis (2010) estimated the average wage of subjects at $4.80[1] per hour and Horton and Chilton (2010) estimated the reservation wage of subjects at $1.38 per hour[2], which is much lower than usual payments of laboratory experiments. But Mason and Watts (2009) and Marge, Banerjee and Rudickny (2010) suggest that AMT wages do not influence quality of results for tasks like image ordering, word puzzles and transcription. According to Mason and Watts (2009), increasing wages only increases the quantity of work. These results are in line with the following papers in this section finding that economic paradigms display similar results in online environments despite the much lower stakes, but also the meta-analysis of Camerer and Hogarth (1999) concluding that stakes have little influence on results of traditional experiments.

The second issue refers to the sample pool generally being more heterogeneous than traditional laboratory subjects, potentially affecting results of economic paradigms. Berinsky, Huber and Lenz (2012) indeed showed that despite Turkers being more representative of the US population than the population of a convenience sample in political sciences, they are still a distorted representation of the general population by being younger, more liberal and more cognitive. Different pools of subjects is not necessary an issue because Exadaktylos, Espín

[1] Corresponding to a wage of $0.08 per minute
[2] Corresponding to a wage of $0.023 per minute

and Brañas-Garza (2013) showed that self-selected students answer the Dictator, Ultimatum and Trust Games similarly than a representative sample of the population of a city.

If differences exist, it would most likely exist economic paradigms detecting them. Regarding cognitive biases and logical fallacies in online experiments, Paolacci, Chandler and Ipeirotis (2010) showed that online subjects also display risk-aversion in gain, risk-seeking in losses, conjunction fallacy and outcome bias. Goodman, Cryder and Cheema (2013) reproduced the result of risk-aversion in gains and risk-seeking in losses, also finding that online subjects display present bias, delay-speedup asymmetry and certainty effect.

Several papers showed that online subjects display the same behavior in economic games as traditional participants. The validity of the online environment to gather datas for a single economic game was demonstrated by Horton, Rand and Zeckhauser (2011) showing identical behavior between Turkers and traditional subjects in the Dictator Game, and replicating the phenomenon of increased cooperation among believer subjects by religiously priming them. Brañas-Garza, Capraro and Rascón Ramírez (2018) collected more than 3500 observations on MTurk to show that the behavior of online subjects in the Dictator Game is similar to standard subjects. Jorrat (2021) obtained substantial cooperation in the Prisoner Dilemma with online subjects, and also increased cooperation when priming it. Brañas-Garza, Jorrat, Kovářík and López (2021) studied behavior of university students in the Dictator Game for stakes of €5, €100 and €1000. The €5 and €1000 conditions were ran in the Lab but the €100 condition had to be run online because of the Covid-19 pandemic. The overall pattern of results indicates a reduction of altruistic behavior when share increases, with the behavior of online subjects answering the €100 condition being consistent with their expected behavior in the Lab.

Others papers showed the possibility of collecting datas with more ambitious designs. Chesney, Chuah and Hoffman (2009) showed that *Second Life* players display behavior in line with laboratory results in the Dictator Game, the Ultimatum Game, the Public Good Game, the Minimum Effort Game and the Guessing Game. Amir, Rand and Kobi Gal (2012) studied the behavior of Turkers when answering social elicitation games. Subjects answered with and without stakes the Dictator Game, the Ultimatum Game, the Trust Game and the Public Good Game. They found that stakes influence results in the expected directions and give results comparable to standard experiments. Arechar, Gächter and Molleman (2018) studied the repeated Public Good Game with and without punishment, showing that the standard behavioral patterns of cooperation and punishment are replicated online. Results also suggest that online subjects are more prosocial : they are more cooperative without punishment, and when given the opportunity to punish they do so less frequently and intensively. Additionally, the high attrition rate of 40% was unrelated with trial outcomes and thus does not threaten the internal validity of datas. In conclusion, these papers demonstrated the ability to obtain qualitative datas with challenging online experiments. They also suggest that heterogeneity of subjects in online environments does not threaten the quality of datas.

This paper focuses on the third issue that refers to experimenters not controlling the experimental environment in which subjects are replying to the experiment. In the context of economic experiments, it means that we are particularly concerned by the decision-making of subjects being affected by specific factors of the online environment: subjects may not attentively read the experimental instructions or be significantly distracted by the environment during the experiment. Conversely, they may make their decisions without taking the time to think about them or answer automatically if they are no longer interested by the experiment. They may also receive help from a partner or use search engines to find answers to questions. And they are more likely to drop-out of the experiment, with this attrition potentially be selective. It is therefore necessary to collect datas suggesting that not controlling the experimental environment does not threaten the quality of datas.

As far as we know, this question has only been investigated by Hergueux and Jacquemet (2015), whose subjects in the lab and online conditions were university students recruited through the same procedure[3], answered the experiment on the same online interface and were paid with the same monetary stakes. Subjects answered social preferences games : the Public Good Game, the Dictator Game, the Ultimatum Game, the Trust Game and the Holt-Laury risk-aversion task. Results showed that subjects answer tasks similarly in both environments and are not more likely to drop-out from the experiment. Questionnaire results also showed that online subjects have similar beliefs on the experiment and self-reported social preferences than laboratory subjects. However, online subjects diverged on three points. First, they answer more rapidly than laboratory subjects, potentially affecting results by increasing the influence of emotions in the decision-making. Second, results contradict the theory of prosociality decreasing with social distance. Online subjects are more altruistic in the Dictator Game, more trusting and trustworthy in the Trust Game. It suggests that the inherent necessity of trust in online transactions outperforms social distance. Third, online subjects are less risk-averse in the Holt-Laury task, behaving more like expected utility maximizers. The experimental design could not identify factors responsible for these differences between conditions. Additionally, the payment method diverged between conditions, potentially influencing results. This paper therefore called for additional comparisons between the lab and online environments to obtain more evidences that lack of control is not an issue.

This is precisely the goal of our experiment. We ran an experiment in which the experimental environment was the only difference between conditions. The recruitment procedure, experimental interface, experimental tasks and payment

[3]Arechar et al (2018) is lacking this characteristic.

method were otherwise strictly identical between conditions. We made subjects perform several standard economic tasks related to the measurement of economic preferences, creating a rich environment likely to elicit any potential differences. We focused on time preferences to study a more general characteristic of decision-making than social preferences. We find that results are identical between conditions, except for laboratory subjects donating more frequently to charities in the Dictator Game. We conclude that lack of control is not an issue and that online environments are as valid as laboratory environments.

2 EXPERIMENTAL DESIGN

The key feature of our design is the use of two experimental treatments only differing on the location where subjects complete the experiment:

- Laboratory: subjects completed the experiment in a standard experimental economics laboratory.
- Online: subjects completed the experiment at home on their personal computer.

Comparing the two conditions allow us to evaluate the potential effect of not controlling subjects.

2.1 Experimental Tasks

Because the potential differences are numerous, we created a rich environment likely to elicit any potential differences by making subjects successively reply to several standard economic tasks:

- The Convex Time Budget (CTB) of Andreoni and Sprenger (2012).
- The Multiple Price Lists (MPL) of Andreoni et al (2015) in a modified version.
- The risk-aversion task (HL) of Holt and Laury (2002).
- The Dictator Game (DG) with subjects donating their total earnings.
- The Cognitive Reflection Task (CRT) of Frederick (2005)
- A Numeracy (Num) task related to percentages.

We ran two experimental sessions in November 2014 and March 2015[4]. Subjects were not told that they will be randomly assigned to Laboratory or Online conditions when registering to participate. We have 257 subjects who completed the experiment (113 Lab, 144 Online)[5]. This sample size allows us to find an

[4]This experiment was the first ever organized at Middlesex University.

[5]A total of 520 students registered to participate in the experiment. The attrition rate was 50.58% (52.92% Lab, 48.57% Online) with a t-test not rejecting the equality between conditions (p=0.217).

effect size of 0.3 SD or higher with a 90% significance level. Any effect below this threshold could be considered a low effect according to Cohen´s effects size. Subjects were on average 23.16 years old and 56.97% of them were female. The average payment was £19.48.

2.2 Questionnaire

Subjects also answered a questionnaire measuring control variables, allowing us to verify if comparing the two conditions is meaningful. First, we verify that subjects are alone when answering the Online experiment. Regarding the location of reply, 75% reply at home, 18.75% at university, 5% somewhere else and 1.25% indicated nothing. It suggests that a significant proportion of subjects could be in presence of others, but when asked about their social environment 98.60% reported being "Alone" or "Mostly Alone", suggesting that Online subjects are as isolated as Lab subjects.

Second, it allows us to verify that samples are comparable. We run difference mean test between groups to check the balance in different control variables. Table 1 provides results of these tests, with Romano-Wolf adjusted $p-value$ for multiple testing in the last column. It suggests that the two groups are similar in their characteristics ($p \geq$0.238), allowing us to estimate the causal effect of online setting on results.
We conclude that the two groups are comparable and that the experiment measures the causal effect of experimental conditions.

Table 1: Balance check

	n	$mean_L$	$mean_O$	$L-O$	$p-value$	$adj.p-value$*
Exercise	255	0.57	0.42	-0.15	0.016	0.238
Medical check up	254	0.71	0.62	-0.09	0.109	0.711
Smoke	256	0.18	0.21	0.03	0.513	0.987
Weight	213	77.27	73.85	-3.42	0.444	0.983
Height	206	166.57	163.20	-3.37	0.215	0.855
Work	256	0.44	0.45	0.01	0.936	0.987
English native	255	0.44	0.58	0.14	0.032	0.374
Age at admission	251	21.42	21.15	-0.27	0.650	0.987
Female	251	0.56	0.58	0.02	0.752	0.987
Numeracy score	257	4.40	4.60	0.21	0.180	0.840
CRT score	224	0.57	0.81	0.24	0.060	0.535
Trust - experimenter	252	0.87	0.89	0.02	0.581	0.987
Trust - donation	243	5.14	4.90	-0.24	0.318	0.937
All nighters	255	9.73	19.71	9.98	0.028	0.355
Want credit card	252	17.43	9.79	7.64	0.075	0.600

* Refers to Romano-Wolf step-down adjusted p-values.

2.3 Methods

Throughout our analysis, we estimate the following simple linear regression model to identify the causal effect of online setting on different outcome variables:

$$y_i = \beta_0 + \beta_1 * online_i + \boldsymbol{\gamma_j} * \boldsymbol{X_{ji}} + \epsilon_i \tag{1}$$

With y_i the outcome variable for each individual, $online_i$ the dummy variable that takes value 1 when subjects answer the experiment in the online environment and 0 when subjects answer the experiment in the laboratory, $\boldsymbol{X_{ji}}$ the vector of control variables and ϵ_i the error term. We will use as control variables *age* which is the age of subjects (in years), *gender* taking value 1 if the subject is female and 0 otherwise, *crt score* which is the number of correct answers of subjects in the CRT and *numeracy score* which is the number of correct answers of subjects in the numeracy task. We will use as dependent variables in our regressions *response time*, *consistency* and *preferences* of subjects in each tasks. We define *response time* as the amount of time (in minutes) needed by subjects to answer the task. We define *consistency* according to the task. Subjects are consistent in MPL and HL if they do not make multiple switching. In both tasks, subjects are supposed to initially choose the left column until they switch and choose the right column. Choosing the left column after choosing the right column is an inconsistent choice named multiple switching. In CTB, consistency is complex to define because the 45 budget choices create 920 opportunities to make an inconsistent choice, rendering too strict the usual requirement of complete consistency. Because a single trial can rapidly increase the number of inconsistent choices, we give some margin of errors to subjects by considering them consistent if at least 80% choices are consistent. We define *preferences* of subjects according to the task. In the CTB task, we are interested by the number of tokens allocated to the early period. In the MPL task, we are interested by the number of early choices made by subjects. In the HL task, we look at the number of safe choices made by subjects and in the DG task the share of monetary payoffs donated to charities.

3 RESULTS

The following section displays the results for response time, consistency and preferences in each task. We find almost no differences between Lab and Online conditions.

3.1 Response Times

Table 2 displays the regressions results for *response time*, suggesting that time needed to answer a task is not influenced by the experimental environment[6].

[6] We adjusted p-values for multiple hypothesis testing in Table 1 because we had significant results. Table 2 does not have significant results.

Table 2: OLS estimations of the impact of online setting on response time.

	(1) TimeCTB	(2) TimeCTB	(3) TimeMPL	(4) TimeMPL	(5) TimeHL	(6) TimeHL	(7) TimeDG	(8) TimeDG
Online	-0.676	-1.568	-0.069	-0.005	0.311	0.247	0.032	0.075
	(1.213)	(1.267)	(0.134)	(0.148)	(0.426)	(0.389)	(0.085)	(0.089)
	[0.578]	[0.217]	[0.604]	[0.973]	[0.467]	[0.527]	[0.704]	[0.401]
Constant	18.651***	23.570***	2.515***	2.255***	5.380***	3.371**	0.991***	0.554
	(0.861)	(4.425)	(0.087)	(0.548)	(0.192)	(1.454)	(0.052)	(0.342)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.021]	[0.000]	[0.106]
Observations	256	218	256	218	254	218	254	219
Adj R-squared	-0.003	0.052	-0.003	0.012	-0.002	0.018	-0.003	0.005
Controls	No	Yes	No	Yes	No	Yes	No	Yes

Note: Time in minutes. Robust standard errors in parentheses and p values in brackets. Adjusted R-Squared are taken from regressions without robust standard errors. Asterisks denote significance level: $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

Column 1 shows that subjects answer on average the CTB task in 18.65 minutes in the Lab and 17.97 minutes in the Online environment[7]. The *online* variable is not significant ($p = 0.578$) and Column 2 shows the same result when adding controls ($p = 0.217$). The equality of distributions across conditions is not rejected by a Mann-Whitney-Wilcoxon test ($p = 0.215$). Column 3 shows that subjects answer on average the MPL task in 2.52 minutes in the Lab and 2.45 minutes in the Online environment with the *online* coefficient not significant ($p = 0.604$)[8]. Column 4 shows that this result is robust to adding controls ($p = 0.973$) and the equality of distributions across conditions is not rejected by a Mann-Whitney-Wilcoxon test ($p = 0.118$). Column 5 shows that subjects answer on average the HL task in 5.38 minutes in the Lab and 5.69 minutes in the Online environment[9]. The *online* coefficient is not significant ($p = 0.467$)) and Column 6 shows that this result is robust to adding controls ($p = 0.527$). A Mann-Whitney-Wilcoxon also shows equality of distributions across conditions ($p = 0.143$). Finally, column 7 shows that subjects take 0.991 minutes to answer the DG in the Lab and 1.023 minutes Online with the *online* coefficient not significant ($p = 0.704$)[10]. Column 8 shows that this result is robust to adding control ($p = 0.401$). We also have a Mann-Whitney-Wilcoxon test not rejecting the equality of distributions across conditions ($p = 0.224$).

Further analysis are presented in the Appendix. Section B.1.1 shows that re-

[7] We excluded one outlier in the Online condition because he took 68.70 minutes to answer the CTB task

[8] We excluded one outlier in the Lab condition because he took 13.53 minutes to answer the MPL task

[9] We excluded three outliers in the Online condition because they respectively took 192.50, 202.30 and 242.08 minutes to answer the HL task

[10] We excluded one outlier in the Lab condition because he took 6.77 minutes to answer the DG task. We also excluded two outliers in the Online condition because they respectively took 6.88 and 9.60 minutes to answer the task.

sponse time in CTB is not different across cognitive levels or by whether subjects are below or above the median time. Regarding the MPL, HL and DG tasks, sections B.1.2, B.1.3 and B.1.4 also show no differences across cognitive levels in these tasks. Additionally, all the effects we found are less than the minimum detectable effect of 0.3 SD in this experiment. Indeed, the effects in absolute terms represent a variation between 0.2% and 8% with respect to the mean of the outcome variables in the Lab. If we increased the number of observations to $n = 1000$, some differences would be significant but their magnitudes are not economically important.

We conclude that the online environment has no impact on average response times.

3.2 Consistency

Table 3 displays the results of regressions on *consistency*, suggesting that is not influenced by the experimental environment. Column 1 shows that we have 40.70% consistent subjects in the Lab and 38.19% consistent subjects in the Online environment for the CTB task. The *online* variable is not significant ($p = 0.684$) and Column 2 shows that this result holds when adding controls ($p = 0.756$). Column 3 shows that we have 87.61% consistent subjects in the Lab and 92.36% in the Online condition for the MPL task, with this difference not being significant ($p = 0.215$). Column 4 shows that this result is robust to adding controls ($p = 0.317$). Column 5 shows that we have 54.01% consistent subjects in the Lab and 57.71% in the Online condition for the HL task, with the *online* coefficient not significant ($p = 0.560$). Column 6 shows that this result holds when adding controls ($p = 0.615$).

Further analysis in Section B.2.1, Section B.2.2 and Section B.2.3 respectively show for the CTB, MPL and HL tasks that we do not find differences across cognitive levels. Section B.2.1 also shows that Online subjects make similar type of inconsistent choices in CTB than Lab subjects, and Table S1 shows that Online subjects are similarly consistent in CTB than Lab subjects when using more strict or more tolerant definitions of consistency in CTB. Section B.2.2 and Section B.2.3 also shows that we do not find differences in consistency across each MPL and HL tasks between conditions. As before, the effects we found are not significant and would not be economically important if we had a larger number of observations because they represent a variation between 5% to 6% with respect to the mean of the outcome variables in the Lab.

We conclude that the online environment has no impact on consistency.

Table 3: OLS estimations of the impact of online setting on consistency in the CTB, MPL and HL tasks.

	(1) ConsCTB	(2) ConsCTB	(3) ConsMPL	(4) ConsMPL	(5) ConsHL	(6) ConsHL
Online	-0.025	-0.021	0.048	0.039	0.037	0.032
	(0.062)	(0.067)	(0.038)	(0.039)	(0.063)	(0.063)
	[0.684]	[0.756]	[0.215]	[0.317]	[0.560]	[0.615]
Constant	0.407***	-0.061	0.876***	0.681***	0.540***	0.001
	(0.046)	(0.218)	(0.031)	(0.146)	(0.047)	(0.221)
	[0.000]	[0.781]	[0.000]	[0.000]	[0.000]	[0.995]
Observations	257	219	257	219	257	219
Adj R-squared	-0.003	0.026	0.002	0.026	-0.002	0.121
Controls	No	Yes	No	Yes	No	Yes

Robust standard errors in parentheses and p values in brackets. Adjusted R-Squared from regressions without robust standard errors. Asterisks denote significance level: $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

3.3 Preferences

3.3.1 Time and risk preferences

Finally, we look at whether answers of subjects in each task (i.e revealed preferences) are different between the Lab and Online conditions. Table 4 displays regressions on number of tokens allocated to the early period in the CTB task, number of early period choices in the MPL task and number of safe choices in the HL task.

Column 1 suggests that on average subjects allocate 33.219 tokens to the early period in the Lab and 33.874 tokens to the early period Online, with the *online* variable not being significant ($p = 0.804$). Column 2 shows that this result holds when adding controls ($p = 0.708$). Figure 1 displays the average amount of tokens allocated to early periods for each combination of start date t and delay k between conditions, showing that the overall patterns are similar between conditions. Further analysis in Section B.3.1 of the online appendix suggest that there are no differences in the number of tokens allocated to early period across cognitive levels, that subjects similarly use present, interior and future allocations between conditions, and that subjects similarly increase their allocations to the future when the interest rate increases between conditions.

Column 3 shows that subjects make on average 9.212 early period choices in the Lab and 9.194 early period choices in the Online environment, with the *online* coefficient not being significant ($p = 0.976$). Column 4 shows that this result holds when adding controls ($p = 0.434$). Figure 2 shows that subjects display similar behavior in each MPL tasks between conditions, rapidly switching to

Table 4: OLS estimations of the impact of online setting on choices of subjects.

	(1) ErlCTB	(2) ErlCTB	(3) ErlMPL	(4) ErlMPL	(5) SafeHL	(6) SafeHL
Online	0.655	1.039	-0.018	0.514	0.954*	0.737
	(2.637)	(2.773)	(0.608)	(0.656)	(0.524)	(0.582)
	[0.804]	[0.708]	[0.976]	[0.434]	[0.070]	[0.206]
Constant	33.219***	46.459***	9.212***	5.652**	11.053***	14.240***
	(1.951)	(9.675)	(0.468)	(2.550)	(0.388)	(1.947)
	[0.000]	[0.000]	[0.000]	[0.028]	[0.000]	[0.000]
Observations	257	219	257	219	257	219
Adj R-squared	-0.004	0.036	-0.004	0.053	0.009	0.020
Controls	No	Yes	No	Yes	No	Yes

Robust standard errors in parentheses and p values in brackets. Adjusted R-Squared are taken from regressions without robust standard errors. Asterisks denote significance level: $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

later allocations when the interest rate becomes sufficient. Further analysis in Section B.3.2 of appendix show that we do not find differences across cognitive levels, Table S2 shows that subjects make similar number of early choices in each MPL tasks between conditions and Figure S11 shows that consistent subjects display similar behavior in each MPL tasks between conditions.

Column 5 shows that subjects make on average 11.053 safe choices in the Lab and 12.007 safe choices Online, with a t-test marginally rejecting equality between conditions ($p = 0.070$). Column 6 shows that this result disappears when adding controls ($p = 0.206$) and a Mann-Whitney-Wilcoxon test does not reject the equality of distributions between environments ($p = 0.126$). Figure 3 shows that subjects display similar behavior in each HL tasks between conditions. The dotted line in each graph display choices of the risk-neutral individual, showing that subjects in both conditions have a significant amount of risk-aversion that is consistent with standard results. Further analysis in Section B.3.3 show that we do not find differences across cognitive levels, Table S3 shows that subjects make a similar number of safe choices in each HL tasks between conditions and Figure S14 shows that consistent subjects display similar behavior in each HL tasks between conditions.

We conclude that the online environment does not affect preferences in the CTB, MPL and HL tasks.

In order to suggest that this result of no differences between conditions has external validity, we need to show that results of Lab subjects are comparable to standard results of the literature. We therefore compare choices of Lab subjects in the CTB with choice of AS subjects. Then, we compare choices of Lab

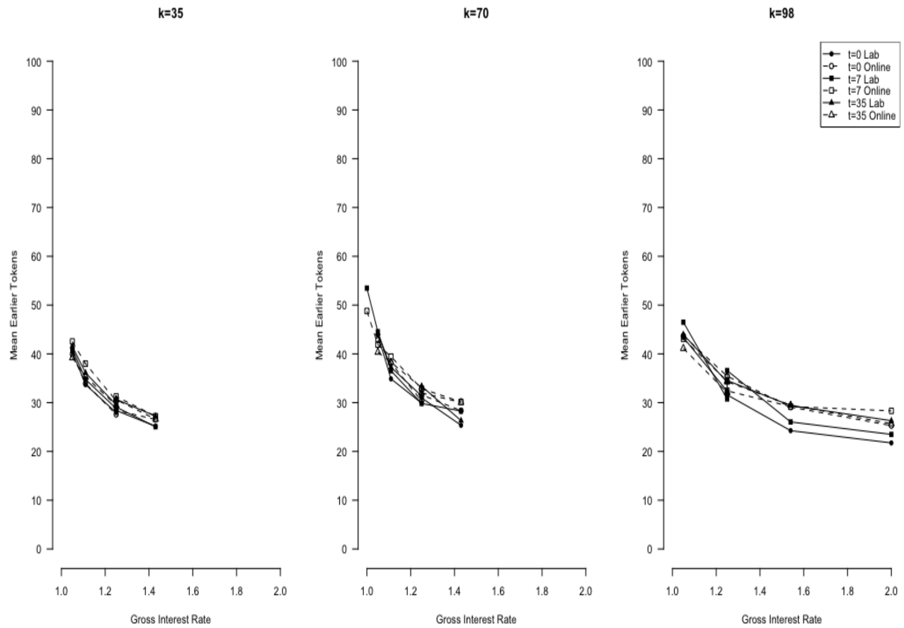


Figure 1: Mean allocations of tokens to the early period in the CTB task by condition

subjects in the MPL and HL with subjects of Andreoni et al (2015) (AKS) who answered a similar experiment. Because 90.63% of AKS subjects are consistent in both tasks, we only include in this analysis Lab subjects that are consistent in the studied task.

Figure S9 shows that Lab subjects allocate slightly differently than AS subjects in the CTB: they initially make less allocations to present but maintain higher allocations to present when interest rate increases. This pattern is reflected in the average amount of tokens allocated to early periods by each group: Lab subjects allocate on average 33.59 tokens to early periods and AS subjects allocate on average 27.58 tokens to early periods, with a t-test rejecting the equality between groups at 1% ($p < 0.001$). Figure S12 shows that Lab subjects display similar behavior than AKS subjects in each MPL tasks, except that they make more early period choices. This result is further supported by t-tests rejecting equality of the number of early choices between groups in MPL1 (Lab: 2.44, AKS: 1.44, $p < 0.001$), MPL2 (Lab: 3.64, AKS: 2.21 , $p < 0.001$) and MPL3 (Lab: 1.97, AKS: 1.25, $p = 0.002$). It is noteworthy that Lab subjects answered MPL3 with start date and delay (t=0, k=98) while AKS subjects answered for (t=0, k=63). Comparing them is still relevant because such long term delays are similarly discounted, although k=98 should generate slightly higher discount. It

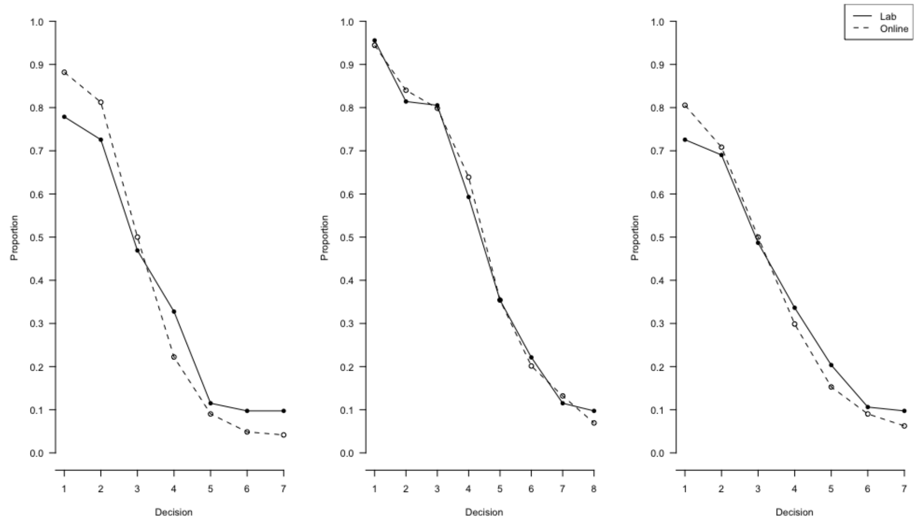


Figure 2: Average number of early choices in each MPL tasks by condition

is therefore a bit more meaningful to find that Lab subjects make more early choices. Figure S15 shows that Lab subjects have a similar behavior than AKS subjects in each HL tasks, with t-tests not rejecting equality of the number of safe choices between groups in HL1 (Lab: 5.79, AKS: 5.56, $p = 0.502$) and HL2 (Lab: 5.62, AKS: 5.37, $p = 0.450$).

We conclude that Lab subjects display similar behavior than standard subjects in the HL task. They also display similar patterns than standard subjects in the CTB and MPL tasks, but make more early period choices.

3.3.2 Altruism

We now turn our attention to the Dictator Game. The endowment of subjects in the DG was their total earnings and they were informed of this when answering the task. Figure 4 presents the distribution of endowments across conditions. We see that means ($p = 0.384$) and distributions (Mann-Whitney-Wilcoxon test, $p = 0.513$) are similar across conditions. Additionally, we find an almost null correlation of r=-0.073 between the endowment of subjects and the percentage of their earnings donated to charity. We follow the analysis made by Brañas-Garza et al (2021) and refer to share as the percentage of earnings donated. Table 5 shows regression results for the DG. Column 1 shows that Lab subjects on average donate 14.87% of their earnings to charities and Online subjects on average donate 11.54% of their earnings, with a t-test not rejecting the equality between condition ($p = 0.292$). Column 2 shows that the *online*

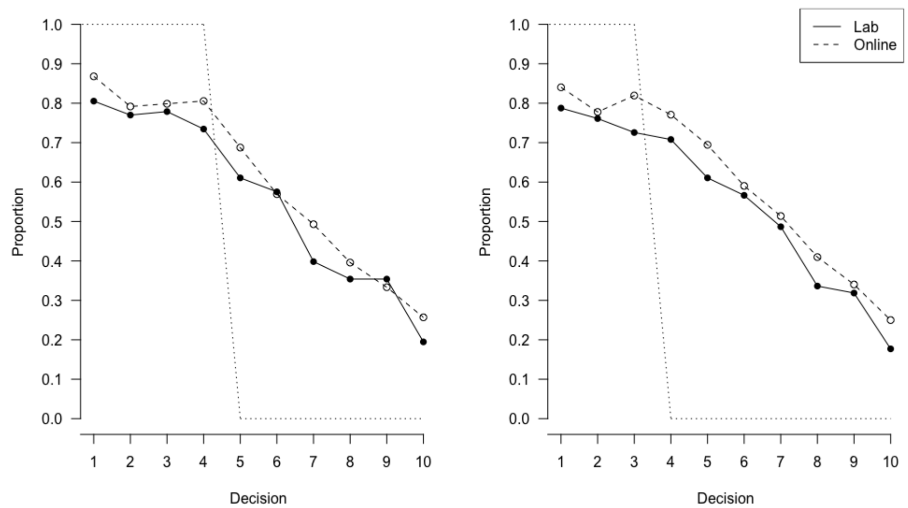


Figure 3: Average number of safe choices in each HL tasks by condition

variable has a negative and marginally significant effect when adding control ($p = 0.070$). Column 3 shows that Lab subjects on average donate €3.42 to charities and Online subjects on average donate €3.18, with a t-test not rejecting the equality between conditions ($p = 0.744$). Column 4 shows that this result is robust to adding controls ($p = 0.273$). Column 5 shows that 38% of Lab subjects and 49.9% of Online subjects donate nothing to charities, with a t-test marginally rejecting the equality between conditions ($p = 0.055$). Column 6 shows that this result becomes significant at 5% when adding controls ($p = 0.033$) and is still marginally significant after correcting the p-value for multiple testing with the Romano-Wolf procedure ($p = 0.088$). The result that the only differences between conditions is the decision to donate is further supported by Mann-Whitney-Wilcoxon tests rejecting the equality of distributions of *share* ($p = 0.020$) and *giving* ($p = 0.057$) variables between conditions. Further analysis in Section B.3.4 support this result by showing that decision to donate to charities and percentage of earnings donated to charities are similar across cognitive levels, and Table S4 shows that there are no differences in donations between conditions for subjects donating less than 50%, 50%, more than 50% or 100% of their earnings.

A potential explanation of our results is the Hawthorne effect: subjects in the Lab condition display more generous behavior because they are in presence of others. Table S5 supports this result by running the same regressions than Table 5 after replacing the *online* explanatory variable by the *withothers* variable.

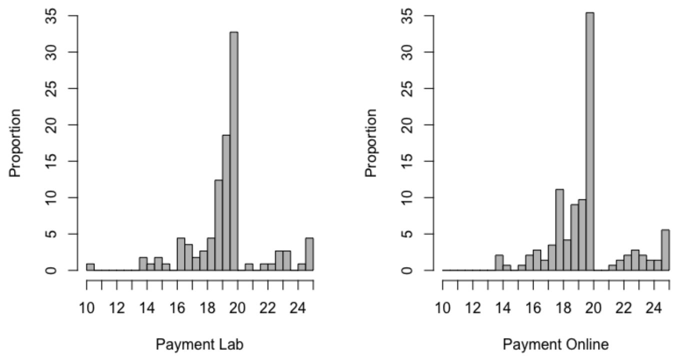


Figure 4: Distribution of experimental payment of subjects

It takes value 1 if subjects declare being "With Others"[11] when answering the experiment or value 0 if subjects declare being "Alone" or "Mostly Alone". Column 5 shows that presence of others significantly decrease the probability to not donate anything by 20% ($p = 0.019$) and Column 6 shows that adding controls make this result significant at 1% ($p = 0.002$). This result remains significant at 5% after correcting the p-value for multiple testing with the Romano-Wolf procedure ($p = 0.028$).

We conclude that the online environment decreases the share of subjects donating in the DG, potentially because feeling the presence of others incite to donate to charities.

In order to validate our results, we compare them with standard results found in the literature. Figure 5 shows the multi-histograms of the share of experimental payments donated to charities by conditions. We see that the overall pattern of the distribution is similar to the one of Figure S18 from the meta-analysis of Engels (2011), except that we have more subjects donating 0% of their earnings and less subjects donating 50% of their earnings. We find that Lab subjects donate on average 14.87% of their earnings while Engels (2011) found in his meta-analysis that subjects donate on average 28.3% of their earnings, with a t-test rejecting the equality of means at 1% ($p < 0.001$). We also find that Online subjects donate on average 11.53% of their earnings while Brañas-Garza, Capraro and Rascón (2018) found that MTurk subjects donate on average 27.3% of their earnings, with a t-test rejecting the equality of means at 1% ($p < 0.001$). We find that our subjects donate less than usual in the DG, however Figure S19 from the meta-analysis of Engels (2011) shows that an average donation of 13% in the Dictator Game is not dramatically low. A potential explanation of our

[11] 31 Lab subjects, 2 Online subjects.

results is that subjects of standard DG experiments are given their endowments, while in our experiment subjects made a large effort to obtain it. They should therefore be less willing to donate.

Table 5: OLS estimations of the impact of online setting on share, giving and type of altruistic behavior in the charity dictator game.

	(1) Share	(2) Share	(3) Giving	(4) Giving	(5) Selfish	(6) Selfish
online	-3.332	-5.776*	-0.261	-0.919	0.119*	0.145**
	(3.158)	(3.169)	(0.798)	(0.837)	(0.062)	(0.068)
	[0.292]	[0.070]	[0.744]	[0.273]	[0.055]	[0.033]
Constant	14.867	-12.999	3.420***	-5.197	0.380*	0.706***
	(2.365)	(13.487)	(0.562)	(3.723)	(0.046)	(0.227)
	[0.292]	[0.336]	[0.000]	[0.164]	[0.055]	[0.002]
Observations	257	219	257	219	257	219
Adj R-squared	0.000	0.065	-0.004	0.076	0.010	0.037
Controls	No	Yes	No	Yes	No	Yes

Robust standard errors in parentheses and p values in brackets. Adjusted R-Squared from regressions without robust standard errors. Asterisks denote significance level: $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

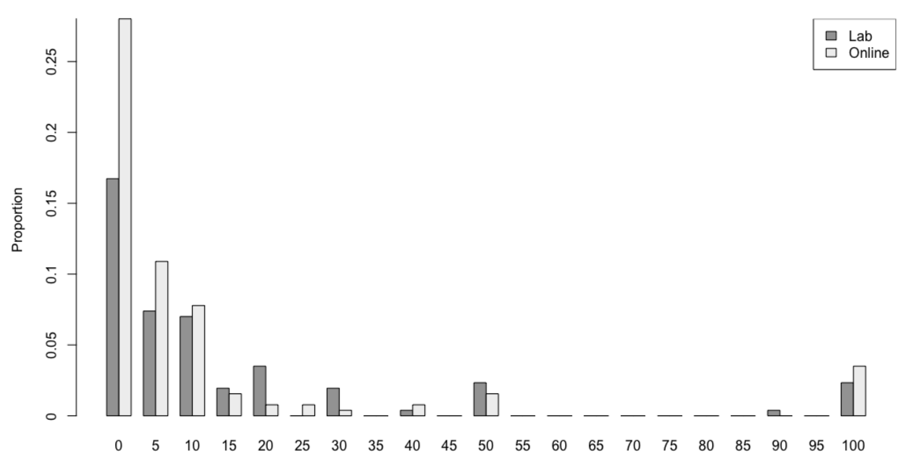


Figure 5: Distribution of donations in the DG by conditions

4 DISCUSSION

In this paper, we investigated whether not controlling the experimental environment influence experimental results. We designed a rich environment made of several standard economic tasks to elicit any potential differences between subjects, then subjects answer these tasks in the Laboratory or Online environments. Results suggest that there are no differences between conditions in the CTB, MPL and HL tasks, but we find that a higher share of Lab subjects donate in the DG. A potential explanation is that the Lab environment influence subjects to modify their behavior because of the presence of others, suggesting Hawthorne effect. This interpretation is supported by regressions showing that being in presence of others decreases the likeliness of donating nothing in the DG. We also compared results of Lab subjects with standard results from the literature and found similar patterns of results in the CTB, MPL and HL tasks, although Lab subjects allocate slightly more to early periods in the CTB and MPL tasks. We find that our subjects donate a smaller share of their earnings than standard subjects in the DG, a potential explanation being that the large effort required by the experiment reduces the willingness of subjects to donate. Overall, we find that there are mostly no differences between conditions and that our subjects display behavior in line with results of the literature.

Comparing our results with the ones of Hergueux and Jacquemet (2015), we remark that they differ on several points. First, we find no differences in response times between conditions for each task, while they find that Online subjects answer the experiment faster than Lab subjects. Second, we find that Online subjects make a similar number of safe choices in the HL task than Lab subjects, while they find that Online subjects make less safe choices in the HL task than Lab subjects. Third, we find that Online subjects are more likely to donate nothing in the DG than Lab subjects, but donate similar share and amounts when giving, while they find that Online subjects are twice more generous than Lab subjects and are also more prosocial in the Ultimatum Game and Trust Game. A potential explanation is that their experiment was focusing on prosocial preferences with subjects being given their endowments, while our subjects answered the DG after a long experiment in which they earned their endowment. It was therefore more costly for our subjects to donate. This comparison therefore suggest that our paper make two contributions complementing the findings of Hergueux and Jacquemet (2015): the first contribution is to find that larger social distance increases selfishness, the second contribution is to obtain results with mostly no differences between Lab and Online environments.

We conclude that results validate the use of online experiments by showing that not controlling the experimental environment does not influence response time, consistency and answers of subjects in several standard economic tasks. More frequent donations to charities in the Lab environment suggest that future studies should investigate whether online environment elicit more truthful behavior in social preferences from subjects by removing the Hawthorne effect.

REFERENCES

Amir, O., Rand, D. G., & Gal, Y. A. K. (2012). Economic games on the internet: The effect of $1 stakes. *PloS ONE, 7*(2), e31461.

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.

Arechar, A. A., Gächter, S., & Molleman, L. (2018). Conducting interactive experiments online. *Experimental Economics, 21*(1), 99-131.

Berinsky, A. J., Huber, G. A., & Lenz, G. S. (2012). Evaluating online labor markets for experimental research: Amazon. com's Mechanical Turk. *Political Analysis, 20*(3), 351-368.

Brañas-Garza, P., Capraro, V., & Rascón-Ramírez, E. (2018). Gender differences in altruism on Mechanical Turk: Expectations and actual behaviour. *Economics Letters, 170*, 19-23.

Brañas-Garza, P., Jorrat, D., Kovářík, J., & López, MC. (2021). Hyper-altruistic behavior vanishes with high stakes. *PLoS ONE, 16*(8), e0255668.

Brañas-Garza, P., Kujal, P., & Lenkei, B. (2019). Cognitive reflection test: Whom, how, when. *Journal of Behavioral and Experimental Economics, 82*, 101455.

Camerer, C. F., & Hogarth, R. M. (1999). The effects of financial incentives in experiments: A review and capital-labor-production framework. *Journal of Risk and Uncertainty, 19*(1), 7-42.

Chesney, T., Chuah, S. H., & Hoffmann, R. (2009). Virtual world experimentation: An exploratory study. *Journal of Economic Behavior & Organization, 72*(1), 618-635.

Engel, C. (2011). Dictator games: A meta study. *Experimental Economics, 14*(4), 583-610.

Exadaktylos, F., Espín, A. M., & Branas-Garza, P. (2013). Experimental subjects are not different. *Scientific Reports, 3*(1), 1-6.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives, 19*(4), 25-42.

Goodman, J. K., Cryder, C. E., & Cheema, A. (2013). Data collection in a flat world: The strengths and weaknesses of Mechanical Turk samples. *Journal of Behavioral Decision Making, 26*(3), 213-224.

Hergueux, J., & Jacquemet, N. (2015). Social preferences in the online laboratory: a randomized experiment. *Experimental Economics, 18*(2), 251-283.

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

Horton, J. J., & Chilton, L. B. (2010, June). The labor economics of paid crowdsourcing. In *Proceedings of the 11th ACM conference on Electronic commerce* (pp. 209-218).

Horton, J. J., Rand, D. G., & Zeckhauser, R. J. (2011). The online laboratory: Conducting experiments in a real labor market. *Experimental Economics, 14*(3), 399-425.

Ipeirotis, P. G. (2010). Analyzing the amazon mechanical turk marketplace. *XRDS: Crossroads, The ACM Magazine for Students, 17*(2), 16-21.

Jorrat, D. (2021). Recruiting experimental subjects using WhatsApp. *Journal of Behavioral and Experimental Economics, 90*, 101644.

Marge, M., Banerjee, S., & Rudnicky, A. I. (2010, March). Using the Amazon Mechanical Turk for transcription of spoken language. In *2010 IEEE International Conference on Acoustics, Speech and Signal Processing* (pp. 5270-5273). IEEE.

Mason, W., & Watts, D. J. (2009, June). Financial incentives and the" performance of crowds". *In Proceedings of the ACM SIGKDD workshop on human computation* (pp. 77-85).

Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on amazon mechanical turk. *Judgment and Decision Making, 5*(5), 411-419.

Online appendix

Lab vs online experiments: no differences

Benjamin Prissé and Diego Jorrat

A Additional Information on the Recruitment and Payment of Subjects

We give additional details on the recruitment and payment of subjects in this section.

Because this experiment was the first ever ran at Middlesex University, we had to build the pool of students interested in participating in economic experiments. In November 2014, we invited by e-mail all the students (1291 individuals) in the Business School of Middlesex University. We obtained replies from 290 individuals who registered to be later contacted for participating in economic experiments. In March 2015, we sent the same invitation e-mail to all the students (21253 individuals) in the six schools of Middlesex University[12]. We obtained replies from 343 subjects who registered to be later contacted for participating in economic experiments.

Among the 520 subjects (240 Lab, 280 Online) who registered to participate in the experiment, 194 subjects (73 Lab, 121 Online) registered to the first experimental session and 326 subjects (167 Lab, 159 Online) to the second experimental session. We have 257 subjects (113 Lab, 144 Online) who completed the experiment, 92 subjects (33 Lab, 59 Online) completed the experiment in the first experimental session and 165 subjects (80 Lab, 85 Online) in the second experimental session. We lost 10 subjects in the first session in what seemed to have been a technical issue. All others subjects who attended the experiment completed it. The attrition rate was 52.58% in the first cohort (54.79% Lab, 51.24% Online) with a t-test not rejecting the equality of attrition rates between conditions (p=0.710), and the attrition rate was 49.39% in the second cohort (52.09% Lab, 46.54% Online) with a t-test not rejecting the equality of attrition rates between conditions (p=0.191). T-tests also do not reject that attrition rates are equal between the two Lab experimental sessions (p=0.899) and the two Online experimental sessions (p=0.55).

The task selected for payment was randomly decided by the roll of a 47-sided dice. If subjects obtained a number between 1 and 45, they were paid according to their choice in the corresponding CTB budget. If they rolled 46, they were paid according to their choice on MPL and rolled another 22-sided dice to determine which decision was chosen for payment. If they rolled 47, they were paid

[12] Art and Design, Business, Health and Education, Law, Media and Performing Art, Science and Technology

according to their choice on the HL task and rolled a 20-sided dice to determine which HL decision was chosen for payment, then rolled another 10-sided dice to determine the result of the random probability.

B Additional Analysis

In this section, we present additional analysis complementing results found in the main paper.

We analyze whether the effect of online setting is heterogeneous or not with the cognitive level of subjects. We assigned each subject to a cognitive level according to their score in the CRT task: low (no correct answer), middle (one correct answer) and high cognitive level (two or three correct answers).

B.1 Response Time

B.1.1 Convex Time Budget

Figure S1 presents results with gray lines specifying 95% CI, confirming that there are no differences in response time by cognitive levels between conditions.

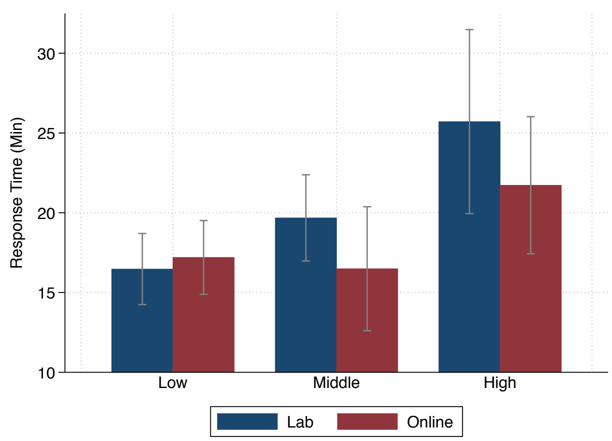


Figure S1: Response Time in the CTB by cognitive level

We can also suspect that some subjects lose time in the Online condition because they are distracted by the environment. These subjects would most likely be among the slowest Online subjects, therefore we separate subjects according to the median time in each condition and compare similar groups across conditions. If we focus on the 50% fastest subjects, the average time is 11.49 minutes in the Lab and 10.65 minutes Online with a t-test not rejecting the equality across conditions (p=0.243). If we focus on the 50% slowest subjects, the average time is 25.79 minutes in the Lab and 25.26 minutes Online with a t-test not rejecting the equality across conditions (p=0.711). It suggests that Online subjects are not less concentrated on the task.

B.1.2 Multiple Price List

Figure S2 shows that there are no differences in MPL response time by cognitive levels between conditions.

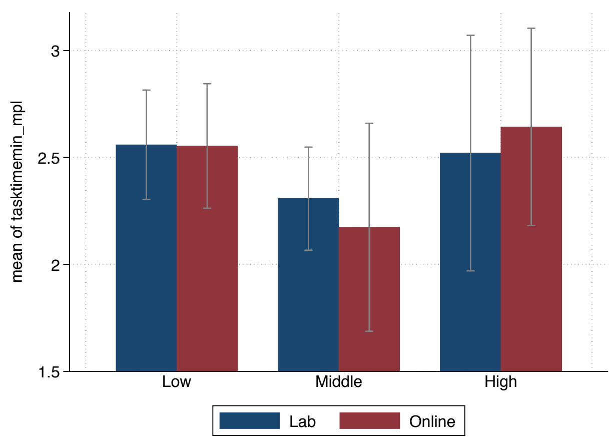


Figure S2: Response time in the MPL by cognitive level

B.1.3 Holt-Laury

Figure S3 shows that there are no differences in terms of response time across by cognitive levels between conditions.

We can also separate subjects in each condition by the time they take to answer the task. Once again, we separate subjects by the median time in each condition (50% fastest, 50% slowest) and compare similar groups across conditions. The fastest subjects take 226.36 seconds in the Lab and 184.90 seconds Online, with a t-test rejecting equality of time between conditions at 1% (p<0.001). The slowest subjects take 418.82 seconds in the Lab and 504.81 seconds Online, with a t-test rejecting equality of time between conditions at 5% (p=0.030). Results therefore suggest that fastest subjects in the Online condition are faster than their Lab counterparts, while slowest subjects in the Online condition are slower than their Lab counterparts.

B.1.4 Dictator Game

Figure S4 shows that there are no differences in terms of response time by cognitive levels between conditions.

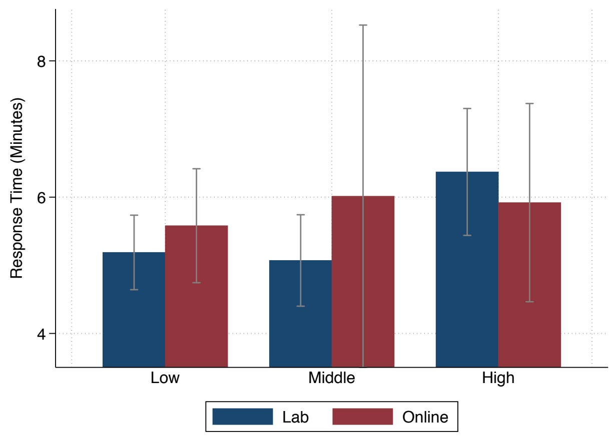


Figure S3: Response Time in the HL task by cognitive level

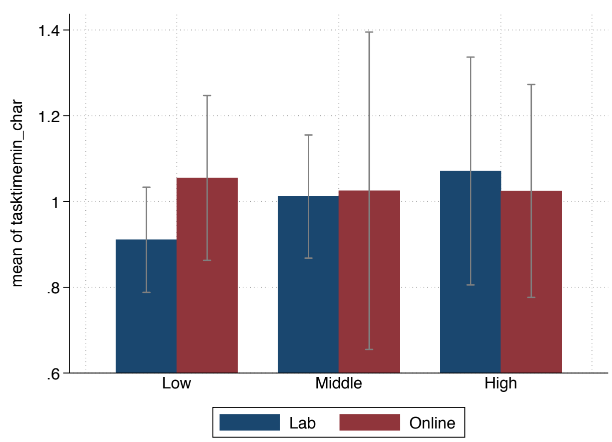


Figure S4: Response Time in the DG by cognitive level

We once again separate subjects by the median time in each condition (50% fastest, 50% slowest) and compare similar groups across conditions. If we focus on the 50% fastest subjects, the average time is 0.62 minutes in the Lab and 0.48 minutes Online, with a t-test rejecting equality between conditions (p<0.001). If we focus on the 50% slowest subjects, the average time is 1.37 minutes in the Lab and 1.56 minutes Online (p=0.122). Therefore, results of the Dictator Game indicate that for subjects in the Online condition, fastest subjects are faster than their Lab counterparts.

B.2 Consistency

B.2.1 Convex Time Budget

Figure S5 shows that there are no differences in terms of consistency over different cognitive levels between Lab and Online conditions.

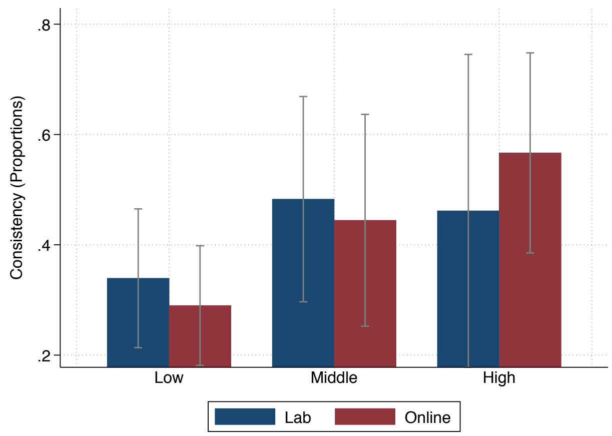


Figure S5: Consistency in the CTB by cognitive level

We can also investigate whether how subjects make inconsistent choices is different across conditions. First, we look at whether the number of inconsistencies in the early (first half of trials[13]) and late (second half of trials[14]) parts of the experiment differ across conditions. The proportion of inconsistencies in the early part of the experiment is 28.68% in the Lab and 27.13% Online with a t-test not rejecting equality across conditions (p=0.572). The proportion of inconsistencies in the late part of the experiment is 27.3% in the Lab and 27.6% Online with a t-test not rejecting equality across conditions (p=0.907). When comparing the percentage of inconsistencies between early and late parts of the experiment within conditions, t-tests do not reject equality for the Lab (p=0.621) and Online (p=0.846). We conclude that the likeliness of subjects making mistakes is not influenced by advancement in the task.

We also look at whether subjects make inconsistencies in similar trials across conditions. We classify trials in three types: Standard (0.XX in the present and 0.20 in the future), 20/25 (0.20 in the present and 0.25 in the future) and 20/20 (0.20 in the present and 0.20 in the future). The proportion of inconsistencies is 28.70% in the Lab and 28.24% Online for Standard trials with a t-test not rejecting equality across conditions (p=0.869), 28.77% in the Lab and 28.70% Online for 20/25 trials with a t-test not rejecting equality across conditions (p=0.976),

[13]until (d=7, k=70) and (a_d=0.18, a_{d+k}=0.20), rationality trial not taken into account.
[14]from (d=7, k=70) and (a_d=0.16, a_{d+k}=0.20)

and 20.76% in the Lab and 25.16% Online for 20/20 trials with a t-test not rejecting equality across conditions (p=0.257). We conclude that subjects are similarly inconsistent in different types of trials across conditions.

Because we decided to consider that subjects are consistent when they make at least 80% consistent choices, it is interesting to study whether the online environment has an impact on consistency when using alternative definitions of consistency. We therefore perform OLS estimations of the effect of the online environment on alternative definitions of CTB consistency. We consider that subjects are consistent if they make at least 70%, 75%, 85%, 90%, 95% or 100% consistent choices. We do not find any significant *online* coefficient, suggesting that the online environment has no impact on CTB consistency.

Table S1: OLS estimations of the impact of online setting on CTB consistency

	(1) Cons70	(2) Cons70	(3) Cons75	(4) Cons75	(5) Cons85	(6) Cons85	(7) Cons90	(8) Cons90	(9) Cons95	(10) Cons95	(11) Cons100	(12) Cons100
online	0.001	0.012	-0.023	-0.018	0.071	0.080	0.058	0.075	0.033	0.039	0.001	0.010
	(0.062)	(0.066)	(0.063)	(0.067)	(0.058)	(0.064)	(0.054)	(0.059)	(0.049)	(0.053)	(0.042)	(0.045)
	[0.985]	[0.862]	[0.711]	[0.791]	[0.225]	[0.212]	[0.277]	[0.204]	[0.496]	[0.460]	[0.979]	[0.824]
Constant	0.575	0.081	0.496	0.001	0.283	-0.150	0.212	-0.107	0.168	-0.084	0.124	0.028
	(0.047)	(0.227)	(0.047)	(0.233)	(0.042)	(0.199)	(0.038)	(0.186)	(0.035)	(0.168)	(0.031)	(0.147)
	[0.000]	[0.722]	[0.000]	[0.996]	[0.000]	[0.451]	[0.000]	[0.566]	[0.000]	[0.619]	[0.000]	[0.847]
Observations	257	219	257	219	257	219	257	219	257	219	257	219
Adj R-squared	-0.004	0.040	-0.003	0.037	0.002	0.038	0.001	0.007	-0.002	0.012	-0.004	-0.016
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

Note: Robust standard errors in parentheses and p values in brackets. Adjusted R-Squared are taken from regressions without robust standard errors. Asterisks denote significance level: $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

B.2.2 Multiple Price List

As before, Figure S6 shows that there is no difference in consistency between Lab and Online conditions across middle and high cognitive levels. We see that for low cognitive level subjects, Lab subjects are less consistent but the difference is marginally significant ($p = 0.087$).

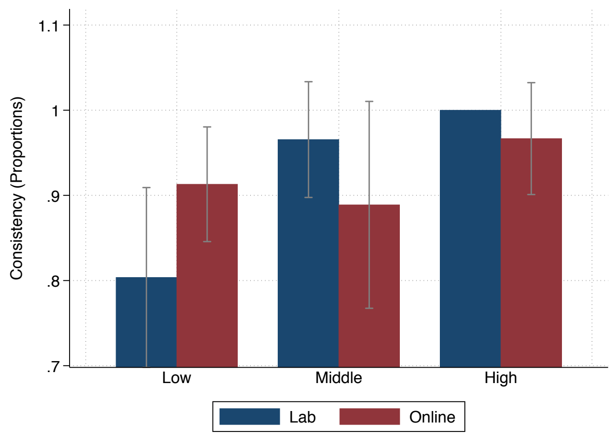


Figure S6: Consistency in the MPL by cognitive level

We also find no differences when comparing the different MPL tasks. We have 92.03% consistent subjects in the Lab and 94.44% consistent subjects Online in the first MPL task, with the difference not being significant (p=0.452). We have 94.69% consistent subjects in the Lab and 95.83% consistent subjects Online in the second MPL task, with the difference not being significant (p=0.672). And we have 90.27% consistent subjects in the Lab and 93.06% consistent subjects Online in the third MPL task, with the difference not being significant (p=0.428).

B.2.3 Holt-Laury

Figure S7 shows that there are no differences in terms of consistency in the HL task by cognitive levels between conditions.

We also find no differences when comparing the different HL tasks. We have 61.06% consistent subjects in the Lab and 61.81% consistent subjects Online in the first HL task, with the difference not being significant (p=0.904). And we have 55.75% consistent subjects in the Lab and 62.5% consistent subjects Online in the second HL task, with the difference not being significant (p=0.277).

We have a relatively high number of subjects that are very inconsistent in at

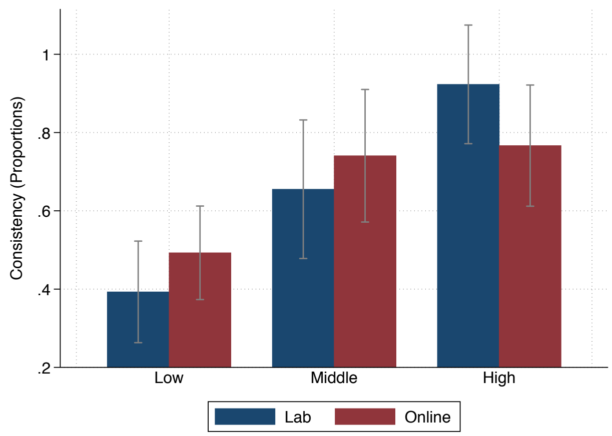


Figure S7: Consistency in the HL task by cognitive level

least one HL task. Subjects are very inconsistent in one HL task if they make at least two inconsistencies in this task. We have 35.40% in the Lab and 33.33% Online subjects that are very inconsistent in the HL task, with a t-test not rejecting equality between conditions ($p = 731$) suggesting that our subjects have difficulties replying to the HL task.

B.3 Preferences

B.3.1 Convex Time Budget

We first look at the number of tokens allocated to early period. Figure S8 shows that there are no differences in terms of number of tokens allocated to the early period by cognitive levels between conditions.

Alternatively, we can study the number of present, interior and future allocations in CTB. Subjects use present allocations on average 2.70 times in the Lab and 3.20 times Online with a t-test not rejecting equality across conditions (p=0.544), subjects use interior allocations on average 33.94 times in the Lab and 32.03 times Online with a t-test not rejecting equality across conditions (p=0.363) and subjects use future allocations on average 8.36 times in the Lab and 9.76 times Online with a t-test not rejecting equality across conditions (p=0.453). We find similar results if we decompose by early payment date (t=0,7,35) or by early and late trials (first half, second half) and run t-tests across conditions. We only find two significant t-tests when comparing t=0 and t=7 in the Online condition (p=0.044) and for both conditions (p=0.020), but both effect sizes are small. We conclude that there are no differences in how subjects allocate their tokens across conditions.

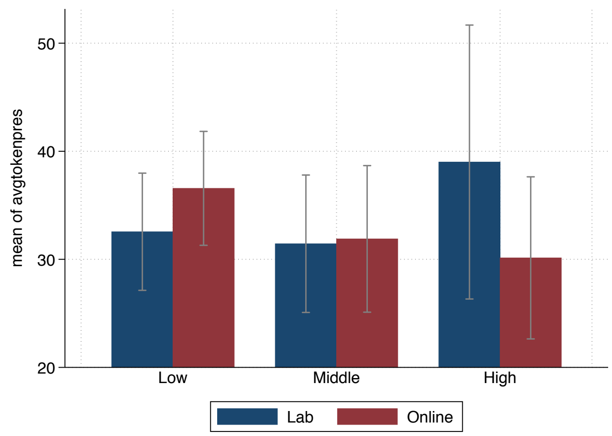


Figure S8: Average amount of tokens allocated to early period in the CTB by cognitive level

We can also study how subjects increase their allocations to the future when advancing to the next trial. More precisely, we measure progression in the future as the difference in allocations to the future between one trial and the previous trial for trials with similar (t,k). The rationality test and 20/25 trials are not included in this comparison. When comparing progressions to the future between the Lab and Online conditions for identical trials with different early dates, we have three t-tests reaching significance: the first progression to the future for t=7 and k=70 (p=0.032), t=35 and k=70 (p=0.070), t=7 and k=98 (p=0.071) with Lab subjects systematically making larger progressions to the future than Online subjects. No other t-test reaches significance.

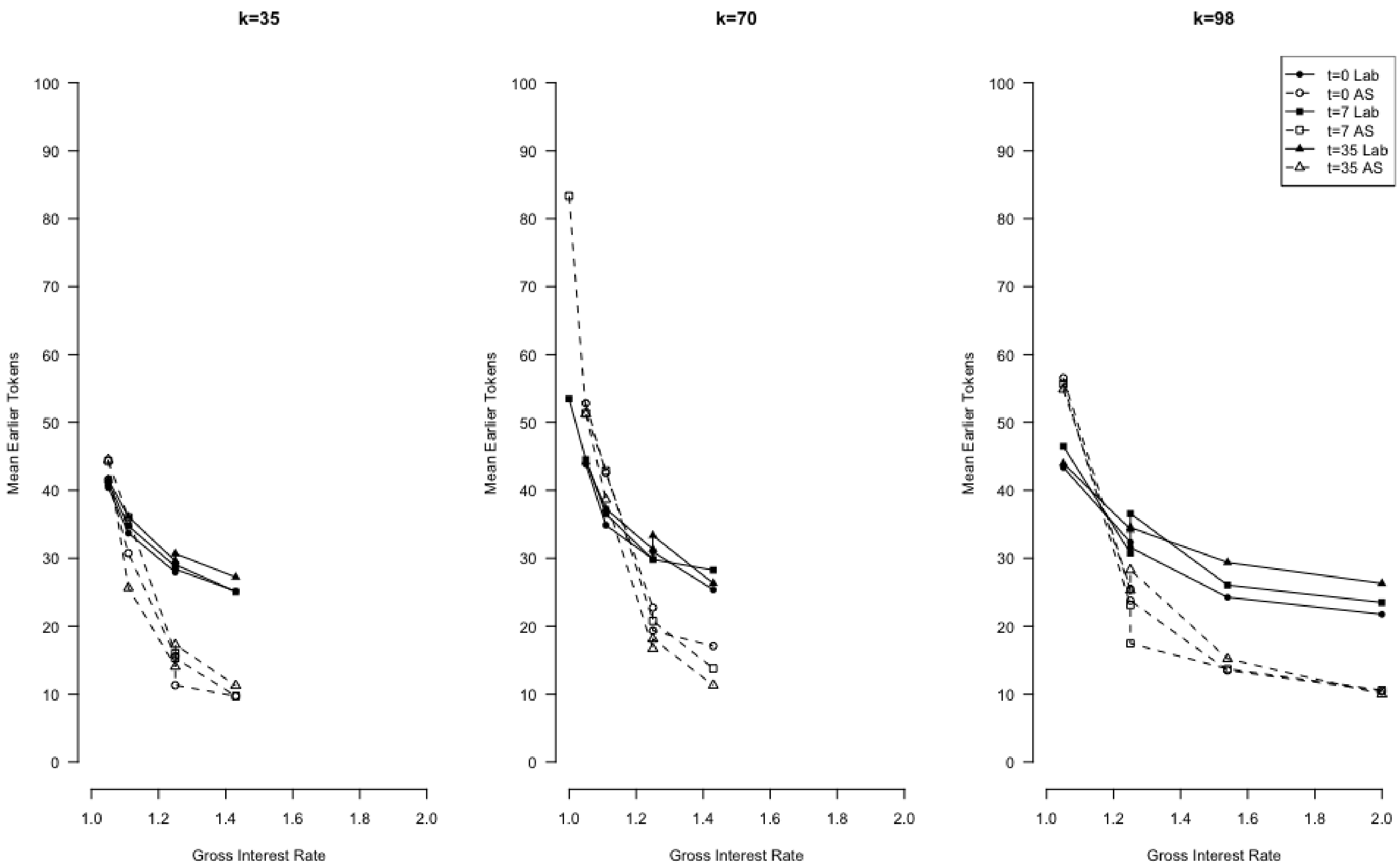


Figure S9: Mean allocations of tokens to early period for Lab subjects and AS subjects

B.3.2 Multiple Price List

Figure S10 shows that there are no differences in terms of number of early period choices in MPL by cognitive levels between conditions.

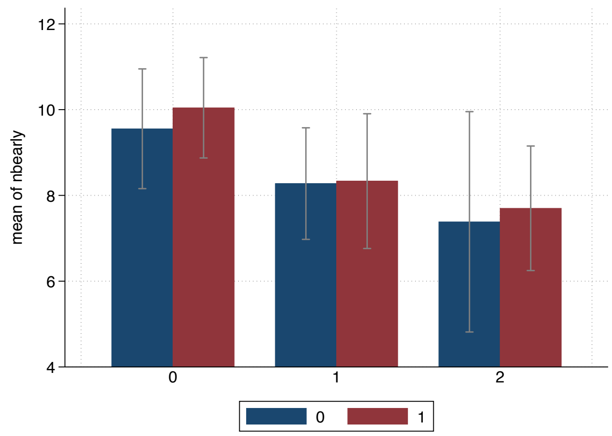


Figure S10: Number of early period choices in the MPL by cognitive level

We can also study whether the online environment affect choices of subjects in the different MPL tasks. Table S2 displays regressions on the number of early choices made by subjects in each MPL task. Column 1 shows that Lab subjects make 2.61 early choices and Online subjects 2.60 early choices in the first MPL task, with the *online* coefficient not being significant ($p = 0.948$). Column 2 shows that this result is robust to adding controls. Column 3 shows that Lab subjects make 3.956 early choices and Online subjects 3.979 early choices in the second MPL task, with the online coefficient not being significant ($p = 0.924$). Column 4 shows that this result is robust to adding controls. Column 5 shows that Lab subjects make 2.646 early choices and Online subjects 2.618 early choices in the third MPL task, with the online coefficient not being significant ($p = 0.907$). Column 6 shows that this result is robust to adding controls. Additionally, Figure S11 displays the choices of consistent subjects in each MPL task to show that they display the characteristic S-shaped pattern. In conclusion, results suggest that subjects reply similarly to the different MPL tasks between conditions.

Table S2: OLS estimations of the impact of online setting on number of early choices in each MPL task

	(1) ErlMPL1	(2) ErlMPL1	(3) ErlMPL2	(4) ErlMPL2	(5) ErlMPL3	(6) ErlMPL3
Online	-0.013	0.180	0.023	0.219	-0.028	0.114
	(0.205)	(0.221)	(0.244)	(0.269)	(0.239)	(0.247)
	[0.948]	[0.416]	[0.924]	[0.416]	[0.907]	[0.644]
Constant	2.611***	1.383	3.956***	2.654***	2.646***	1.615*
	(0.167)	(0.842)	(0.181)	(0.976)	(0.188)	(0.932)
	[0.000]	[0.102]	[0.000]	[0.007]	[0.000]	[0.084]
Observations	257	219	257	219	257	219
Adj R-squared	-0.004	0.020	-0.004	0.032	-0.004	0.085
Controls	No	Yes	No	Yes	No	Yes

Robust standard errors in parentheses and p values in brackets. Adjusted R-Squared from regressions without robust standard errors. Asterisks denote significance level: $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

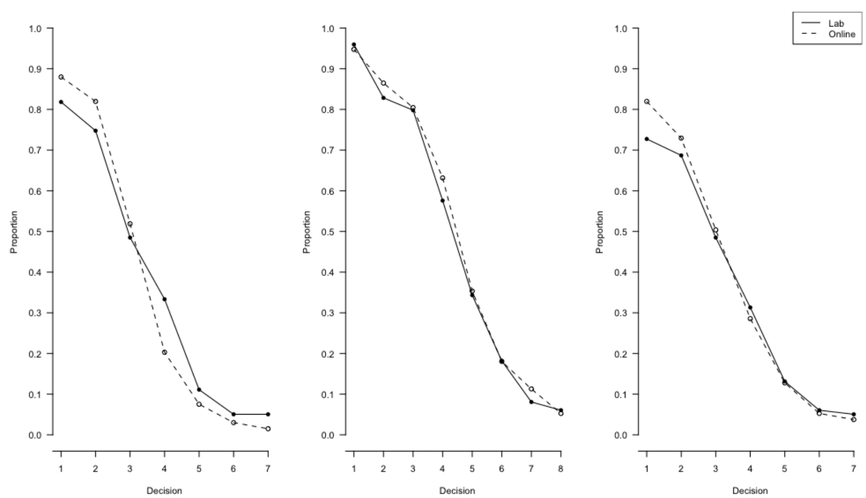


Figure S11: Number of early choices in each MPL task for consistent subjects

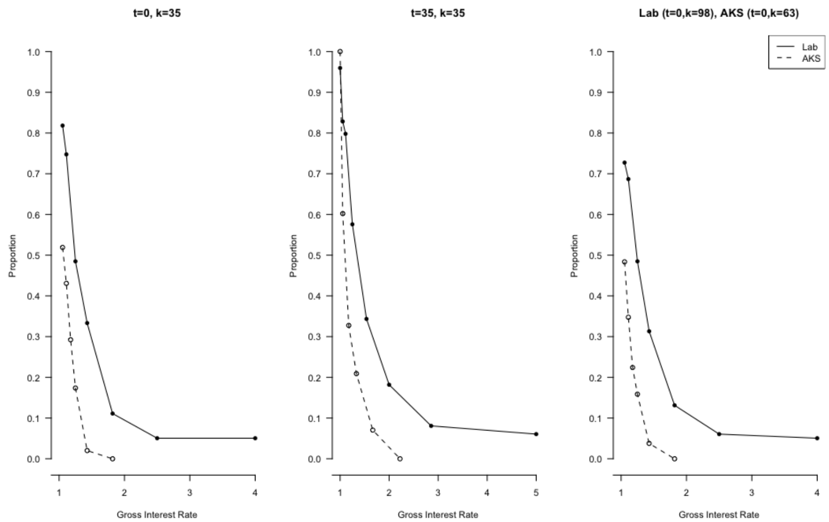


Figure S12: Number of early choices in each MPL task for consistent Lab subjects and AKS subjects

B.3.3 Holt Laury

Figure S13 shows that there are no differences in terms of number of safe choices by cognitive levels between conditions.

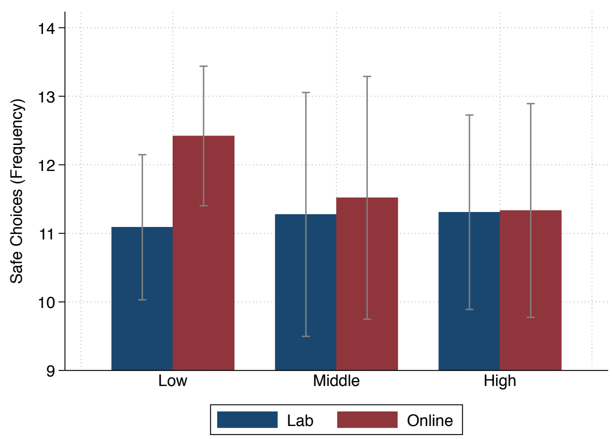


Figure S13: Number of safe choices in the HL task by cognitive level

We can also compare the number of safe choices in the different HL tasks between conditions. Table S3 displays regressions on the number of safe choices made by subjects in each HL task. Column 1 shows that Lab subjects make 5.575 safe choices and Online subjects 6.000 safe choices in the first HL task, with the *online* coefficient not being significant ($p = 0.133$). Column 2 shows that this result is robust to adding controls. Column 3 shows that Lab subjects make 5.478 safe choices and Online subjects 6.007 safe choices in the second MPL task, with the *online* coefficient being marginally significant ($p = 0.053$). Column 4 shows that this result disappear when adding controls. Additionally, Figure S14 displays the number of safe choices in each HL task for consistent subjects to show that subjects correctly answering the task display the characteristic S-shaped pattern. In conclusion, results suggest that subjects reply similarly to the different HL tasks between conditions.

Table S3: OLS estimations of the impact of online setting on number of early choices in each HL task

	(1) SafeHL1	(2) SafeHL1	(3) SafeHL2	(4) SafeHL2
Online	0.425	0.347	0.529*	0.390
	(0.281)	(0.269)	(0.273)	(0.301)
	[0.133]	[0.269]	[0.053]	[0.197]
Constant	5.575***	7.751***	5.478***	6.490***
	(0.209)	(1.054)	(0.200)	(0.986)
	[0.000]	[0.000]	[0.000]	[0.000]
Observations	257	219	257	219
Adj R-squared	0.005	0.027	0.011	0.013
Controls	No	Yes	No	Yes

Robust standard errors in parentheses and p values in brackets. Adjusted R-Squared from regressions without robust standard errors. Asterisks denote significance level: $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

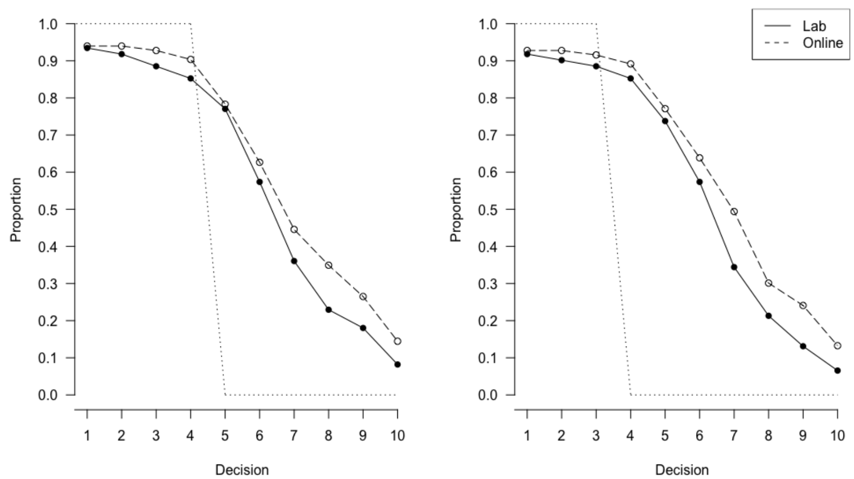


Figure S14: Number of safe choices in each HL task for consistent subjects

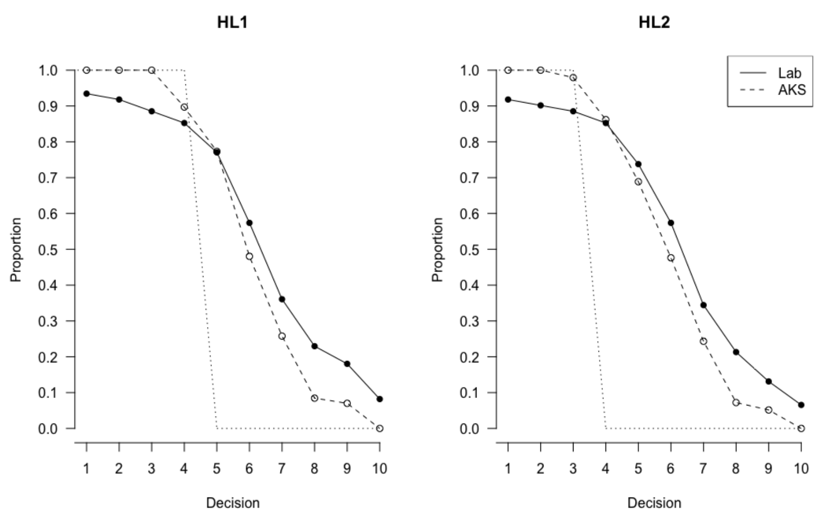


Figure S15: Number of safe choices in each HL task for consistent Lab subjects and AKS subjects

B.3.4 Dictator Game

Figure S16 shows that there are no differences in decision to donate to charity by cognitive levels between conditions.

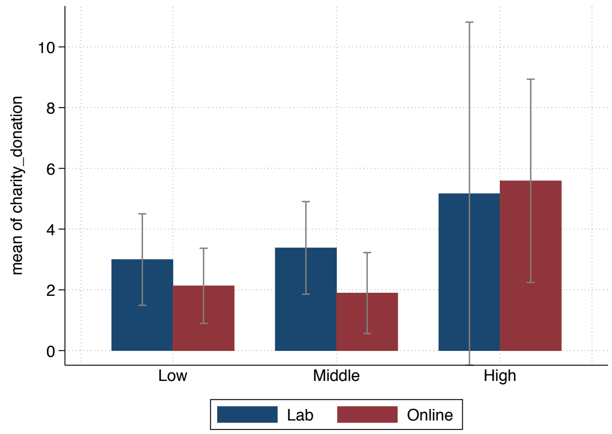


Figure S16: Decision to donate to charities in the DG by cognitive level

Figure S17 shows that there are no differences in terms of percentage of earnings donated to charities between conditions for Low and High cognitive abilities subjects. Regarding subjects with Middle cognitive abilities, Lab subjects donate 6.36% more to charities than Online subjects with a marginally significant t-test rejecting equality between conditions at 10% ($p = 0.083$).

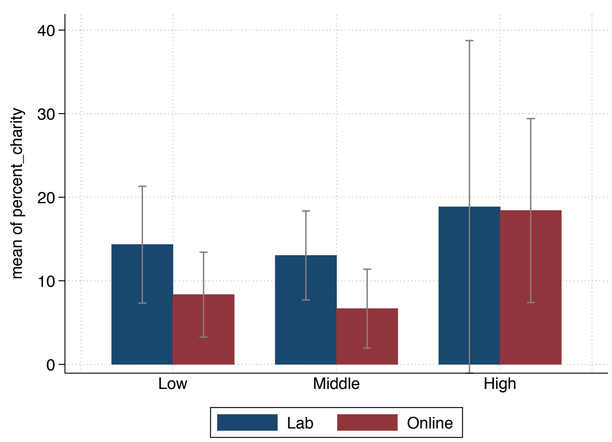


Figure S17: Percentage of earnings donated to charities in the DG by cognitive level

Table S4 displays regressions measuring the effect of the online environment on the likeliness to donate 50% ($share = 50$), 100% ($share = 100$), less than 50% ($share < 50$) and more than 50% ($share > 50$) of your experimental payments to charities in the DG. We find no significant coefficient for the *online* variable, further suggesting that differences between conditions are driven by subjects donating nothing in the Online environment.

Table S4: OLS estimations of the impact of online setting on share and type of altruistic behavior in the DG

	(1) Share=50	(2) Share=50	(3) Share=100	(4) Share=100	(5) Share<50	(6) Share<50	(7) Share>50	(8) Share>50
online	-0.025	-0.032	0.009	-0.015	0.025	0.057	0.001	-0.025
	(0.025)	(0.029)	(0.029)	(0.029)	(0.039)	(0.041)	(0.030)	(0.030)
	[0.317]	[0.275]	[0.749]	[0.603]	[0.521]	[0.165]	[0.986]	[0.416]
Constant	0.053**	-0.013	0.053**	-0.176	0.885***	1.193***	0.062***	-0.179
	(0.021)	(0.088)	(0.021)	(0.141)	(0.039)	(0.154)	(0.023)	(0.139)
	[0.013]	[0.879]	[0.013]	[0.215]	[0.000]	[0.000]	[0.007]	[0.199]
Observations	257	219	257	219	257	219	257	219
R-squared	0.000	-0.011	-0.003	0.032	-0.002	0.038	-0.004	0.044
Controls	No	Yes	No	Yes	No	Yes	No	Yes

Robust standard errors in parentheses and p values in brackets. Adjusted R-Squared from regressions without robust standard errors. Asterisks denote significance level: $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

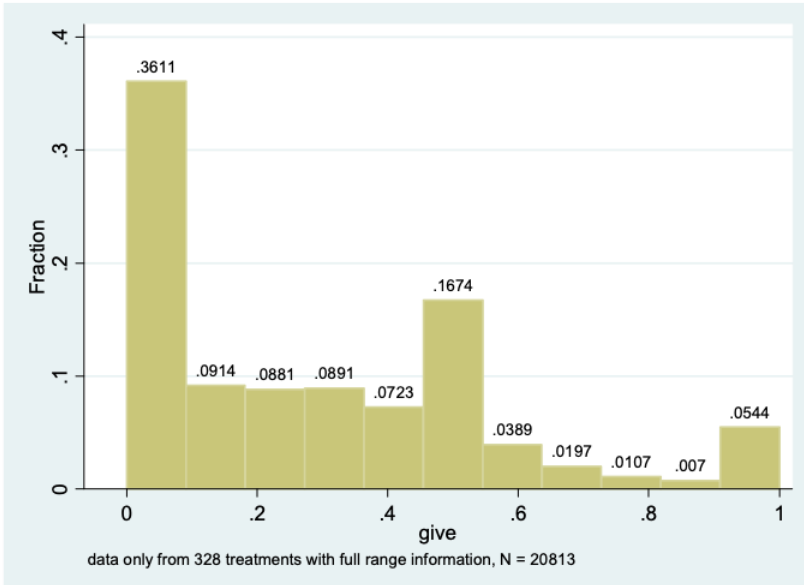


Figure S18: Distribution of share of earnings donated in the DG from the meta-analysis of Engels (2011)

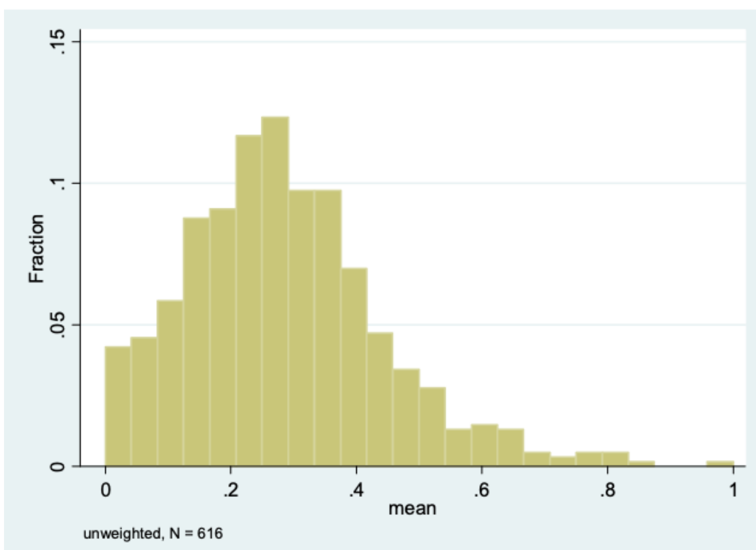


Figure S19: Distribution of the average share of earnings donated in DG experiments from the meta-analysis of Engels (2011)

We further investigate whether the presence of others influence decision to donate to charities in the DG. Column 1 to 4 in Table S5 shows that presence of others has no effect on the share of earnings or amount of money donated to charities. However, Column 5 shows that the presence of others reduce by 20% the likeliness to not donate anything, with this effect being significant at 5% ($p = 0.019$). Column 6 shows that adding controls make this result significant at 1% ($p = 0.002$).

Table S5: OLS estimations of the impact of presence of others on donation to charities

	(1) Share	(2) Share	(3) Giving	(4) Giving	(5) Selfish	(6) Selfish
withothers	2.773	5.119	0.239	0.792	-0.200**	-0.271***
	(4.154)	(4.678)	(0.957)	(1.106)	(0.085)	(0.086)
	[0.505]	[0.275]	[0.803]	[0.474]	[0.019]	[0.002]
Constant	12.409***	-15.925	3.204***	-5.641	0.473***	0.786***
	(1.695)	(13.421)	(0.446)	(3.742)	(0.034)	(0.228)
	[0.000]	[0.237]	[0.000]	[0.133]	[0.000]	[0.001]
Observations	253	217	253	217	253	217
Adj R-squared	-0.003	0.053	-0.004	0.071	0.014	0.047
Controls	No	Yes	No	Yes	No	Yes

Robust standard errors in parentheses and p values in brackets. Adjusted R-Squared from regressions without robust standard errors. Asterisks denote significance level: $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

B.3.5 CRT and Numeracy test scores

Table S6 displays regressions measuring the effect of the online environment on CRT and Numeracy test scores. We use the same control variables than in previous regressions, except *crt score* and *numeracy score* when they are the dependent variable. Column 1 shows that on average Lab subjects answer 0.571 questions correctly and Online subjects answer 0.809 questions correctly in the CRT, with this difference being significant at 5% ($p = 0.050$). Column 2 shows that this difference remains marginally significant when adding controls ($p = 0.074$). Column 3 shows that on average Lab subjects answer 4.398 questions correctly and Online subjects answer 4.604 questions correctly in the Numeracy test, with this difference not being significant ($p = 0.183$). Column 4 shows that this result holds when adding controls ($p = 0.452$). Results suggest that the online environment influence results in the CRT but not the Numeracy test. A potential explanation of the increased CRT score in the online environment is that subjects can easily find questions and their answers on a search engine. Additionally, the meta-analysis of Brañas-Garza, Kujal and

Lenkei (2019) found that subjects answer 1.198 questions correctly in the CRT. We test the equality of the number of correct answers in CRT by conditions with this number, and find that t-tests reject equality for both Lab subjects ($p < 0.001$) and Online subjects ($p < 0.001$). It suggests that our subjects did not concentrate when answering the CRT, potentially because they answered it inside the questionnaire after a long experiment.

Table S6: OLS estimations of the impact of online setting on CRT and Numeracy test scores

	(1) CRTScr	(2) CRTScr	(3) NumScr	(4) NumScr
Online	0.238**	0.206*	0.206	0.106
	(0.121)	(0.115)	(0.154)	(0.140)
	[0.050]	[0.074]	[0.183]	[0.452]
Constant	0.571***	-0.606	4.398***	4.359***
	(0.075)	(0.392)	(0.118)	(0.360)
	[0.000]	[0.124]	[0.000]	[0.000]
Observations	224	219	257	219
Adj R-squared	0.012	0.129	0.003	0.147
Controls	No	Yes	No	Yes

Robust standard errors in parentheses and p values in brackets. Adjusted R-Squared from regressions without robust standard errors. Asterisks denote significance level: $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

C Recruitment Material

The following section displays the material used to recruit subjects.

C.1 Recruitment Letter

Because this experiment was the first ran in Middlesex University, it was necessary to build the pool of subjects. We sent the following invitation to participate to the 22544 students in the six schools of the university[15] :

Subject: Earn **$$$$** participating in economics experiments.

"Dear student:

We would like to invite you to participate in an experiment on decision-making in the first week of December. The experiment is organized by researchers in

[15] Art and Design, Business, Health and Education, Law, Media and Performing Art, Science and Technology

the Economics Department here at Middlesex University. By participating you will be able to earn real money. The payment mentioned below will only be made to those selected to participate.

The amount of money that you can earn will depend on the structure of the task, your own decisions, those made by others and chance.

For this particular experiment:

- You will be paid a show-up fee.
- There is a minimum payment of £12 including the show-up fees.
- The experiment will last no longer than 90 minutes.

To register your interest in participating in this experiment please send an email to experiments@mdx.ac.uk with your name and preferred email address for further instructions. We will then contact you to inform how to proceed.

Yours sincerely,

Professor Pablo Branas-Garza and Dr Karen Khachatryan Economics Department Middlesex University Business School"

C.2 Follow-Up Recruitment Letter

Students who replied and showed their interest to participate in our experiment were then sent this follow-up recruitment letter :

"Dear student:

Thank you for your interest in participating in our experiment.

To be eligible for this experiment, you need to meet these criteria:

> 1. To participate in the experiment you must agree to provide certain personal information, such as your name and telephone number, which will be treated in full confidence. To this end you must read and sign a more general informed consent form available online at http://www.gstk.eu/raer
>
> 2. To receive payments for this experiment you must have a UK bank account and a UK mobile phone number. However, we will not ask you to provide us with your bank account details.

If you agree to participate, please sign the informed consent form online now.

If you agree to participate but do not meet the second criteria, please inform us of this now by replying to this email.

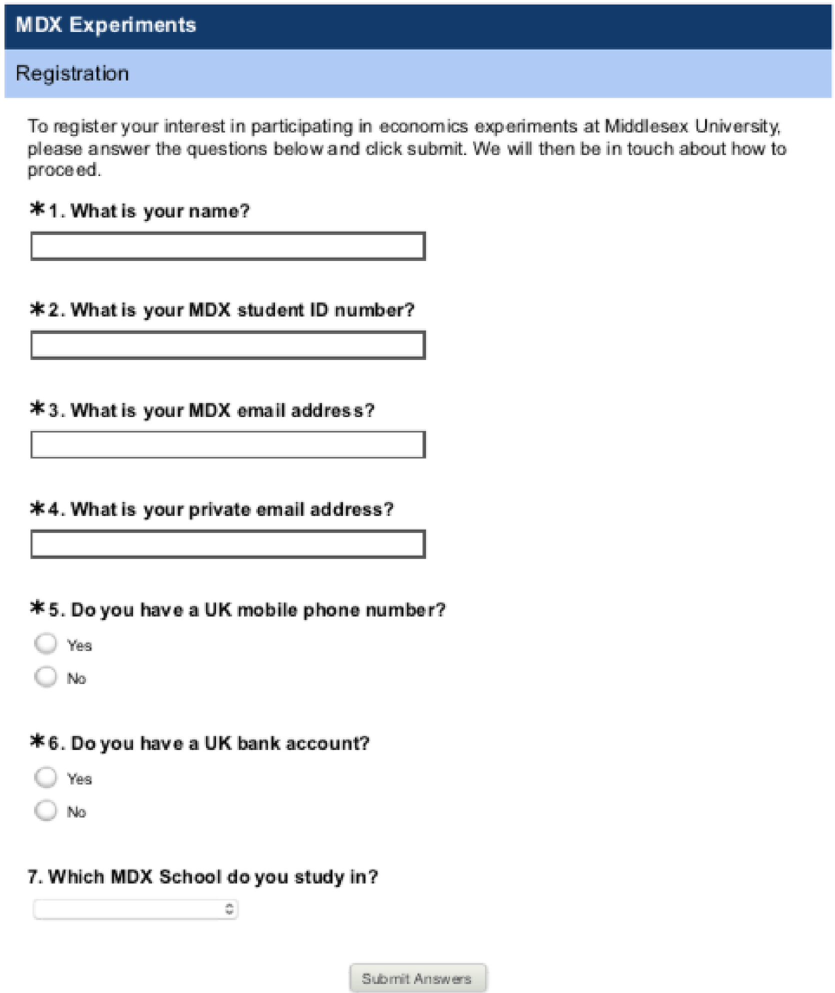
Once you have signed the informed consent form, we will be in touch with further information about possible dates, times and location of the experiment.

Yours sincerely,

Professor Pablo Branas-Garza and Dr Karen Khachatryan Economics Department Middlesex University Business School"

C.3 Registration Survey and Consent Form

Subjects replied the survey in Figure S20 and signed the consent form in Figure S21 to indicate their participation in the experiment.



MDX Experiments

Registration

To register your interest in participating in economics experiments at Middlesex University, please answer the questions below and click submit. We will then be in touch about how to proceed.

***1. What is your name?**

***2. What is your MDX student ID number?**

***3. What is your MDX email address?**

***4. What is your private email address?**

***5. Do you have a UK mobile phone number?**

Yes

No

***6. Do you have a UK bank account?**

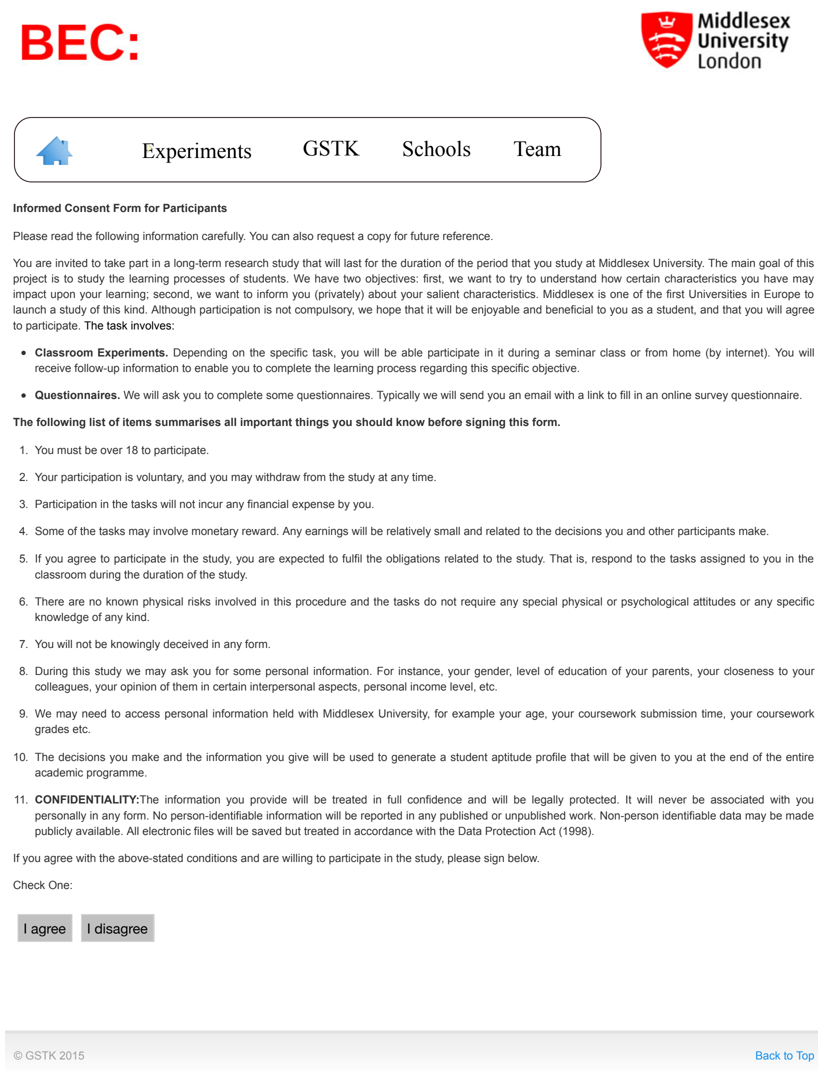
Yes

No

7. Which MDX School do you study in?

Submit Answers

Figure S20: Screenshot of the registration survey



BEC:

Middlesex University London

Experiments GSTK Schools Team

Informed Consent Form for Participants

Please read the following information carefully. You can also request a copy for future reference.

You are invited to take part in a long-term research study that will last for the duration of the period that you study at Middlesex University. The main goal of this project is to study the learning processes of students. We have two objectives: first, we want to try to understand how certain characteristics you have may impact upon your learning; second, we want to inform you (privately) about your salient characteristics. Middlesex is one of the first Universities in Europe to launch a study of this kind. Although participation is not compulsory, we hope that it will be enjoyable and beneficial to you as a student, and that you will agree to participate. The task involves:

- **Classroom Experiments.** Depending on the specific task, you will be able participate in it during a seminar class or from home (by internet). You will receive follow-up information to enable you to complete the learning process regarding this specific objective.
- **Questionnaires.** We will ask you to complete some questionnaires. Typically we will send you an email with a link to fill in an online survey questionnaire.

The following list of items summarises all important things you should know before signing this form.

1. You must be over 18 to participate.
2. Your participation is voluntary, and you may withdraw from the study at any time.
3. Participation in the tasks will not incur any financial expense by you.
4. Some of the tasks may involve monetary reward. Any earnings will be relatively small and related to the decisions you and other participants make.
5. If you agree to participate in the study, you are expected to fulfil the obligations related to the study. That is, respond to the tasks assigned to you in the classroom during the duration of the study.
6. There are no known physical risks involved in this procedure and the tasks do not require any special physical or psychological attitudes or any specific knowledge of any kind.
7. You will not be knowingly deceived in any form.
8. During this study we may ask you for some personal information. For instance, your gender, level of education of your parents, your closeness to your colleagues, your opinion of them in certain interpersonal aspects, personal income level, etc.
9. We may need to access personal information held with Middlesex University, for example your age, your coursework submission time, your coursework grades etc.
10. The decisions you make and the information you give will be used to generate a student aptitude profile that will be given to you at the end of the entire academic programme.
11. **CONFIDENTIALITY:**The information you provide will be treated in full confidence and will be legally protected. It will never be associated with you personally in any form. No person-identifiable information will be reported in any published or unpublished work. Non-person identifiable data may be made publicly available. All electronic files will be saved but treated in accordance with the Data Protection Act (1998).

If you agree with the above-stated conditions and are willing to participate in the study, please sign below.

Check One:

I agree I disagree

© GSTK 2015 Back to Top

Figure S21: Screenshot of the consent form

D Screenshots of the Experiment (Instructions and Tasks)

The following section displays screenshots of the experiment as it was seen by subjects. It includes both instructions and the tasks themselves. Figure S22 show that subjects were first welcomed in the experiment and explained the payment procedure. Then Figure S23 shows that subjects were then explained how the experimental payment will be determined before leaving the first screen.

Welcome and thank you for participating in this experiment on economic decision-making!

You must not press any of your browser buttons (e.g. stop, refresh, reload, back, forward) at any time. If you do so you will exit the experiment.

Payment Explanation

You will earn money by participating in this experiment. Any payments due will be made directly to your bank account. We will only need your mobile phone number to make the bank transfer. To enable us to pay you using just your mobile number we will use Barclays Pingit App.

If you have already registered and linked your phone number to your bank account through either Barclays Pingit App or Paym then you don't have to do anything else besides providing us with your mobile phone number at the end of the experiment.

If you have not yet registered and linked your mobile phone number to your bank account, when we make a payment to you, you will receive a text message with further instructions. You will be instructed to either (1) download the Barclays Pingit app to your mobile phone and register your account details or (2) link your phone number to your bank account online using a computer (web browser). To receive instant payments into your bank account through your mobile phone number, you do not need to be a Barclays customer. Any UK current account will work.

In this experiment, we use $ as an experimental currency unit. However, you will be paid in pound sterling, £. Each $1.00 is worth £0.60. For example, if you earn $10 in the experiment, it means you will be paid 10 times 0.6 = £6.00. As another example, if you earn $25, it means you will be paid £15.00 (25 times 0.6).

Figure S22: Screenshot of payment explanations

Subjects then started the CTB task. Figure S24 shows the CTB instructions presented to subjects and Figure S25 shows the example of CTB decision screen presented to subjects. Subjects were explained that the task consisted in allocating tokens between an earlier and a later date. They were also explained how to read the calendar and decision interface, and were given precision on how to interact with the decision interface :

Earning Money

To begin, you will be given a $10 thank-you payment, just for participating in this experiment! You will receive this thank-you payment in two equally sized payments of $5 each. The two $5 payments will be made to you at two different times. These times will be determined in the way described below.

In this experiment, you will make 47 choices over how to allocate money between two points in time, one time is "earlier" and one is "later." Both the earlier and later times will vary across decisions. This means you could be receiving payments as early as today, and as late as the end of April, or possibly two other dates in between. Once all 47 decisions have been made, the computer will **randomly select one of the 47 decisions as the decision-that-counts**. We will use the decision-that-counts to determine your actual earnings. Note, since all decisions are equally likely to be chosen, you should make each decision as if it will be the decision-that-counts.

When calculating your earnings from the decision-that-counts, we will add to your earnings the two $5 thank-you payments. Thus, you will always get paid at least $5 at the chosen earlier time, and at least $5 at the chosen later time.

IMPORTANT: All payments you receive will arrive to your bank account through your mobile phone number. That includes payments that you receive today as well as payments you may receive at later dates.

As a reminder to you, the day before you are scheduled to receive one of your payments, we will send you a text message notifying you that the payment is coming.

By now you know how to contact us at experiments@mdx.ac.uk. However, if one of your payments is not received you should immediately contact Professor Pablo Branas-Garza at branasgarza@gmail.com or Dr. Karen Khachatryan at karen.g.khachatryan@gmail.com (tel. 07506276716), and we will instantly pay you.

Continue <-- Please click here to continue to the next screen

Figure S23: Screenshot of payment determination

The Experiment

This is an experiment on decision making with 47 choices, presented in three parts. In this first part you will make a series of 45 decisions about how to divide a set of tokens between two dates. Tokens will later be exchanged for money. The tokens you allocate to later date will always be worth more money than tokens you allocate to the earlier date. This process is best described by an example.

Below is a sample Decision Screen, like what you will see in the next decision screen.

Choosing the Decision-that-Counts

The first decision on the screen shows the choice to allocate 100 tokens between January 12th and February 16th. Notice that today's date is highlighted in yellow on the calendar above the tab. Also note that the earlier date (January 12th) is highlighted in green while the later date (February 16th) is highlighted in blue. In each decision the dates are highlighted so that you can easily see when the decision begins and ends.

In this decision, each token you allocate to January 12th is worth $0.16, while each token you allocate to February 16th is worth $0.20. So, if you allocate all 100 tokens to January 12th you will earn $16 on this date, and nothing on February 16th. You are also free to allocate some tokens to the earlier date and some to the later date. For instance, if you allocate 50 tokens to January 12th and 50 to February 16th, you will earn $8.00 on January 12th and $10.00 on February 16th. Remember that however you allocate the tokens, any earnings will be added to your $5 thank-you payment for both the earlier and later dates. So, even if you allocate all your tokens to one of the dates, you will still receive a payment of at least $5 on both the earlier and later dates.

Notice that you can navigate through all 45 decisions by using the tabs at the top of the decision screen. Notice also that, on the right, the computer automatically calculates how much you will receive on both the earlier and later dates, if this is chosen as the decision-that-counts. You can revise your choices as much as you like. Once you are satisfied with all of your decisions, you can click on the "Submit Decisions" button.

Figure S24: Screenshot of CTB instructions

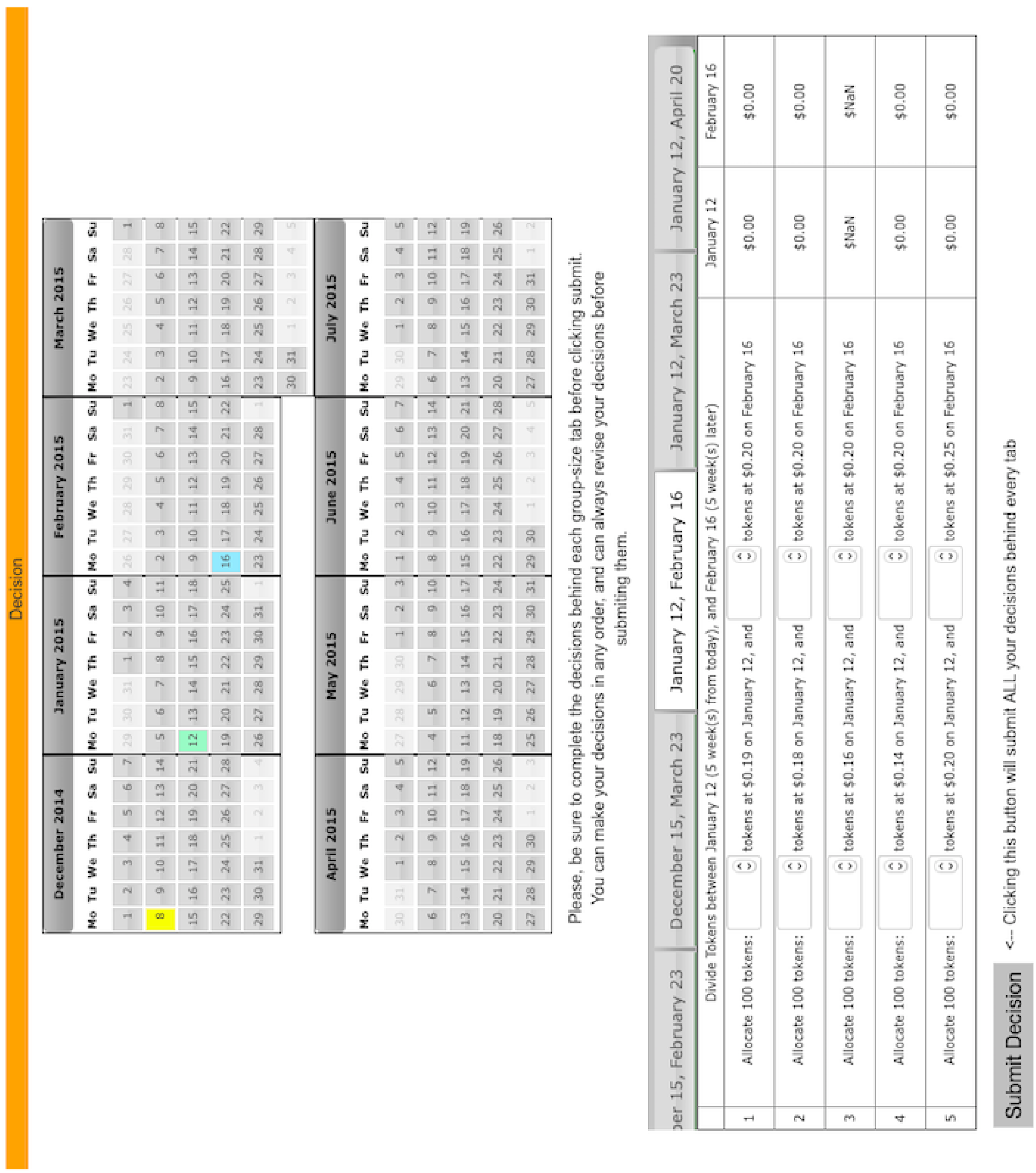


Figure S25: Screenshot of the CTB example

Subjects were then given two examples to better understand how to differently allocate in CTB and how to read the automatically calculated payments. Figure S26 shows the first explanatory example and Figure S27 shows the second explanatory example.

Consider another example. The screen below shows an allocation of 25 tokens for January 12th and 75 tokens for February 16th. Notice how the cash values are automatically computed. Suppose this decision was chosen as the decision-that-counts for payment. Then this subject would be paid $4.00 on January 12th and $15.00 on February 16th. The person's earlier payment of $4.00 + $5.00 (thank-you payment) = $9.00 would be paid on January 12th. The person's later payment of $15.00 + $5.00 (thank-you payment) = $20.00 will be paid on February 16th.

Please, be sure to complete the decisions behind each group-size tab before clicking submit.
You can make your decisions in any order, and can always revise your decisions before submiting them.

mber 15, February 23 | December 15, March 23 | January 12, February 16 | January 12, March 23 | January 12, April 20

	Divide Tokens between January 12 (5 week(s) from today), and February 16 (5 week(s) later)				January 12	February 16
1	Allocate 100 tokens:		tokens at $0.19 on January 12, and	tokens at $0.20 on February 16	$0.00	$0.00
2	Allocate 100 tokens:		tokens at $0.18 on January 12, and	tokens at $0.20 on February 16	$0.00	$0.00
3	Allocate 100 tokens:	25	tokens at $0.16 on January 12, and 75	tokens at $0.20 on February 16	$4.00	$15.00
4	Allocate 100 tokens:		tokens at $0.14 on January 12, and	tokens at $0.20 on February 16	$0.00	$0.00
5	Allocate 100 tokens:		tokens at $0.20 on January 12, and	tokens at $0.25 on February 16	$0.00	$0.00

Submit Decision <-- Clicking this button will submit ALL your decisions behind every tab

Figure S26: Screenshot of the first CTB example

Suppose instead that the following choice was made: 40 tokens for January 12th and 60 tokens for February 16th. Then, if this decision was chosen as the decision-that-counts, this subject's earlier payment of $11.40 (= $6.40 + $5 thank-you payment) would be made on January 12th and the later payment of $17 (= $9 + $5 thank-you payment) would be made on February 16th.

Please, be sure to complete the decisions behind each group-size tab before clicking submit.
You can make your decisions in any order, and can always revise your decisions before submiting them.

mber 15, February 23 | December 15, March 23 | January 12, February 16 | January 12, March 23 | January 12, April 20

	Divide Tokens between January 12 (5 week(s) from today), and February 16 (5 week(s) later)				January 12	February 16
1	Allocate 100 tokens:		tokens at $0.19 on January 12, and	tokens at $0.20 on February 16	$0.00	$0.00
2	Allocate 100 tokens:		tokens at $0.18 on January 12, and	tokens at $0.20 on February 16	$0.00	$0.00
3	Allocate 100 tokens:	40	tokens at $0.16 on January 12, and 60	tokens at $0.20 on February 16	$6.40	$12.00
4	Allocate 100 tokens:		tokens at $0.14 on January 12, and	tokens at $0.20 on February 16	$0.00	$0.00
5	Allocate 100 tokens:		tokens at $0.20 on January 12, and	tokens at $0.25 on February 16	$0.00	$0.00

Submit Decision <-- Clicking this button will submit ALL your decisions behind every tab

Figure S27: Screenshot of the second CTB example

Before answering the CTB task, subjects were provided a summary of the experimental instructions that is displayed in Figure S28.

Important: The number of tokens to allocate is the same across all decisions, i.e. 100.

Important: Tokens allocated to the earlier and later times will vary across the 45 decisions (from $0.10 to $0.20 for the earlier time, and from $0.20 to $0.25 for the later time), but tokens allocated to the later time will always be worth more than the tokens allocated to the earlier time.

After completing the decisions on the first tab, move on to the second tab and so forth. Make sure to complete all the nine tabs and take your time on each decision. You can come back and change your decisions at any time before submitting them.

Each decision is numbered. At the end of the experiment, the computer will draw one random number to determine which decision will actually be the decision-that-counts for payment.

Important: Since each of the 45 decisions is equally likely to be chosen as the decision that counts, you should treat each decision independently, imagining that it will in fact be chosen as the decision-that-counts.

Important things to remember

- You will get $10 in thank-you payments just for participating.
- The thank-you payment will be split $5 at the earlier time and $5 at the later time.
- You will make 45 decisions about how to allocate tokens between an earlier time and a later time.
- The tokens you allocate to the later time will always be worth more than the tokens you allocate to the earlier time.
- One of the 47 decisions you make will be randomly selected to be the decision-that-counts for payment.
- You will get a text message either on the day or the day before your payment is to arrive.
- Your payment, made into your bank account, will arrive through the mobile phone number you provide.
- Barclays Pingit guarantees instant payments.
- You must not press any of your browser buttons (e.g. stop, refresh, reload, back, forward) at any time.

Figure S28: Screenshot of the end of the CTB instructions

Subjects then replied to the CTB task. We show screenshots of the decision screen for each of the nine intertemporal budget choices proposed to subjects :

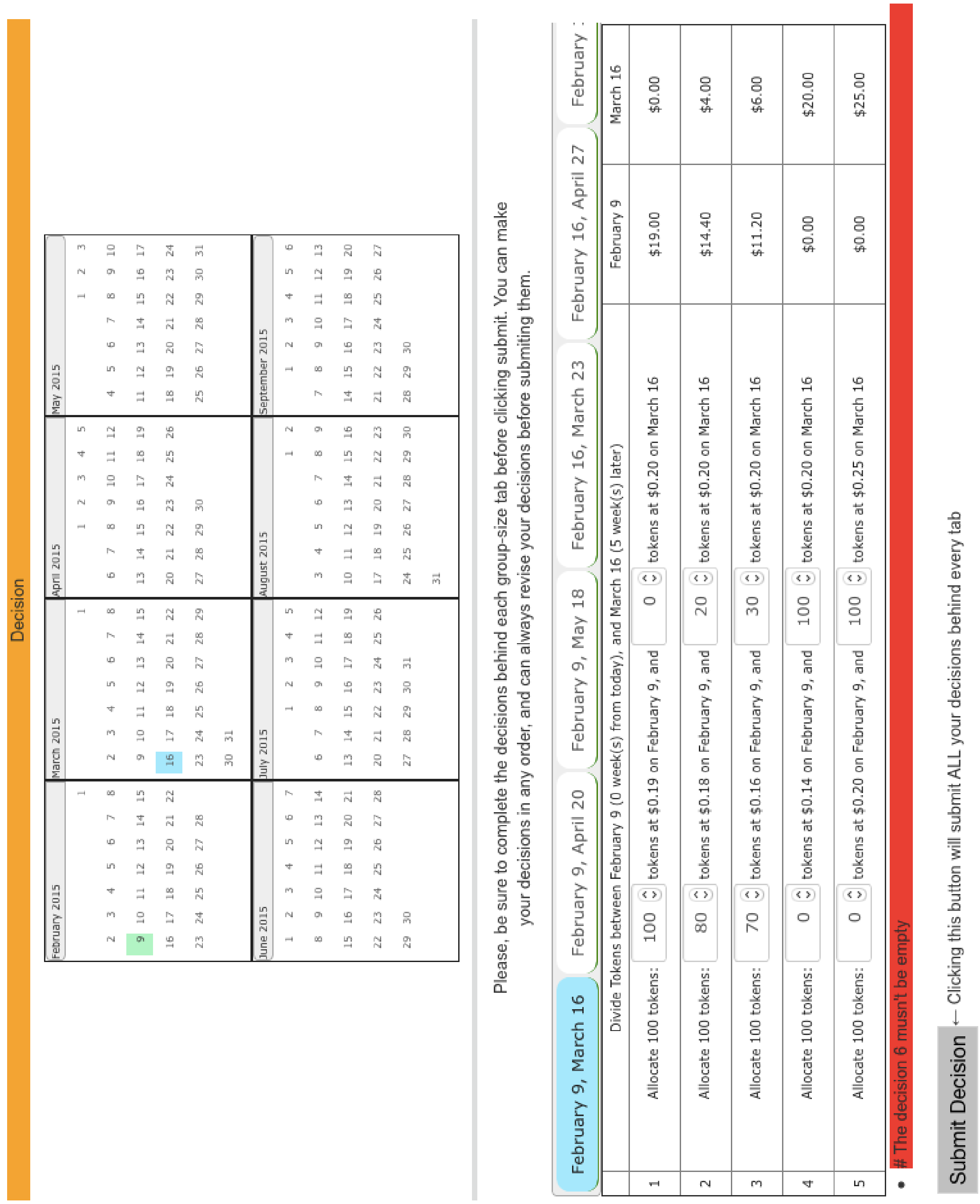


Figure S29: Screenshot of CTB1

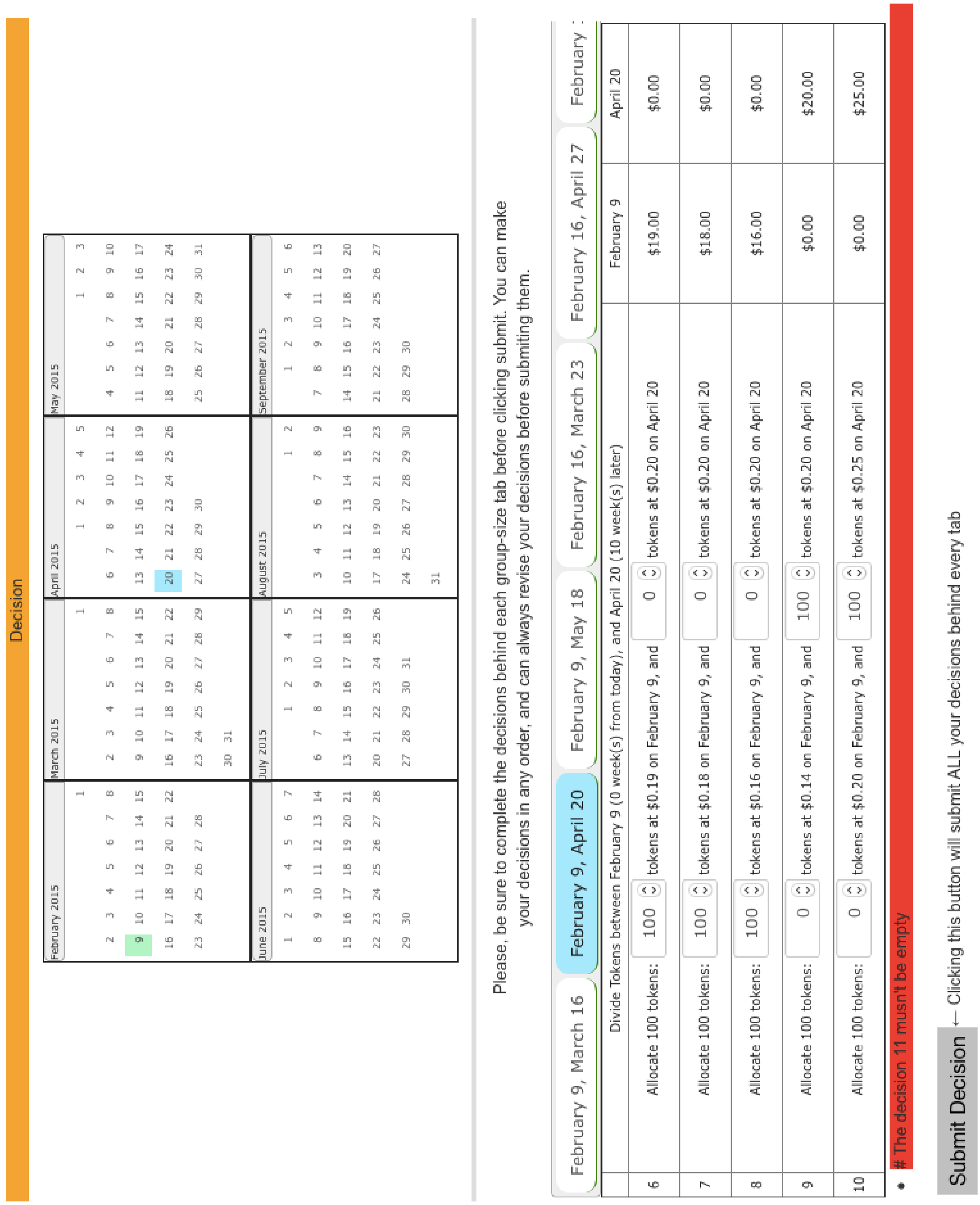


Figure S30: Screenshot of CTB2

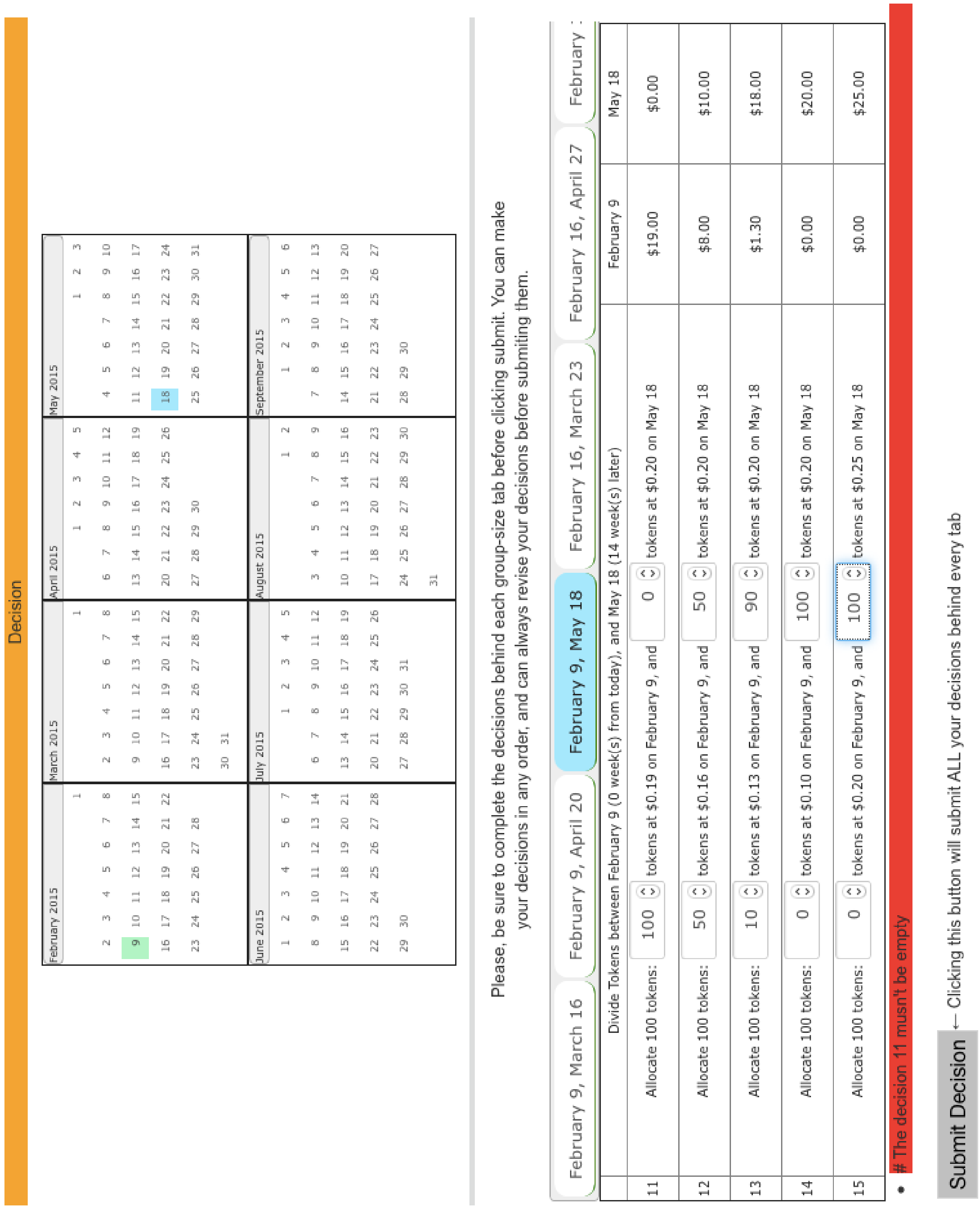


Figure S31: Screenshot of CTB3

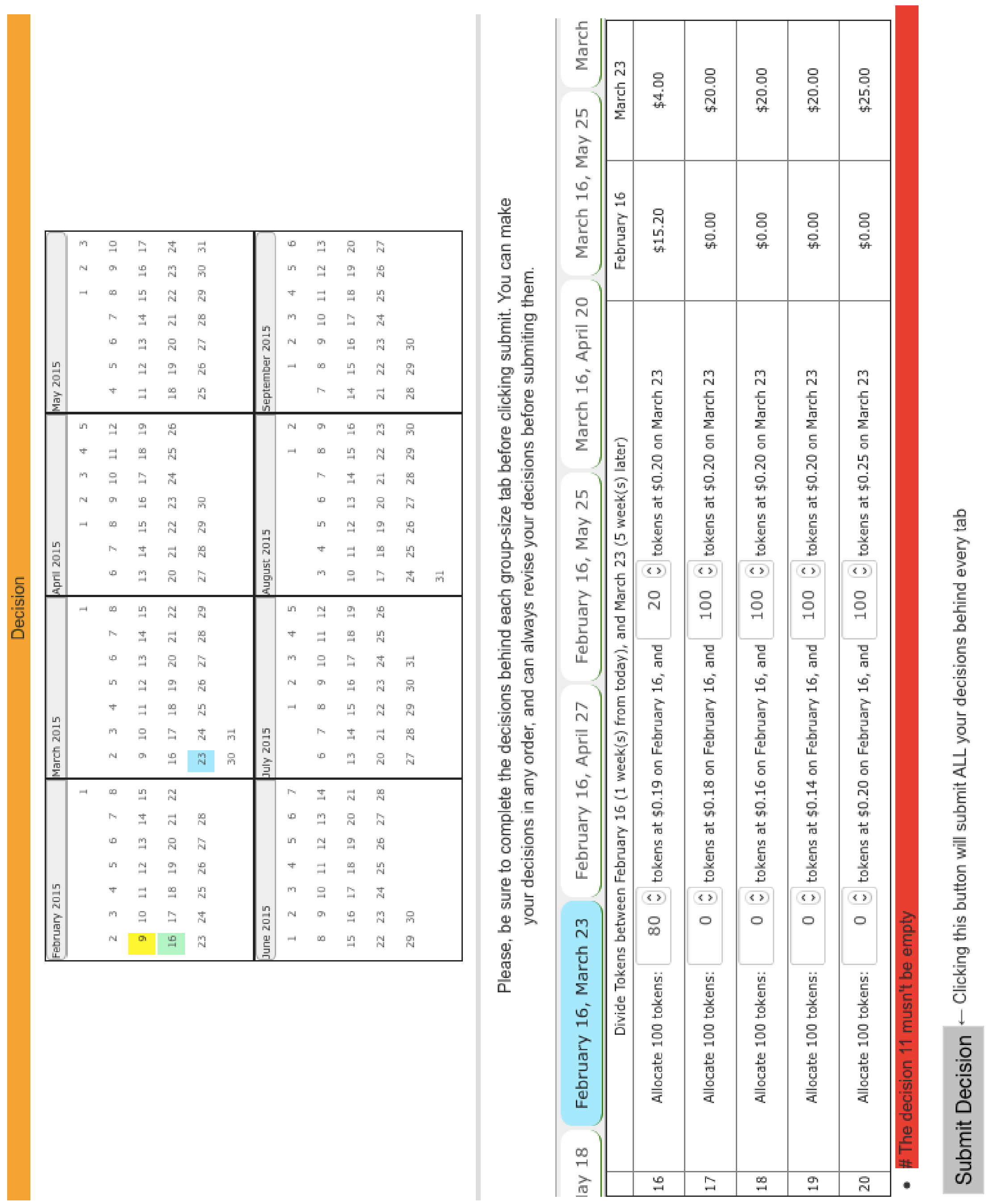


Figure S32: Screenshot of CTB4

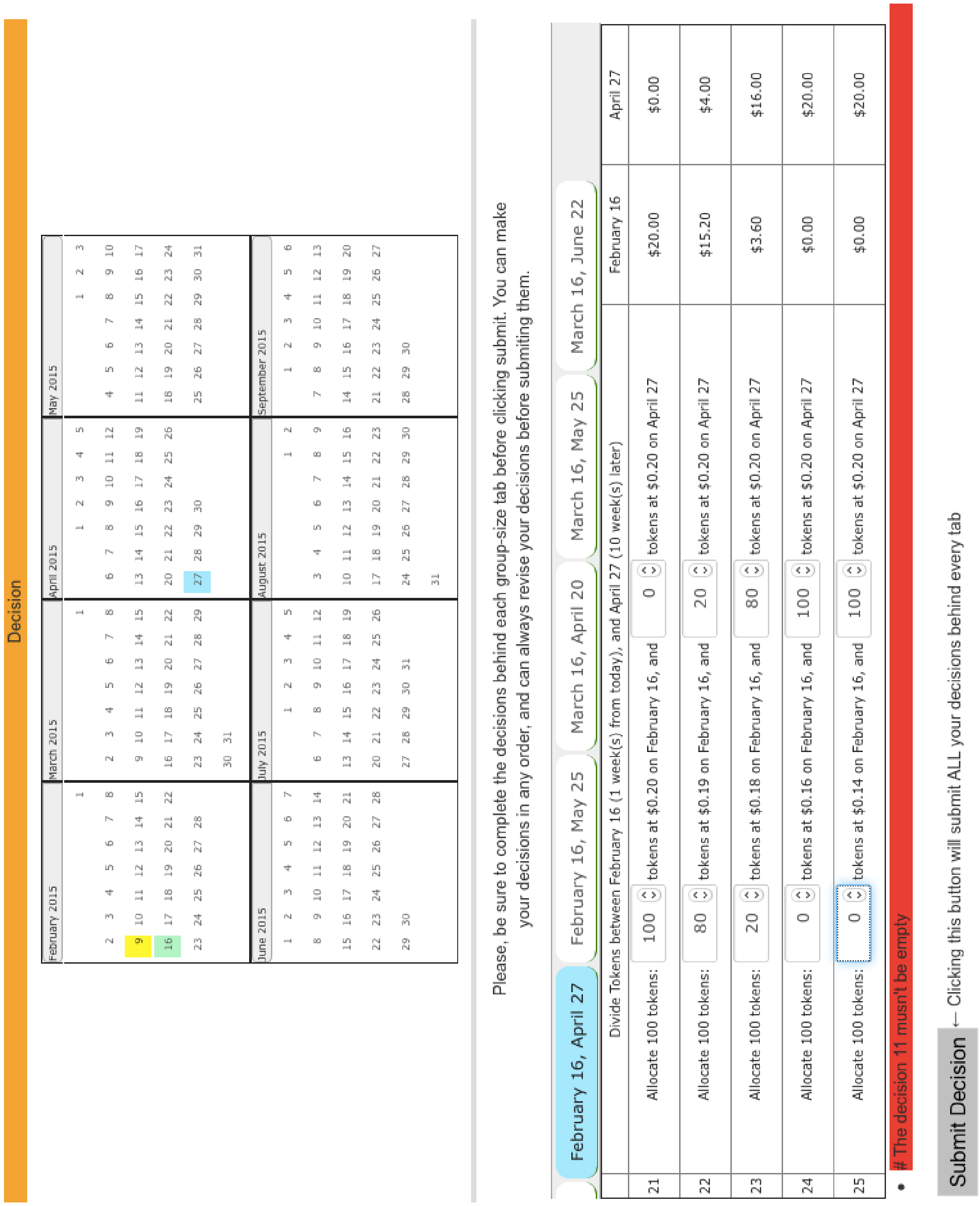


Figure S33: Screenshot of CTB5

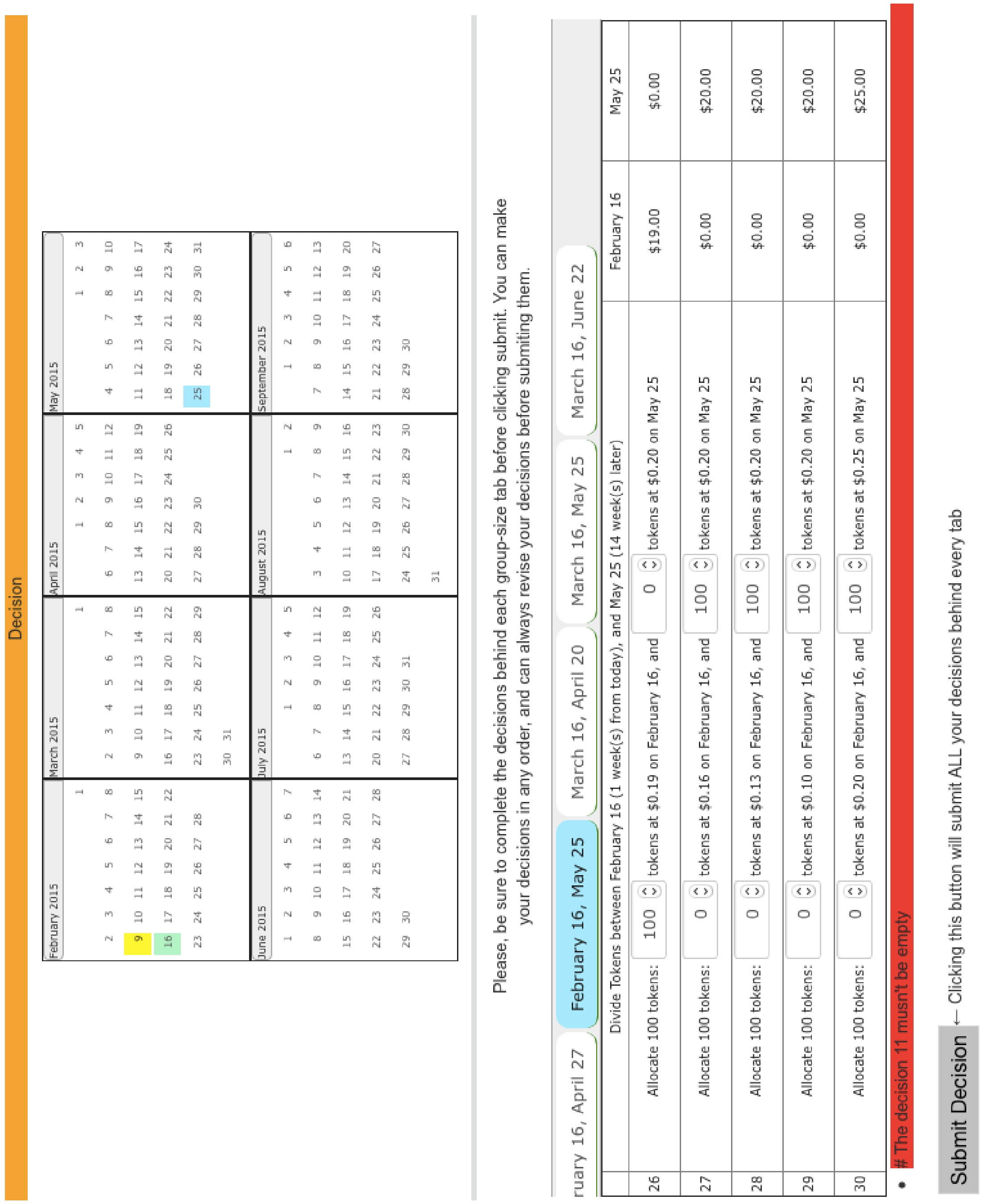


Figure S34: Screenshot of CTB6

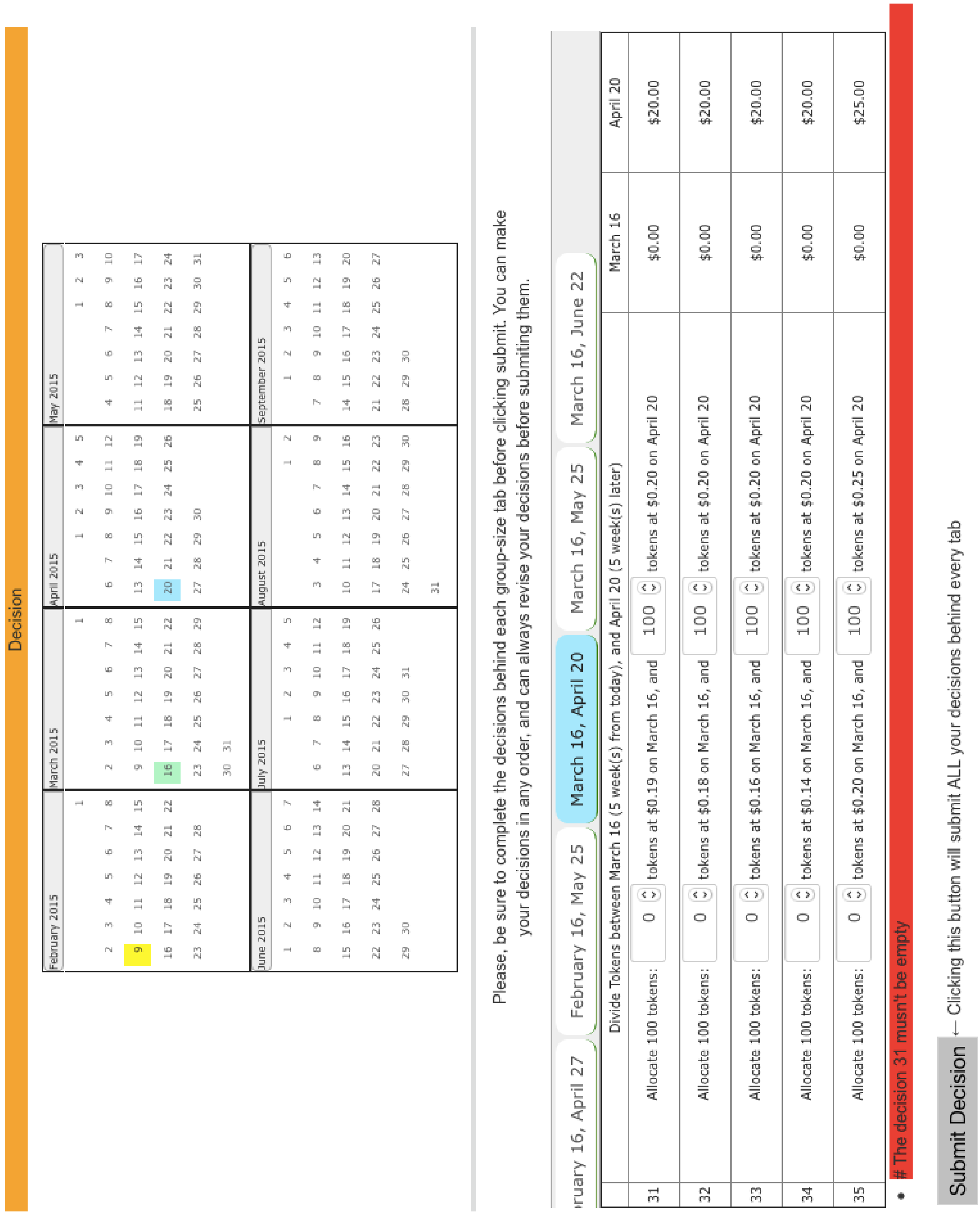


Figure S35: Screenshot of CTB7

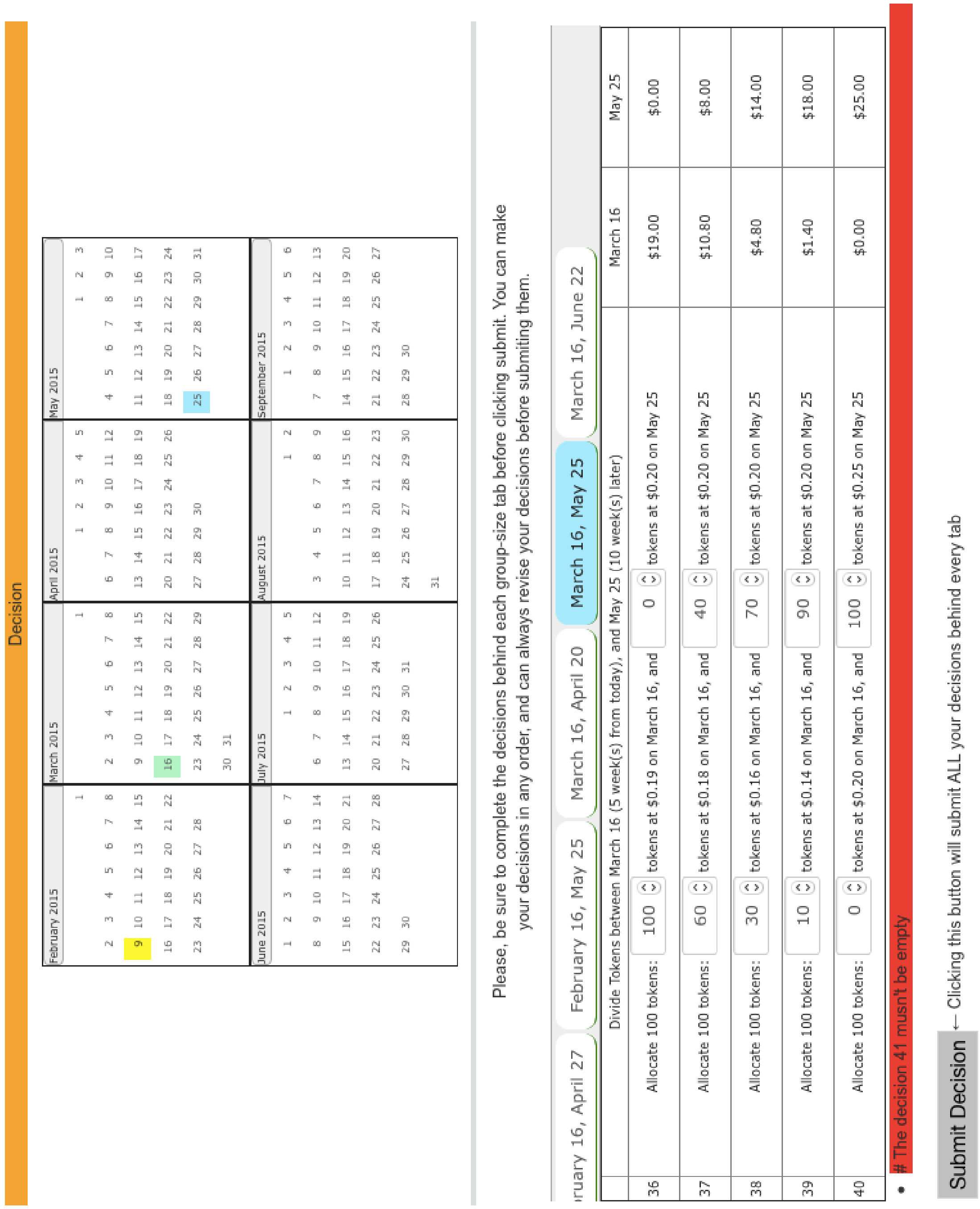


Figure S36: Screenshot of CTB8

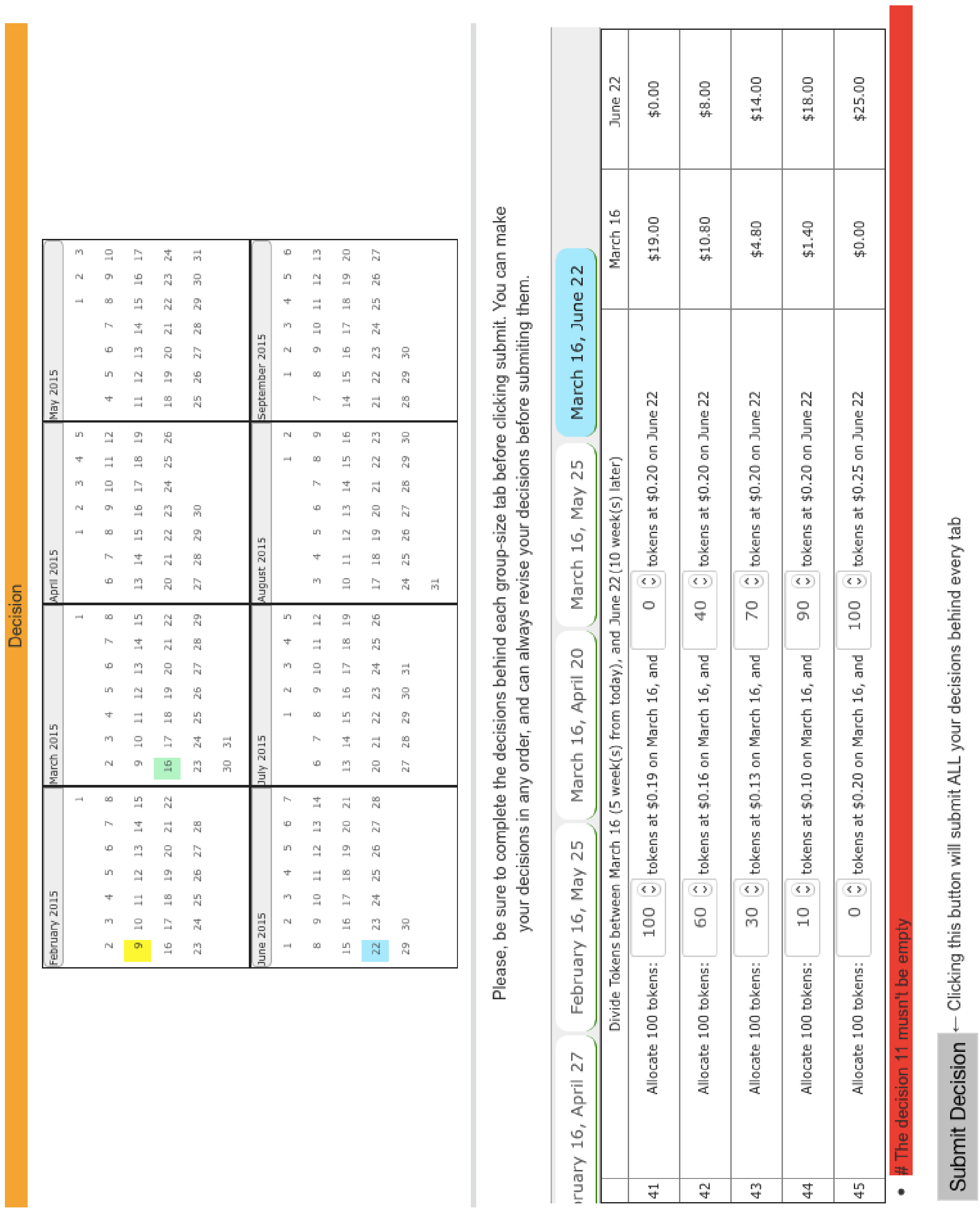


Figure S37: Screenshot of CTB9

Subjects then answered the Holt-Laury task. Figure S38 shows instructions of the task and Figure S39 displays the task example.

How It Works:

In the following two screens you are asked to choose between options: Option A or Option B.?On each screen you will make ten choices, one on each row. For each decision row you will have to choose either Option A or Option B. You make your decision by clicking on the circle next to the option you prefer more. You may choose A for some decision rows and B for other rows, and you may change your decisions and make them in any order.

There are a total of 20 decisions on the following two screens. These decisions represent one of the 47 choices you make in the experiment. If the number 46 is drawn, these decisions will determine your payoffs. If the number 46 is drawn, a second number will also be randomly drawn from 1 to 20. This will determine which decision (from 1 to 20) on the following two screens is the decision-that-counts. The option you choose (either Option A or Option B) in the decision-that-counts will then be played. You will receive your payment from the decisionÐthat-counts today. Your $5 sooner and later thank-you payments, however, will still be paid as before. The sooner payment will be made today and the later payment will be made in 5 weeks.

Playing the Decision-That-Counts:

Your payment in the decision-that-counts will be determined by throwing a 10-sided die (generated by the computer). Now, please look at Decision 1 below (the same decisions as on the next screen). Option A pays $10.39 if the throw of the ten-sided die is 1, and it pays $8.31 if the throw is 2-10. Option B yields $20 if the throw of the die is 1, and it pays $0.52 if the throw is 2-10. The other Decisions are similar, except that as you move down the table, the chances of the higher payoff for each option increase. In fact, for Decision 10 in the bottom row, the die will not be needed since each option pays the highest payoff for sure, so your choice here is between $10.39 or $20.

Remember that each decision could be the decision-that-counts! It is in your interest to treat each decision as if it could be the one that determines your payoff.

Figure S38: Screenshot of HL instructions

Decision	Option A						Option B					
		If the die reads	You receive	and	If the die reads	You receive		If the die reads	You receive	and	If the die reads	You receive
1	○	1	$10.39		2-10	$8.31	○	1	$20		2-10	$0.52
2	○	1-2	$10.39		3-10	$8.31	○	1-2	$20		3-10	$0.52
3	○	1-3	$10.39		4-10	$8.31	○	1-3	$20		4-10	$0.52
4	○	1-4	$10.39		5-10	$8.31	○	1-4	$20		5-10	$0.52
5	○	1-5	$10.39		6-10	$8.31	○	1-5	$20		6-10	$0.52
6	○	1-6	$10.39		7-10	$8.31	○	1-6	$20		7-10	$0.52
7	○	1-7	$10.39		8-10	$8.31	○	1-7	$20		8-10	$0.52
8	○	1-8	$10.39		9-10	$8.31	○	1-8	$20		9-10	$0.52
9	○	1-9	$10.39		10	$8.31	○	1-9	$20		10	$0.52
10	○	1-10	$10.39		-	$8.31	○	1-10	$20		-	$0.52

Start

Figure S39: Screenshot of the HL example

Figure S40 shows the first HL task and Figure S41 the second HL task, both answered.

Decision	Option A						Option B					
		If the die reads	You receive	and	If the die reads	You receive		If the die reads	You receive	and	If the die reads	You receive
1	●	1	$10.39		2-10	$8.31	○	1	$20		2-10	$0.52
2	●	1-2	$10.39		3-10	$8.31	○	1-2	$20		3-10	$0.52
3	●	1-3	$10.39		4-10	$8.31	○	1-3	$20		4-10	$0.52
4	●	1-4	$10.39		5-10	$8.31	○	1-4	$20		5-10	$0.52
5	●	1-5	$10.39		6-10	$8.31	○	1-5	$20		6-10	$0.52
6	●	1-6	$10.39		7-10	$8.31	○	1-6	$20		7-10	$0.52
7	●	1-7	$10.39		8-10	$8.31	○	1-7	$20		8-10	$0.52
8	○	1-8	$10.39		9-10	$8.31	●	1-8	$20		9-10	$0.52
9	○	1-9	$10.39		10	$8.31	●	1-9	$20		10	$0.52
10	○	1-10	$10.39		-	$8.31	●	1-10	$20		-	$0.52

Back Continue

Figure S40: Screenshot of the first HL task

Decision	Option A						Option B					
		If the die reads	You receive	and	If the die reads	You receive		If the die reads	You receive	and	If the die reads	You receive
11	●	1	$13.89		2-10	$5.56	○	1	$25		2-10	$0.28
12	●	1-2	$13.89		3-10	$5.56	○	1-2	$25		3-10	$0.28
13	●	1-3	$13.89		4-10	$5.56	○	1-3	$25		4-10	$0.28
14	●	1-4	$13.89		5-10	$5.56	○	1-4	$25		5-10	$0.28
15	●	1-5	$13.89		6-10	$5.56	○	1-5	$25		6-10	$0.28
16	○	1-6	$13.89		7-10	$5.56	●	1-6	$25		7-10	$0.28
17	○	1-7	$13.89		8-10	$5.56	●	1-7	$25		8-10	$0.28
18	○	1-8	$13.89		9-10	$5.56	●	1-8	$25		9-10	$0.28
19	○	1-9	$13.89		10	$5.56	●	1-9	$25		10	$0.28
20	○	1-10	$13.89		-	$5.56	●	1-10	$25		-	$0.28

back Submit <-- Clicking this button will submit ALL your decisions

Figure S41: Screenshot of the second HL task

Subjects then answered the MPL task. Figure S42 displays instructions of the task. Figure S43 shows the first MPL task, Figure S44 the second MPL task and Figure S45 the third MPL task, all answered.

How It Works:

On this screen you are asked to choose between smaller payments closer to today and larger payments further in the future. For each row, choose one payment: either the smaller, sooner payment or the larger, later payment. There are 22 decisions in total. Each decision has a number from 1 to 22.

NUMBERS 1 THROUGH 7: Decide between payment today and payment in five weeks

NUMBERS 8 THROUGH 15: Decide between payment today and payment in fourteen weeks

NUMBERS 16 THROUGH 22: Decide between payment today and payment in ten weeks

This screen represents one of the 47 choices you make in the experiment. If the number 47 is drawn, this screen will determine your payoffs. If the number 47 is drawn, a second number will also be drawn from 1 to 22. This will determine which decision (from 1 to 22) on this screen is the decision-that-counts. The payment you chose (either sooner or later) in the decision that counts will be added to either your earlier $5 thank-you payment or your later $5 thank-you payment.

Remember that each decision could be the decision that counts! Treat each decision as if it could be the one that determines your payment.

Start

Figure S42: Screenshot of MPL instructions

TODAY VS. FIVE WEEKS FROM TODAY

WHAT WILL YOU DO IF YOU GET A NUMBER BETWEEN 1 AND 7?

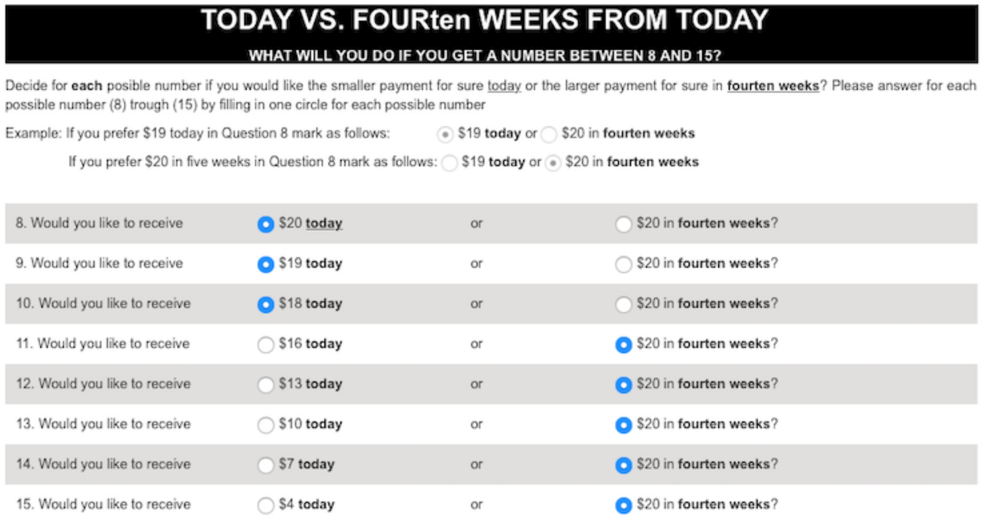
Decide for **each** posible number if you would like the smaller payment for sure today or the larger payment for sure in **five weeks**? Please answer for each possible number (1) trough (7) by filling in one circle for each possible number

Example: If you prefer $19 today in Question 1 mark as follows: $19 **today** or $20 in **five weeks**

If you prefer $20 in five weeks in Question 1 mark as follows: $19 **today** or $20 in **five weeks**

1. Would you like to receive	$19 **today**	or	$20 in **five weeks**?
2. Would you like to receive	$18 **today**	or	$20 in **five weeks**?
3. Would you like to receive	$16 **today**	or	$20 in **five weeks**?
4. Would you like to receive	$14 **today**	or	$20 in **five weeks**?
5. Would you like to receive	$11 **today**	or	$20 in **five weeks**?
6. Would you like to receive	$8 **today**	or	$20 in **five weeks**?
7. Would you like to receive	$5 **today**	or	$20 in **five weeks**?

Figure S43: Screenshot of the first MPL task



TODAY VS. FOURten WEEKS FROM TODAY

WHAT WILL YOU DO IF YOU GET A NUMBER BETWEEN 8 AND 15?

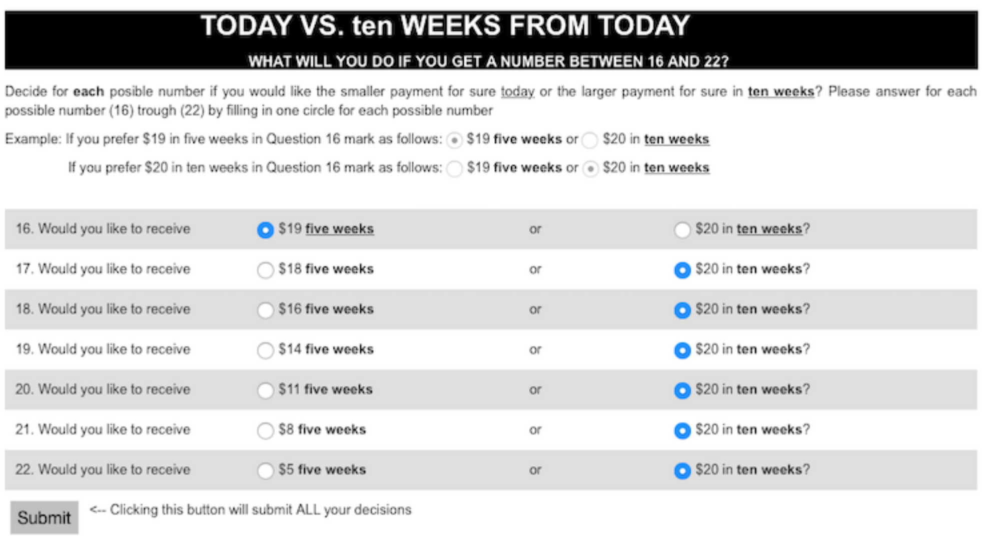
Decide for **each** posible number if you would like the smaller payment for sure today or the larger payment for sure in **fourten weeks**? Please answer for each possible number (8) trough (15) by filling in one circle for each possible number

Example: If you prefer $19 today in Question 8 mark as follows: $19 **today** or $20 in **fourten weeks**

If you prefer $20 in five weeks in Question 8 mark as follows: $19 **today** or $20 in **fourten weeks**

8. Would you like to receive	$20 today	or	$20 in **fourten weeks**?
9. Would you like to receive	$19 **today**	or	$20 in **fourten weeks**?
10. Would you like to receive	$18 **today**	or	$20 in **fourten weeks**?
11. Would you like to receive	$16 **today**	or	$20 in **fourten weeks**?
12. Would you like to receive	$13 **today**	or	$20 in **fourten weeks**?
13. Would you like to receive	$10 **today**	or	$20 in **fourten weeks**?
14. Would you like to receive	$7 **today**	or	$20 in **fourten weeks**?
15. Would you like to receive	$4 **today**	or	$20 in **fourten weeks**?

Figure S44: Screenshot of the second MPL task



TODAY VS. ten WEEKS FROM TODAY

WHAT WILL YOU DO IF YOU GET A NUMBER BETWEEN 16 AND 22?

Decide for **each** posible number if you would like the smaller payment for sure today or the larger payment for sure in **ten weeks**? Please answer for each possible number (16) trough (22) by filling in one circle for each possible number

Example: If you prefer $19 in five weeks in Question 16 mark as follows: $19 **five weeks** or $20 in **ten weeks**

If you prefer $20 in ten weeks in Question 16 mark as follows: $19 **five weeks** or $20 in **ten weeks**

16. Would you like to receive	$19 **five weeks**	or	$20 in **ten weeks**?
17. Would you like to receive	$18 **five weeks**	or	$20 in **ten weeks**?
18. Would you like to receive	$16 **five weeks**	or	$20 in **ten weeks**?
19. Would you like to receive	$14 **five weeks**	or	$20 in **ten weeks**?
20. Would you like to receive	$11 **five weeks**	or	$20 in **ten weeks**?
21. Would you like to receive	$8 **five weeks**	or	$20 in **ten weeks**?
22. Would you like to receive	$5 **five weeks**	or	$20 in **ten weeks**?

Submit <-- Clicking this button will submit ALL your decisions

Figure S45: Screenshot of the third MPL task

Figure S46 displays the screen in which subjects discovered the decision chosen for payment after being reminded the payment method.

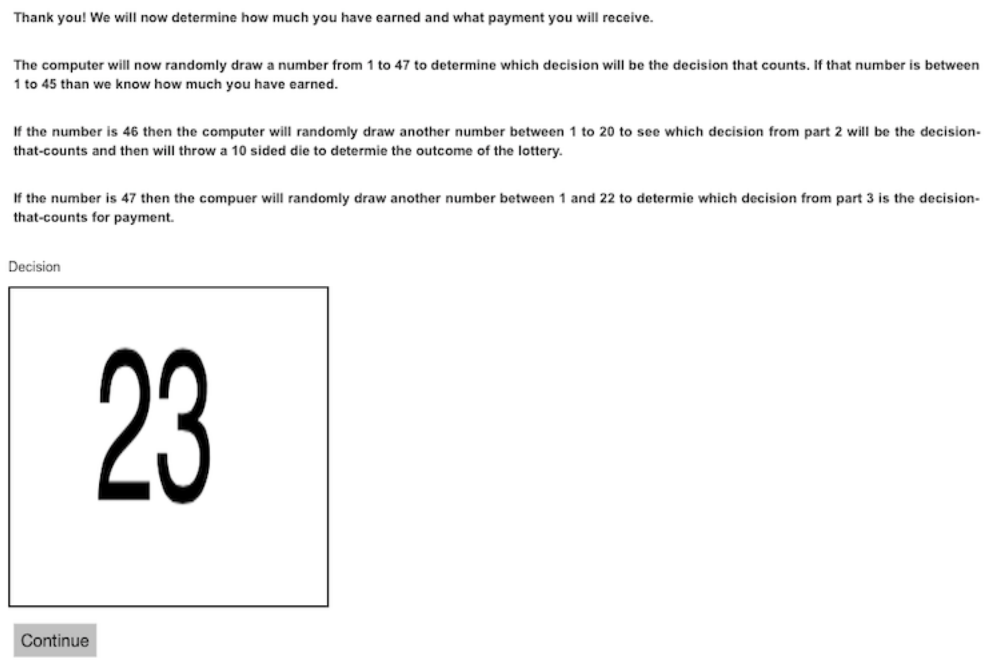


Figure S46: Screenshot of decision chosen for payment

Figure S47 displays the screen in which subjects discovered their experimental payment and how it was calculated :

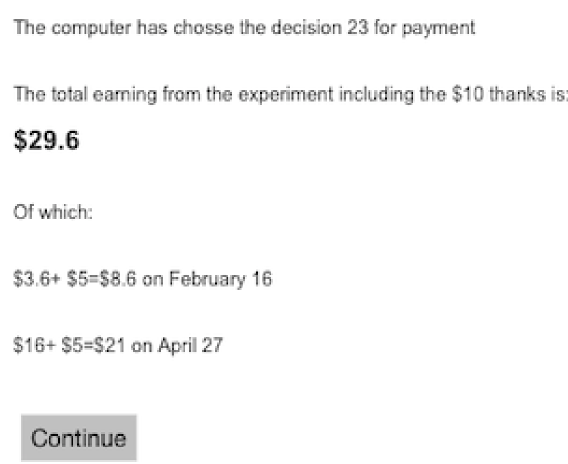
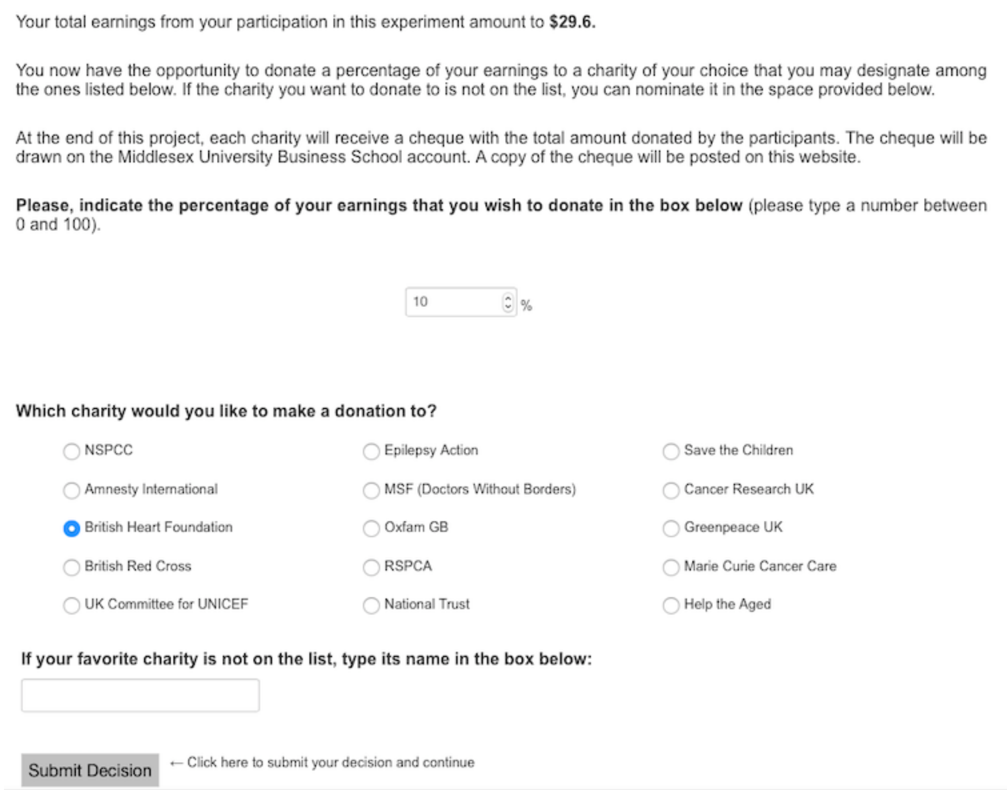


Figure S47: Screenshot of the experimental payment

Subjects answered the Dictator Game in the next screen. Figure S48 shows the task interface in which subjects indicated their donation to charities and selected the charity that will receive their donation :


Your total earnings from your participation in this experiment amount to **$29.6.**

You now have the opportunity to donate a percentage of your earnings to a charity of your choice that you may designate among the ones listed below. If the charity you want to donate to is not on the list, you can nominate it in the space provided below.

At the end of this project, each charity will receive a cheque with the total amount donated by the participants. The cheque will be drawn on the Middlesex University Business School account. A copy of the cheque will be posted on this website.

Please, indicate the percentage of your earnings that you wish to donate in the box below (please type a number between 0 and 100).

10 %

Which charity would you like to make a donation to?

NSPCC	Epilepsy Action	Save the Children
Amnesty International	MSF (Doctors Without Borders)	Cancer Research UK
British Heart Foundation	Oxfam GB	Greenpeace UK
British Red Cross	RSPCA	Marie Curie Cancer Care
UK Committee for UNICEF	National Trust	Help the Aged

If your favorite charity is not on the list, type its name in the box below:

Submit Decision ← Click here to submit your decision and continue

Figure S48: Screenshot of the Dictator Game

Finally, subjects were shown the screen in Figure S49 displaying their final experimental payoffs after donating to charities :

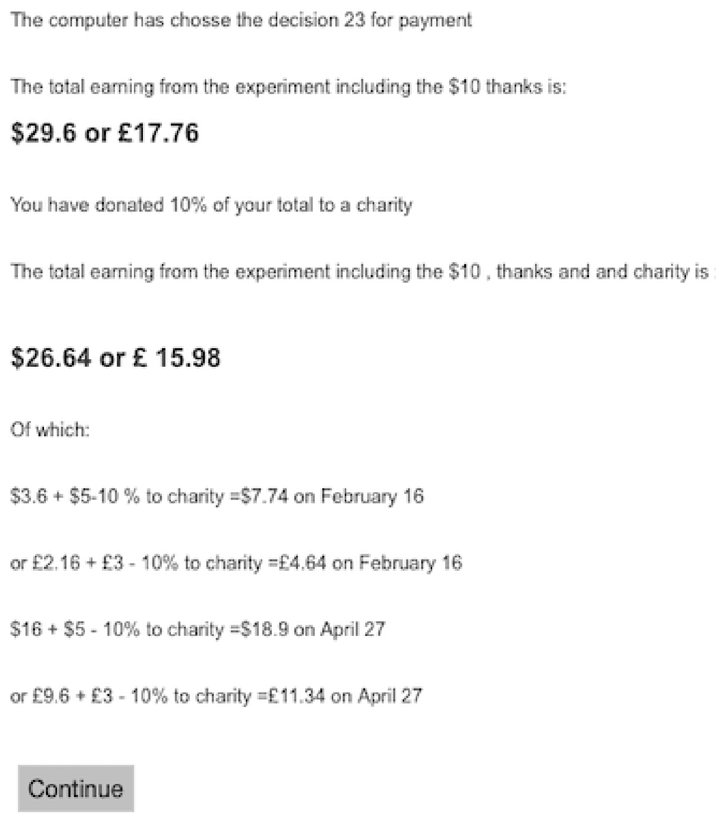
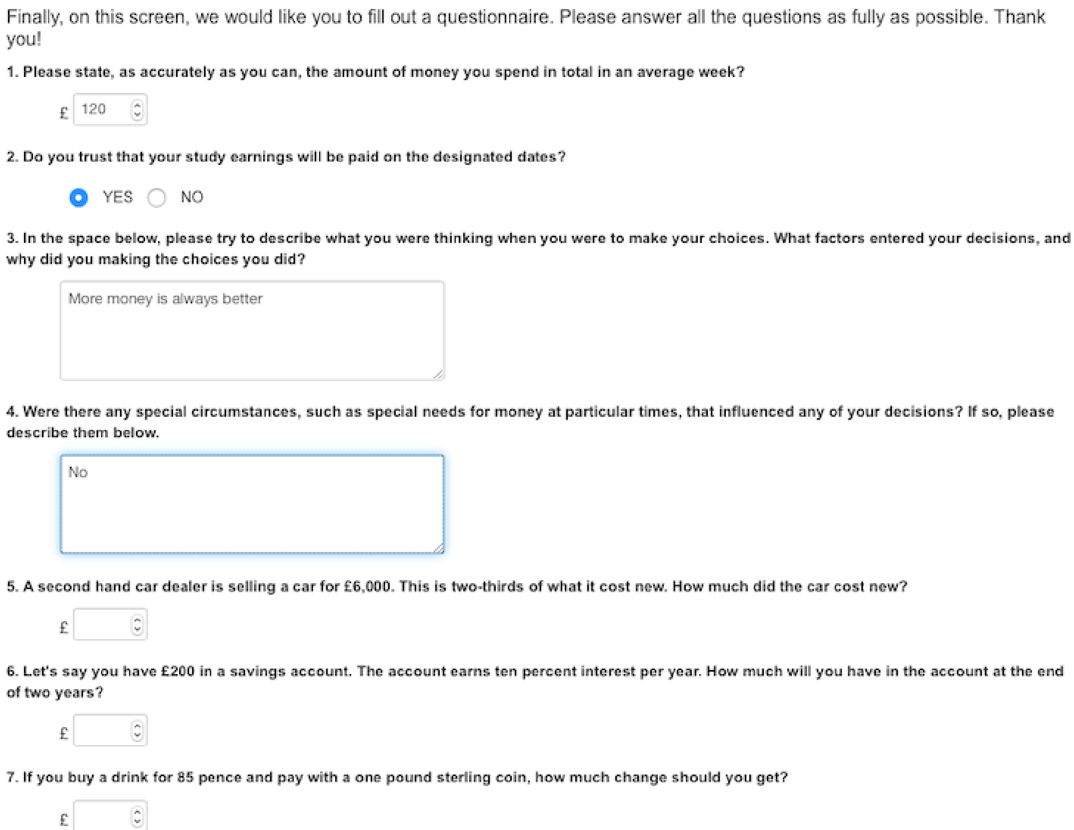


Figure S49: Screenshot of the final experimental payment

Subjects then answered a questionnaire. Figure S50 shows questions from 1 to 7, Figure S51 shows questions from 8 to 14, Figure S52 shows questions from 15 to 19, Figure S53 shows questions from 20 to 27, Figure S54 shows questions from 28 to 33, Figure S55 shows questions from 34 to 42 and Figure S56 shows question 43.



Finally, on this screen, we would like you to fill out a questionnaire. Please answer all the questions as fully as possible. Thank you!

1. Please state, as accurately as you can, the amount of money you spend in total in an average week?

£ 120

2. Do you trust that your study earnings will be paid on the designated dates?

YES NO

3. In the space below, please try to describe what you were thinking when you were to make your choices. What factors entered your decisions, and why did you making the choices you did?

More money is always better

4. Were there any special circumstances, such as special needs for money at particular times, that influenced any of your decisions? If so, please describe them below.

No

5. A second hand car dealer is selling a car for £6,000. This is two-thirds of what it cost new. How much did the car cost new?

£

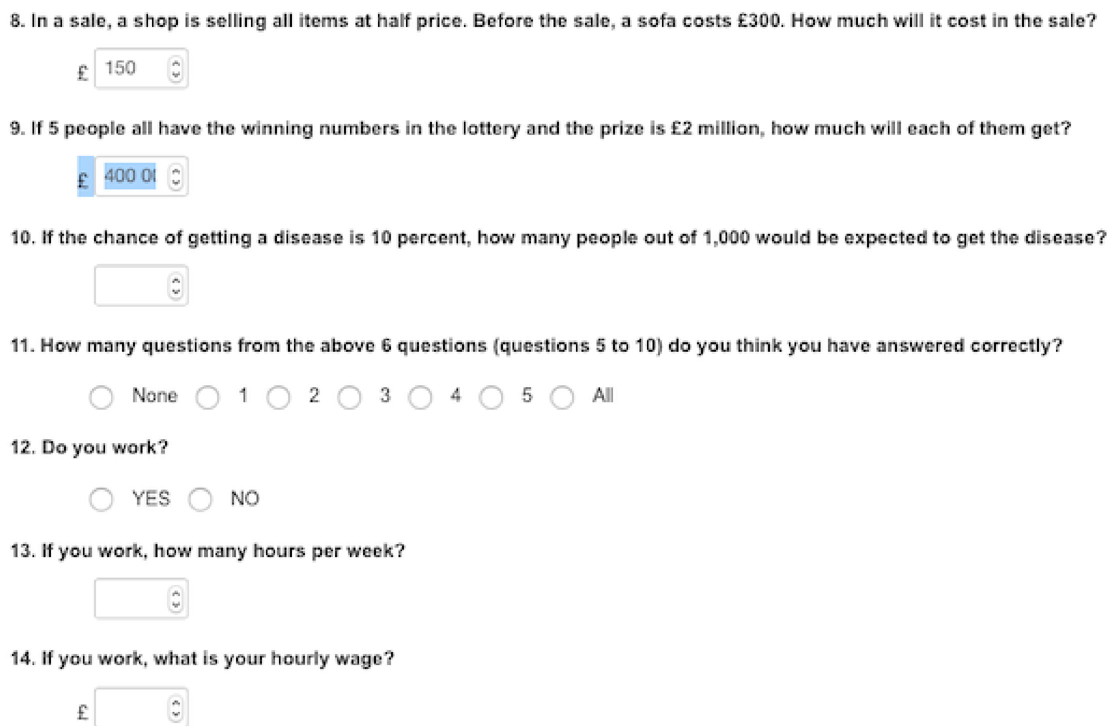
6. Let's say you have £200 in a savings account. The account earns ten percent interest per year. How much will you have in the account at the end of two years?

£

7. If you buy a drink for 85 pence and pay with a one pound sterling coin, how much change should you get?

£

Figure S50: Screenshot of questions 1 to 7



8. In a sale, a shop is selling all items at half price. Before the sale, a sofa costs £300. How much will it cost in the sale?

£ 150

9. If 5 people all have the winning numbers in the lottery and the prize is £2 million, how much will each of them get?

£ 400 0(

10. If the chance of getting a disease is 10 percent, how many people out of 1,000 would be expected to get the disease?

11. How many questions from the above 6 questions (questions 5 to 10) do you think you have answered correctly?

None 1 2 3 4 5 All

12. Do you work?

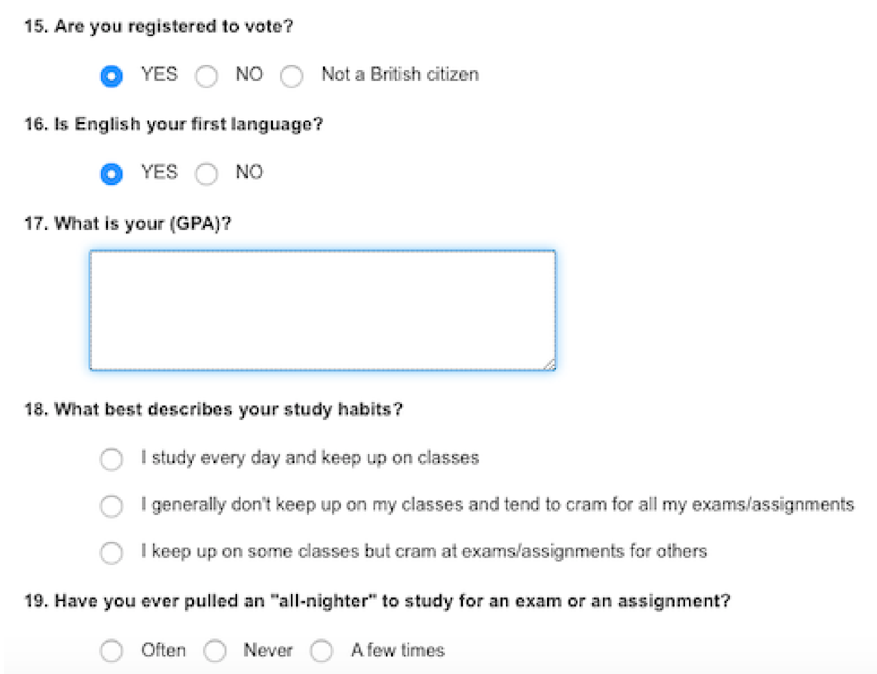
YES NO

13. If you work, how many hours per week?

14. If you work, what is your hourly wage?

£

Figure S51: Screenshot of questions 8 go 14



15. Are you registered to vote?

YES NO Not a British citizen

16. Is English your first language?

YES NO

17. What is your (GPA)?

18. What best describes your study habits?

I study every day and keep up on classes

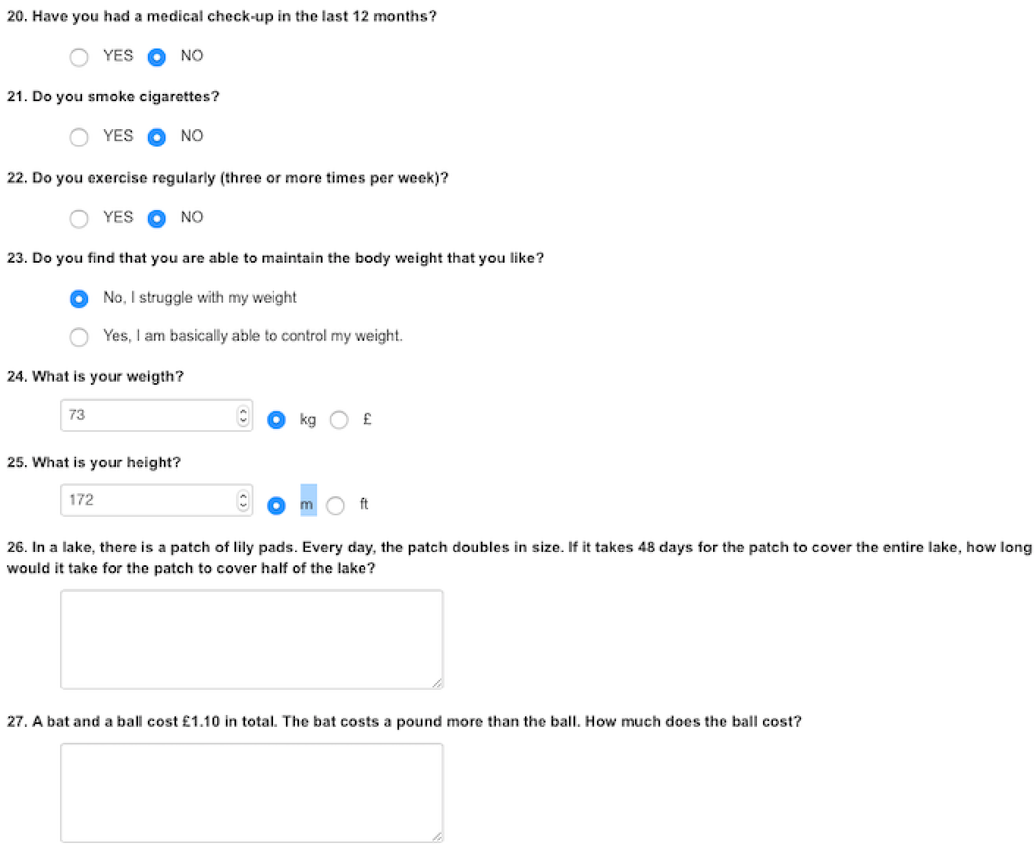
I generally don't keep up on my classes and tend to cram for all my exams/assignments

I keep up on some classes but cram at exams/assignments for others

19. Have you ever pulled an "all-nighter" to study for an exam or an assignment?

Often Never A few times

Figure S52: Screenshot of questions 15 to 19



20. Have you had a medical check-up in the last 12 months?

YES NO

21. Do you smoke cigarettes?

YES NO

22. Do you exercise regularly (three or more times per week)?

YES NO

23. Do you find that you are able to maintain the body weight that you like?

No, I struggle with my weight

Yes, I am basically able to control my weight.

24. What is your weigth?

73 kg £

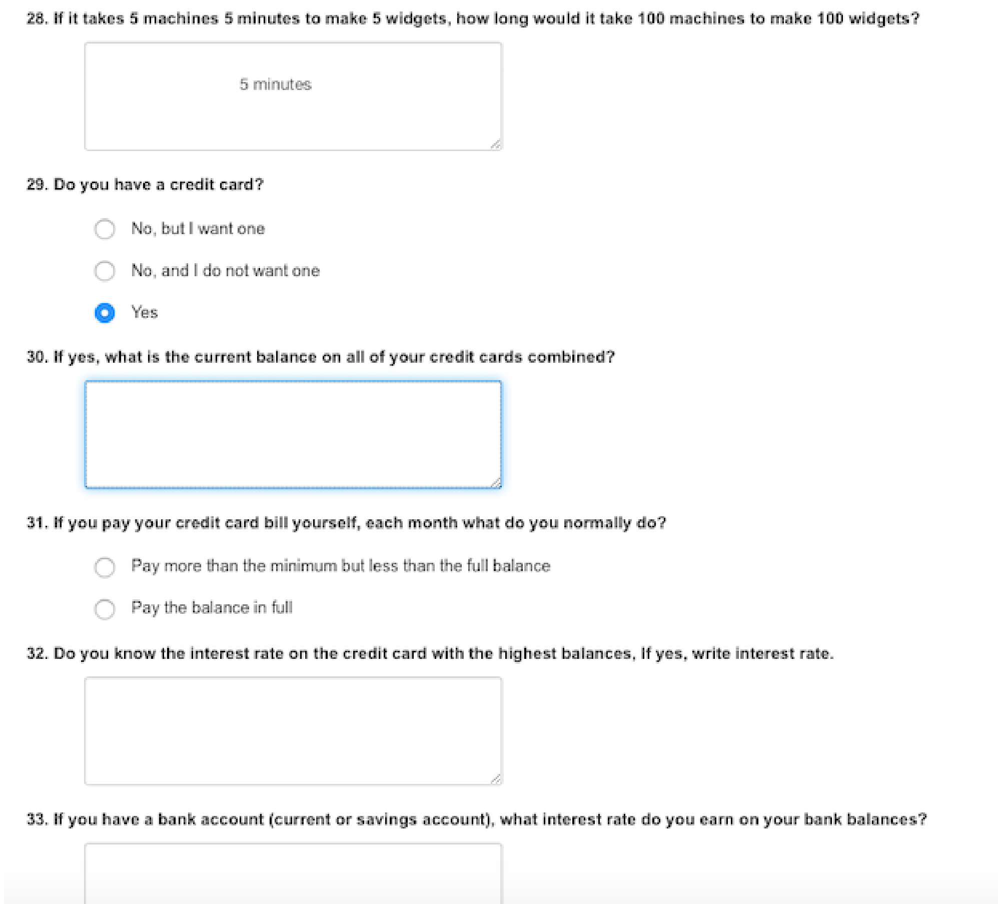
25. What is your height?

172 m ft

26. In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake?

27. A bat and a ball cost £1.10 in total. The bat costs a pound more than the ball. How much does the ball cost?

Figure S53: Screenshot of questions 20 to 27



28. If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets?

5 minutes

29. Do you have a credit card?

- No, but I want one
- No, and I do not want one
- Yes

30. If yes, what is the current balance on all of your credit cards combined?

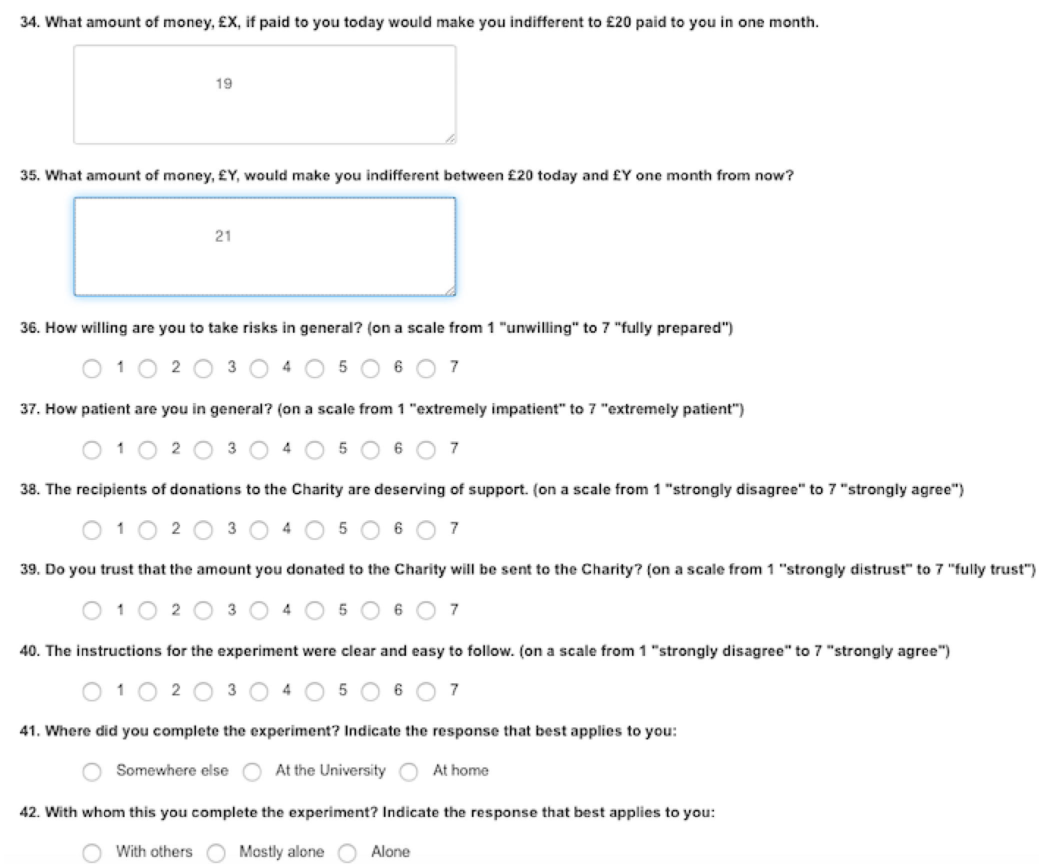
31. If you pay your credit card bill yourself, each month what do you normally do?

- Pay more than the minimum but less than the full balance
- Pay the balance in full

32. Do you know the interest rate on the credit card with the highest balances, If yes, write interest rate.

33. If you have a bank account (current or savings account), what interest rate do you earn on your bank balances?

Figure S54: Screenshot of questions 28 to 33



34. What amount of money, £X, if paid to you today would make you indifferent to £20 paid to you in one month.

19

35. What amount of money, £Y, would make you indifferent between £20 today and £Y one month from now?

21

36. How willing are you to take risks in general? (on a scale from 1 "unwilling" to 7 "fully prepared")

○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7

37. How patient are you in general? (on a scale from 1 "extremely impatient" to 7 "extremely patient")

○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7

38. The recipients of donations to the Charity are deserving of support. (on a scale from 1 "strongly disagree" to 7 "strongly agree")

○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7

39. Do you trust that the amount you donated to the Charity will be sent to the Charity? (on a scale from 1 "strongly distrust" to 7 "fully trust")

○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7

40. The instructions for the experiment were clear and easy to follow. (on a scale from 1 "strongly disagree" to 7 "strongly agree")

○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7

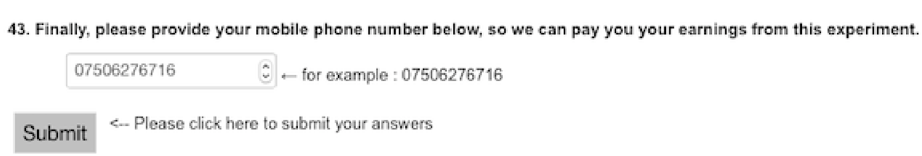
41. Where did you complete the experiment? Indicate the response that best applies to you:

○ Somewhere else ○ At the University ○ At home

42. With whom this you complete the experiment? Indicate the response that best applies to you:

○ With others ○ Mostly alone ○ Alone

Figure S55: Screenshot of questions 34 to 42



43. Finally, please provide your mobile phone number below, so we can pay you your earnings from this experiment.

07506276716 ← for example : 07506276716

Submit <-- Please click here to submit your answers

Figure S56: Screenshot of question 43