**Economics Working Paper Series**

**2022/005**

# Rituals of Reason: A Choice-Based Approach to the Acceptability of Lotteries in Allocation Problems

Elias Bouacida and Renaud Foucart

The Department of Economics
Lancaster University Management School
Lancaster LA1 4YX
UK

# Rituals of Reason: A Choice-Based Approach to the Acceptability of Lotteries in Allocation Problems*

Elias Bouacida[†] Renaud Foucart[‡]

December 20, 2021

**Abstract**

We study revealed preferences towards the use of random procedures in allocation mechanisms. We report the results of an experiment in which subjects vote on a procedure to allocate a reward to half of them. The first possibility is an explicitly random device: the result of a lottery. The second is either an unpredictable procedure they could interpret as meritocratic, or one that is obviously arbitrary. We run all treatments with and without control. We identify an aversion to lotteries and clearly arbitrary procedures across treatments, even though, on aggregate, subjects do not believe any procedure to give them a higher probability of success and there is no correlation between beliefs and outcomes. In line with the literature, we also find evidence of a control premium in most procedures.

**Keywords:** lotteries, mechanism design, allocation problems, procedures, tie-breaking rule

**JEL-Code:** D01, D78, D91

# 1 Introduction

A frequent issue in allocation problems is the choice of a procedure to break ties when no better criterion is available. While many of the solutions provided by economists involve a role for randomization and the use of lotteries,[1] random tie-breakers are often unpopular in practice. The goal of this paper is to provide incentivized evidence on individual preferences toward the role of lotteries in allocation problems. In an experiment, we ask subjects to choose between two unpredictable procedures, one clearly random and the

*We thank Micael Castanheira, Konstantinos Georgalos, Rustamdjam Hakimov, Kim Kaivanto, Dorothea Kubler, Patrick Legros, Jean-Marc Tallon and Eyal Winter for useful discussions and comments. We thank seminar participants at Lancaster University, ESA 2020 world meeting, University Paris 8, University Paris Dauphine and Université libre de Bruxelles for feedback and suggestions.

†Université Paris 8, elias.bouacida@univ-paris8.fr

‡Lancaster University Management School, r.foucart@lancaster.ac.uk

[1] Randomization is widely accepted by economists as the fairest way to deal with ties in the criteria-based allocation, albeit there is a debate on the best way to implement it, (see for instance Budish et al., 2013 for a review, and Erdil and Ergin, 2008; Kesten and Ünver, 2015; Abdulkadiroglu et al., 2019 for the case of school choice). Lotteries can also have efficiency benefits: Basteck et al. (2021) show for instance that the introduction of a lottery quota strengthens truth-telling in the deferred acceptance mechanism.

other not, to allocate a reward to themselves. We document a relative aversion to the explicit use of lotteries, as long as the alternative procedure follows the *rituals of reason* (Elster, 1989, p.37):[2] albeit unpredictable, it is reminiscent of meritocratic procedures and subjects may be able to interpret it as such. This aversion cannot be explained by overconfidence and can be mitigated by offering control over the lottery.

Experimental evidence shows that lotteries are perceived as a fair procedure and help make unequal outcomes socially acceptable (Bolton et al., 2005; Schmidt and Trautmann, 2019). Famous examples of allocation problems solved by lotteries are the military draft, the selection of jurors, the allocation of social housing in several US cities, the Rotating Savings and Credit Associations (ROSCAs), or the green card lottery. Lotteries are however not ubiquitous, and human beings are often hostile to see their fate decided by the toss of a coin. In USA law, the Administrative Procedure Act (APA) forbids the use of random devices in administrative decisions.[3] Several European countries have witnessed a large opposition to school choice mechanisms involving randomization.[4] The UK government's "school admission code (2014)" explicitly states (p.14) that "local authorities must not use random allocation as the principal oversubscription criterion".[5] Medical researchers often report difficulties to run random control trials (RCT), as doctors oppose allocating a treatment at random.[6] Even economists are reluctant to use lotteries for themselves. To allocate funding, academics spend a lot of time ranking research projects of similar quality and reach conclusions that are as good as random (Cole et al., 1981; Graves et al., 2011; Pier et al., 2018). Proposals to save time and money by adding random tie breakers to the criteria used for peer review (Greenberg, 1998; Brezis, 2007; Fang and Casadevall, 2016; Roumbanis, 2019; Avin, 2019) are however largely ignored (Barnett, 2016).[7]

We want to understand why and to what extent it is hard to use lotteries in practical allocation problems. The main novelty of our approach is to ask non-hypothetical questions: subjects' earnings during the experiment depend on the allocation mechanism the majority of them choose. An individual displays a preference for a non-random procedure

[2] "We have a strong reluctance to admit uncertainty and indeterminacy in human affairs. Rather than accept the limits of reason, we prefer the rituals of reason."

[3] The APA [requires court to] set aside agency action that is "arbitrary, capricious, [or] an abuse of discretion." (Vermeule, 2015, p.475) The US Supreme court (Judulang v. Holder, 132 S. Ct. 476, 485 [2011] [Kagan, J.]) rules that the use of a random device is arbitrary and thus illegal (p.11), regardless of how costly it is to provide a rationalization (p.21).

[4] In France, the system of allocation of students to universities (APB) used lotteries to break ties when capacity was reached. It was criticized and replaced in 2018 by another mechanism with a lot more criteria. In French-speaking Belgium, a 2009 attempt to randomly allocate students in oversubscribed high schools lasted only a year after being dubbed by some parents and the media the "lottery law" and replaced by a set of criteria still in use today.

[5] This debate followed the introduction in 2008 of a random tie-breaking rule in Brighton and Hove, after a criterion of distance to the school. Lotteries are also used as last-resort criteria in cities such as Birmingham. The issue remains a political talking point, and the 2017 Conservative Party manifesto explicitly committed the government to (p.50): "never introduce a mandatory lottery-based school admissions policy."

[6] In the medical literature, randomization is seen as ethical if a state of *equipoise* is reached: a consensus that none of the possible outcomes of the lottery is *ex ante* better than the others (Lilford and Jackson, 1995). It is however very difficult for a doctor not to seek an additional rationalization instead of declaring two possible treatments to be *ex ante* equivalent (Donovan et al., 2014).

[7] The Health Research Council of New Zealand is a notable exception, running a pilot random allocation for its early career "Explorer" grants since 2013. The Volkswagen fundation in Germany is another one.

if, between two *payoff-equivalent* options, she prefers the non-random one. This definition excludes an important reason why people may dislike lotteries: a majority of subjects can vote against lotteries simply because they expect to perform better in the alternative procedure. The preference for the non-random procedure can be *absolute* or *relative*. By a relative preference, we mean that subjects oppose a lottery because they prefer a specific alternative procedure, for instance, one they interpret as fair or meritocratic. By an absolute preference, we mean that subjects prefer even a completely arbitrary procedure to an explicitly random device. Understanding individual preferences towards lotteries can help policy makers increase the social acceptability of random tie breakers when they are part of an optimal mechanism.

In our experiment, subjects vote between two mechanisms allocating a reward to half of the participants in their group. As alternatives to a lottery, we offer three other procedures. We call the procedures without explicit randomization *criteria*. The first criterion is the outcome of a modified game of rock-paper-scissors (*RPS*), in which subjects compete against each other. The RPS procedure follows the rituals of reason: albeit unpredictable, it is reminiscent of a meritocratic procedure. In the second criterion, *Paintings*, subjects have to guess, from pairs of paintings selected by the second author of this paper, the ones preferred by the first author. This criterion also follows the rituals of reason but adds the dimension of a procedure depending on the choices of a central decision maker on top of competition between subjects. As the choices of the experimenters are transparent and made before the experiment, it does not involve any reciprocal preferences or betrayal aversion. Finally, the third criterion, *Time*, is the outcome of a completely arbitrary and intractable algorithm that does not follow the rituals of reason. Its only potential appeal is the absence of a randomization device.

For all three criteria and the lottery, we run two different treatments. In the first one, subjects make no decision at all. In the second, we give them control over some parameters or *strategies*. It allows us to control for the well-known result that subjects prefer to be in charge, even if the decisions they take are perfectly meaningless (Owens et al., 2014; Bartling et al., 2014; Bobadilla-Suarez et al., 2017; Ferreira et al., 2020). This concept is known in the literature as a *control premium* or the *intrinsic value of decision rights*. It is related to the idea of the *illusion of control* (Langer, 1975; Sloof and von Siemens, 2017). The latter explains the preference for control by overconfidence, an explanation we can rule out in our experiment, while the former constitutes a direct preference for decision rights.

The number of participants earning the reward is the same in all procedures and all procedures involved the same number of choices in the treatments with control. While our subjects are in general overconfident, their overconfidence is roughly similar across treatments. Finally, all the procedures are equally unpredictable: we found no correlation between the beliefs of our subjects about their performance and their actual performance.

We find evidence of widespread relative preference for non-random procedures. 61% of our subjects vote for the two procedures following the rituals of reason (RPS and Paintings) over lotteries. Absolute preference for non-random procedure is less frequent, as only 43% of the participants prefer our completely arbitrary criterion Time over the lottery. As we expect from the literature on control, relative preference for non-random procedures is stronger when those involve some form of control, but this effect is small and not always significant. Perhaps surprisingly, control matters more for the lottery (11% increase in lottery choice with control) than for any of the criteria.

We ran the experiment using Amazon Mechanical Turk (AMT thereafter), a crowdsourced online sample of workers popular in economics (Rand et al., 2012; Kuziemko et al., 2015) and other social sciences. We chose this pool of participants for two main reasons. First, while the results obtained on AMT typically go in the same direction as those of subjects participating in lab experiments, they are closer to being representative of the general population (Snowberg and Yariv, 2021). Moreover, the incentives in our study address the main concerns generally associated with AMT, such as the risk of participants not telling the truth (see Hauser et al., 2019 for a review). Second, we recruited 1,324 participants with well-understood characteristics who completed the main experiment over three sessions, something we could never have done with the usual pool of subjects of our university lab.

The novelty of our approach is to provide to the subjects of an allocation problem an incentivized choice between a lottery and a criterion. Non-incentivized survey evidence shows that people are reluctant to use a random device to determine the outcome of important hypothetical decisions involving other people (Keren and Teigen, 2010) when other criteria are available. Oberholzer-Gee et al. (1997) also report survey evidence that a market mechanism is the only less acceptable procedure than a lottery for the allocation of a nuclear waste facility. Finally, Eliaz and Rubinstein (2014) provide the characteristics that make subjects perceive a random procedure as fair. We return to these characteristics when introducing our lottery treatment.

An important feature of our study is that we focus on preferences for allocation problems in which the decision makers are also the subjects of the mechanism. It is not the case, for instance, for the military conscription, the choice of jurors, school choice, administrative decisions, or the funding of research projects, which mostly do not use lotteries.

On the contrary, in migration policy such as the green card lottery or the tie-breaking rule of the short-term skilled migration programs H-1B in the USA (Pathak et al., 2020), the participants are, by definition, not US citizens. Their preferences have no direct impact on the social acceptability of the random procedure by the US public. In the allocation of social housing, the decision to randomize is mostly made by people unlikely to apply for one. Similarly, when economists run RCTs, they expect the treatment to benefit a share of the population, and to have no effect or even a negative one on those who were not treated (Aldashev et al., 2017; Deaton and Cartwright, 2018; Heckman, 2020). Whether the subjects of the RCTs would have preferred another allocation mechanism is generally not discussed.

Our approach also differs from the well-known result that individuals often choose to rely on a randomization device to make their choice when indifferent or indecisive between two alternatives (see Agranov and Ortoleva, 2021 for a recent example). This *preference for randomization* has also been observed in school choice (Dwenger et al., 2018), an allocation problem in which many subjects seem however reluctant for the mechanism itself to use lotteries.

Our results are related to the concept of *outcome bias*, a tendency to interpret success by merit and effort and ignore the role of luck (Frank, 2016; Brownback and Kuhn, 2019). Hence, subjects may be willing to interpret any device that is not explicitly a lottery as more meritocratic, even if it is in practice completely unpredictable. Once a reward has been received for reasons perceived as more meritocratic, it is then valued more (Loewenstein and Issacharoff, 1994). This point is also related to the literature on *source uncertainty*: individuals treat uncertainty differently depending on the mechanism generating it (see, for instance, Heath and Tversky, 1991; Fox and Tversky, 1995;

Abdellaoui et al., 2011).

In Section 2, we present the design of the experiment. Section 3 outlines the main results. We relate those results to previous experiments in Section 4. Finally, we provide some suggestions on how to make random tie-breaking rules socially acceptable and conclude in Section 5.

# 2 Design of the Experiment

We ran the experiment on Amazon Mechanical Turk (AMT), on June 22, July 7 and August 31, 2021, using oTree (Chen et al., 2016). We ran a pilot experiment (see Appendix A) on April 29 and May 5, 2020. In the main experiment, subjects receive a fixed payment of \$0.80, and half of them receive an additional payment of \$1.60. We also offer additional rewards to incentivize the elicitation of beliefs and ambiguity aversion. The median time the 1,324 participants spend in the experiment is 4 minutes and 41 seconds. They receive a median hourly payment of \$19.61,[8] way above the minimum wage and perceived as a high payment by AMT workers.[9]

We randomly allocate subjects among 12 treatments of on average 110 participants. The experiment is composed of three parts. First, we introduce two procedures to the subjects: one is explicitly random, and the other is not. They also choose their strategies for the procedure(s) in which they have control (if any). In the second part, we ask them to vote on which procedure to use to allocate a reward to half of the members of their group. We implement the choice of the majority, at the end of the experiment. In the third part, we ask them incentivized questions on their beliefs about their absolute and relative ranks in the mechanism, and one incentivized measure of their ambiguity aversion. Finally, we ask non-incentivized demographic and feedback questions. The complete instructions and screen shots are presented in Online Appendix D. You can test the experiment by following this link: https://bouacida-foucart.herokuapp.com/room/readers.

For all the treatments, we provide subjects with a personal ID code. The next day, we post all the strategies and results on a website owned by the University of Lancaster, to allow subjects to check their strategies, as well as our procedures (see Online Appendix D). There is an active community of AMT workers exchanging tips and news on dedicated message boards, so this is a way to increase transparency and make participants aware that we are not trying to deceive them.[10] The payments were also made the next day, on their AMT account. It is not unusual for AMT workers as requester often want to check the work before payment.

We first describe the explicitly random procedure and then the three non-random ones, i.e, the *criteria*.

[8]Following the results of Snowberg and Yariv (2021) who showed using a battery of experimental games that doubling the rewards on AMT has no effect on the outcome, we do not expect our experimental evidence to depend on the exact level of the rewards.

[9]For instance, on Turkerview a discussion platform for AMT workers, we are rated as of 29$^{\text{th}}$ September 2021 as "Workers feel this requester pays well".

[10]While we did not expect all subjects to spend time checking their results, we have anecdotal evidence that some did, as they subsequently contacted us by email to discuss them.

## 2.1 An Explicitly Random Procedure: the DC-5 Lottery

Our random procedure is a bet on whether each of the 5 numbers in the next-day results of a state lottery, the DC-5 lottery, are odd or even. The state lottery is a transparent procedure run by a third party. If anything, our subjects should trust this procedure as being the least biased of all the ones we offer, as they can directly verify the results on the website of the lottery and we cannot influence it. There is also no ambiguity in this procedure as the probabilities of success are common knowledge, and equal to 50% for all subjects.

Our lottery satisfies the criteria identified by Eliaz and Rubinstein (2014) as characterizing a fair random procedure. First, all individuals are treated equally ex-ante: they either face the same decision (in the control treatment), or the bets we allocate them all have the exact same probability of being correct. Second, it allows all individuals to take part in the procedure whatever the realization of the random elements. Third, it delays any asymmetry in the treatment of participants to as late a stage as possible in the procedure, as the lottery only happens the day after the experiment takes place. Fourth, there is no psychological burden associated with the perception that the individual who executes a random device bears some responsibility for its outcome, as the state lottery is not run by a specific individual. Fifth, it uses a conventional, familiar means of randomization, as lotteries are a very old and widespread procedure. Finally, it respects the divine providence as manifested in the realization of the random device, as we reward those who managed to have the correct bets.

In the treatment with control, subjects can specify the sequence they prefer. In the treatment without control, we draw a sequence for them, for instance *Odd, Odd, Even, Even, Odd*. In those treatments without control, we do not specify how the sequence was chosen. In practice, whenever we gave subjects a sequence, we used an equiprobable distribution on possible choices.

## 2.2 Criteria: Procedures Without an Explicit Random Device

The following three procedures do not exhibit explicit randomness in the allocation. While we argue that all those procedures are in practice largely unpredictable, their results do not involve any formal lottery.

### 2.2.1 A Decentralized Ritual of Reason: Rock-Paper-Scissors ("RPS")

The objective of this procedure is to respect Elster (1989)'s rituals of reason, while still being largely random in practice. Each subject has five actions, each one taken from the set {Rock, Paper, Scissors}. In the treatment without control, we provide the subjects with a sequence of actions. In the treatment with control, we let them choose their actions. We play the actions of each subject against all other subjects in their treatment. As in the traditional game, Rock wins against Scissors, Scissors against Paper, and Paper against Rock. The subjects wins if they win more rounds than their opponent. In case of a tie, the first winner of a round wins the game. In the rare event where both players have chosen exactly the same strategy, we consider it as neither a win nor a loss. We then rank all participants by their number of wins. The main difference with a traditional RPS game is that all strategies are chosen in advance and a subject uses the same strategies against all other subjects.

The game can be perceived as meritocratic, at least in the version with control. It is possible for instance to exploit a well-known bias of RPS: too many players choose Rock (44.0% in our sample), and not enough Scissors (23.5%), in particular in the first round, so that choosing Paper (chosen by 32.5% of subjects) gives a higher probability of winning. However, our subjects do not exploit those biases. The pattern is similar to what is observed in the real world. For instance, human players on the website https://roshambo.me[11] choose Rock 35.3% of the time in the first round, and Scissors 29.1% of the time (see Online Appendix A for details). The difference in magnitude is likely to arise because players on the website chose to play a game of RPS, contrary to our subjects. They are therefore more likely to play regularly and have a notion of the best strategies, at least for some of them.

### 2.2.2 A Centralized Ritual of Reason: Guessing the Paintings ("Paintings")

This procedure aims at being unpredictable and respecting the rituals of reason, while adding a specific role for the experimenter. Indeed, a particularity of RPS is that besides setting the rules, the experimenter has no role in the allocation. In many of the examples described in the introduction, such as the evaluation of research proposals, the unpredictable character of the results is however precisely due to the heterogeneity of tastes of those in charge of making the selection.

In this treatment, we explain to the subjects that one of the two experimenters, Renaud, has chosen 5 pairs of paintings, each pair by the same artist, and that the second experimenter, Elias, has chosen in each pair his favourite painting. We display our names and pictures together with the pairs of paintings. To win, subjects have to guess the paintings chosen by Elias. We rank participants according to the number of paintings they have guessed correctly. The first half gets the reward. In case of tie, we use the first pair of paintings, then the second and so on. The actions are thus a sequence of five paintings. In the treatment without control, we provide the subjects with a selection of paintings. In the treatment with control, we let them choose their actions.

Subjects could download a password protected PDF copy of Elias' choices and the password is revealed to them the next day alongside the results of the experiments. We provide this document for transparency, but also to clearly separate the effect of being the subject of the tastes of an experimenter from a possible effect of reciprocity or betrayal aversion had the choices of Elias been made after the experiment took place.

### 2.2.3 An Arbitrary Procedure: Time

Many of the real world algorithms used in allocation mechanisms are complicated and sometimes opaque. For instance, in France, the algorithm allocating students to university between 2009 and 2017 was simply not public. Citizens could find out it was a modified deferred acceptance mechanism but were not provided any information on the details.[12] The current procedure features a transparent central algorithm, but the algorithms used by universities to rank students are not disclosed. Simply put, citizens have a general

[11] We thank Lasse Hassing for giving us access to their data.

[12] In 2017, the French public authority overseeing data privacy law (CNIL) summoned the French government to explain to participants to the mechanism the precise algorithm used to rank them (decision 201-053, 30 August 2017). See also the description of the algorithm of Admission Post Bac (APB) provided by matching in practice, https://www.matching-in-practice.eu/university-admission-practices-france/.

idea of what the mechanism aims to achieve and are aware that the choice is made by algorithms, but know little more than that. While, for ethical reasons, we chose to be transparent on the procedure used, we thought it could be interesting to see how subjects react to a completely arbitrary, complicated, and intractable procedure.

In the treatment without control, we provide subjects with a code corresponding to the last five digits of a time, in hours, minutes and seconds.[13] We rank the code of all subjects using an algorithm based on the number of odd digits. The full algorithm is detailed in the Online Appendix D.1. The default is that the longest part of the procedure is hidden, but subjects can click on a button to view the rest, similar to the terms and condition of use of many online services. In the treatment with control, we asked subjects to choose their time, which is then transformed into their five digits code. In both cases, the procedure is completely unpredictable but does not use any formal random device. It is totally new to subjects and difficult to interpret as meritocratic.

### 2.3 Experimental Sample and Main Differences

Our 1,324 subject were randomly allocated to 12 different treatments, summarized in Table 1. For each of our three criteria, we have a version with and without control over the actions. And for each of these six treatments, we have two versions: one in which the alternative is a lottery with control, and one in which it is a lottery without control. We aim at measuring not only the absolute preference for the different criteria, but also the impact of giving subjects control over each of them. In term of demographic characteristics, our subjects report to be 61% male, 79% USA residents, and 75% employed.[14]

Our criteria differ from the lottery in three different ways. First, by definition, by the absence of an explicitly random device. Second, criteria may be perceived as more uncertain. While the proportion of high rewards is common knowledge and the same for all procedures, and while the results are unpredictable, the probability of winning for each individual is only clearly defined as 50% for the lottery, while it is not explicit for the other procedures and therefore ambiguous. Third, the lottery is the most transparent of our procedures, as it is operated by a third party. If we assume most subjects are ambiguity averse and value transparency, we would expect any aversion for lotteries we identify in this experiment to constitute a lower bound estimate.

## 3 Results

### 3.1 Rituals of Reason

We provide general summary statistics of the different choices and beliefs in Table 2. Most subjects vote in favour of the criteria over the lottery for the two treatments following the rituals of reason, but they dislike the criteria when the procedure is completely arbitrary. Subjects choose RPS and Paintings 61.0% of the time, while they choose Time 43.4% of the time.[15] This result relates to the literature on outcome biases and meritocracy:

[13]In practice, it was the time at which they started the experiment in Washington D.C.'s time zone, but we did not specify it in order to keep the framing as neutral and vague as possible.

[14]The proportion do not vary significantly between different treatments, except for the proportion of USA residents between the Time and RPS treatments, where the p-value of the Fisher test is $< 0.001$. We do not believe it influences our results, as the regression Table 7 in Section 3.4 shows.

[15]The p-values of the one sample two-sided t-test of equality to 50% are $< 0.001$ and 0.008 respectively.

Table 1: Size of the sample for each treatment.

| | | | Lottery | |
|---|---|---|---|---|
| | | **Control?** | Yes | No |
| **Criteria** | RPS | Yes | 99 | 98 |
| | | No | 140 | 152 |
| | Paintings | Yes | 89 | 97 |
| | | No | 114 | 123 |
| | Time | Yes | 113 | 102 |
| | | No | 91 | 106 |

Table 2: Percentage of the each criteria being chosen and beliefs.

| | **Criteria Chosen** | **Criteria Believed Better** | **Criteria Believed $\geq 1/2$** | **Lottery Believed $\geq 1/2$** |
|---|---|---|---|---|
| RPS | 61.8% | 48.5% | 60.7% | 59.5% |
| Paintings | 59.8% | 49.4% | 65.5% | 58.6% |
| Time | 43.4% | 54.6% | 58.3% | 55.8% |
| Aggregate | 55.4% | 50.7% | 61.5% | 58.1% |

subjects like to interpret their successes as the result of effort or talent, but not of luck. Hence, they value a procedure that gives them a simple way to do so. The difference in the choice of the criteria is not significant between RPS and Paintings, with a p-value of the Fisher test of equal proportions of 0.59. We thus cannot identify any difference due to the role of the experimenters.

Looking at the second column, we get a clear insight that these choices correspond to actual preferences and cannot be explained by overconfidence. Indeed, in our incentivized measure of beliefs, the share of subjects who expect to perform better in a criteria compared to the lottery is not significantly different from 50% (which is, on aggregate, the correct estimate), while the share of people actually voting for a criterion to be used is significantly above.[16] We are not measuring the overconfidence in the general performance of individuals, but relative overconfidence in one procedure over the other. In the third and fourth column, we see that subjects are overconfident in all procedures, but overconfidence is not higher in any of them, with the exception of Time.[17] Overconfidence would however lead our subjects to select Time more often, while it is in practice the least chosen procedure.

In the pilot experiment (see Appendix A), we used two different belief elicitation methods asking for more precise probabilities, and reached similar results.[18] Importantly,

[16]The p-value of the one sample two-sided t-test of subjects saying that they are more likely to win with the criteria than 50% is 0.62. The p-value of the two sample two-sided t-test of equality of the means between picking a criteria and saying criteria are better is 0.014.

[17]The minimal p-value of the Fisher test of believing it is more likely to win in one procedure compared to the other is 0.076, when comparing Time and RPS.

[18]A possible weakness of the elicitation method of the main experiment is that it does not satisfy the No Complementarity at the Top condition for incentive compatibility given by Azrieli et al. (2018): Subjects could conceivably hedge between getting the reward in the allocation mechanism and getting the

Table 3: Correlation between choices, beliefs and performance.

| | **Criteria Really Better Believed Better** | **Lottery Rank $\geq 1/2$ Believed $\geq 1/2$** | **Criteria Rank $\geq 1/2$ Believed $\geq 1/2$** |
|---|---|---|---|
| RPS | -0.01 | -0.00 | 0.01 |
| Paintings | -0.00 | -0.01 | 0.03 |
| Time | -0.03 | 0.00 | -0.01 |
| Aggregate | -0.01 | -0.01 | -0.00 |

Table 4: Beliefs and choices of the criteria.

| | **Beliefs Criteria $\geq 1/2$ Lottery $\leq 1/2$** | **Beliefs Criteria $\leq 1/2$ Lottery $\geq 1/2$** | **P-value** |
|---|---|---|---|
| RPS | 79.1% | 46.2% | <0.001 |
| Paintings | 67.0% | 43.7% | 0.004 |
| Time | 50.6% | 45.3% | 0.614 |
| Aggregate | 66.8% | 45.2% | <0.001 |

no subject reported exactly a 50% probability, so if some subjects are indifferent (or randomize) between two options, it is no due to them believing their expected payoffs are identical.

To verify our conjecture that the results of our procedures are unpredictable, we compare whether subjects expect to win the reward in a given procedure (the last two colums of Table 2) with the actual outcome. The last two columns of Table 3 report the correlation between these beliefs and the actual performance of subjects. We find no significant correlation. While we cannot rule out that some participants correctly guessed they were going to win (for instance, by using known biases in the game of RPS, or by correctly guessing Elias' preferences on artwork), on aggregate their predictions are as good as random. The first column looks at the beliefs compared to the relative performance. Again, we could not find any correlation between subjects' beliefs to perform relatively better in a procedure and their actual performance.

Table 4 looks at the influence of beliefs on the choice of a criterion instead of the lottery. We find that subjects who expect to be among the first half of the participants (and thus to win the reward) in the criterion, but not in the lottery, are significantly more likely to vote in favour of the criterion than those who expect to win the lottery but not the criterion. Among those who either expect to win in both, or expect to lose in both mechanisms, we look at the procedure in which they expected to perform better. We do not find a significantly higher share of subjects voting in favour of the criterion when expecting to perform better. While this data tell us that at least some subjects make their choice based on their expected odds, it also shows that those are far from being

payment for the belief elicitation. Our results are consistent with those found in the pilot, hedging does not seem to have played a major role in our experimental belief elicitation. Additionnally, as subjects did not know that a belief elicitation was coming, the only possible hedge happened at that stage and therefore cannot jeopardize the main results.

the only factor. Among subjects expecting to win in the lottery and not the criterion, around 45% still choose a criterion, while around a third choose the lottery even if they only expect to win the criterion. Looking at the stated preferences subject give us to explain their choice (see Appendix B), 24% claim to have taken their decision based on expected odds only and 28% mention the odds as at least one of the reasons for their choices.

This result may also reflect a lack of confidence of our subjects about their beliefs. After all, those beliefs are uncorrelated with their actual performance, so that it makes perfect sense to put little weight on them. The slightly different design of our pilot study (see Appendix A) allows us to address this point. As our subjects are playing against a random set of strategies from an existing database of RPS players, we compute and communicate their probability of winning in one treatment before asking them to choose a given procedure. Everything else equal, telling subjects that their probability of winning RPS is higher than 50% increase their probability of choosing RPS by 25 percentage points. This effect is however almost identical to the one of expecting to perform better in the absence of information on the true probability. Providing reliable information on the probability of winning only sorted the subjects, with the best performing ones more likely to choose the RPS, but does not affect the aggregate share of subjects choosing the criterion.

It is also striking that the arbitrary criterion Time is the only one for which beliefs have no measurable influence on choices. This reinforces the idea that subjects do not interpret its results as meritocratic and do not see these beliefs as meaningful.

## 3.2 The Role of Control

Following the literature on the control premium and the intrinsic value of decision rights (Owens et al., 2014; Bartling et al., 2014; Bobadilla-Suarez et al., 2017; Ferreira et al., 2020), we conjecture that subjects are more likely to vote for a procedure on which they have some control. Table 5 shows the share of subjects voting for a criterion in each of our 12 treatments. The two first columns correspond to the treatments in which subjects have control over the criterion. The second and fourth columns correspond to the treatments in which subjects have control over the lottery. The results show that control indeed matters for the lottery, as on average subjects choose the lottery with control 50% of the time, compared to 39% without.[19] It is not the case in general for the criteria, with a p-value of the Fisher test of 0.27.

When we separate the criteria, control does not matter for the arbitrary criterion Time (p-value of 0.98) but matters for the criteria following the rituals of reason with an increase in choice of these criteria by 7% (p-value of the Fisher test of 0.047). Overall, control matters more for the lottery than for these two criteria,[20] with an 11% increase in the choice of the lottery compared to a 7% increase for the criteria, when given control. It is thus sufficient to give subjects control over an explicitly random procedure to make it relatively more attractive.

[19] The p-value of the Fisher Exact test of equality of the proportion is $< 0.001$.

[20] A possible explanation for this higher preference for control in lotteries is that subjects may misunderstand the probability of some sequences in a lottery, such as five Even for instance. We however show in Online Appendix C that the actual sequence has no influence on the choice of procedure in the treatments without control over the Lottery.

Table 5: The influence of control.

| Control | **Criteria** | **Yes** | **Yes** | **No** | **No** |
|---|---|---|---|---|---|
| | **Lottery** | **No** | **Yes** | **No** | **Yes** |
| Criteria | RPS | 74% | 59% | 65% | 51% |
| | Paintings | 66% | 60% | 63% | 52% |
| | Time | 46% | 42% | 48% | 37% |
| | Aggregate | 62% | 52% | 60% | 48% |

Table 6: Influence of aversion to ambiguity

| | **Ambiguity Averse/Neutral** | **Ambiguity Loving** | **P-value** |
|---|---|---|---|
| RPS | 64.0% | 55.0% | 0.101 |
| Paintings | 61.0% | 56.0% | 0.439 |
| Time | 43.3% | 43.8% | 1.0 |
| Aggregate | 56.7% | 51.7% | 0.134 |

## 3.3 Additional results

### 3.3.1 Risk and Uncertainty

While all our treatments offer the same proportion of winners, and even if, on aggregate, we do not observe relatively more confidence for any of our treatments, the nature of uncertainty is different. Our lottery is a classic example of risk (assuming subjects understood the probabilities correctly). Regardless of the sequence chosen, all participants have a known probability of winning of 50%. For all the other treatments, subjects know that half of them would receive the payoff, but may be unable to tell what is their probability of winning.

In the context of our experiment, it implies that any subject displaying ambiguity aversion should, all other things held equal, prefer the lottery. We compare in Table 6 our results with an incentivized measure of ambiguity aversion. We use the classical Ellsberg two urns choice: subjects choose between a risky urn with 50 red balls and 50 black balls and an urn with an unknown composition of red and black balls. They also choose the winning colour. Ambiguity attitudes have no significant influence on the observed choices. If anything, any effect is in the direction of subjects choosing more often the criteria when they are not ambiguity loving.

### 3.3.2 Time Spent on Different Procedures

There is a large variation in the time our subjects spent in the experiment. These differences may be particularly important in our context if they influence the choice of procedure. We believe this can happen for two reasons. First, subjects may make a cost-benefit analysis and decide to avoid spending time on uncommon procedures, leading them to vote for the lottery due to a lack of information over the criterion. This can be the case in particular for our arbitrary algorithm Time, as it is clearly unconventional. The same reasoning may also hold for the Paintings procedure, as it may take some time to look at the pictures and think about which one Elias might prefer. Second, if there is a difference in the time spent in the different procedures, our subjects may develop a preference for those procedures in which they have invested more time. Of course, as it is an online experiment, the measured time may be unreliable, some subjects may be distracted, or decide to leave their computer to come back to the task later. There is no reason for this to happen more often in one task or another, so that it should not impact the comparison between treatments, but only create outliers with long time spent.

Figure 1 shows boxplots of the time, in seconds, spent by our subjects on the page describing a given procedure and choosing actions if required. It shows that those two concerns cannot play an important role in our experiment. On the first point, we observe that, if anything, subjects spend more time on the unusual procedures, in particular when given control over it. On the second, the most chosen criteria (RPS) is actually the one on which, on average, subjects spend the less time, on par with the Lottery, which is also a common procedure. In the absence of control, subjects spend on average the same time for each procedure, except for Paintings, perhaps because it is less usual, and because some subjects may enjoy looking at them (at least, this is what several participants reported in the comments). Albeit also very unusual, Time is judged more quickly as reading the complete procedure is an active choice of subjects and the procedure is not meaningful.

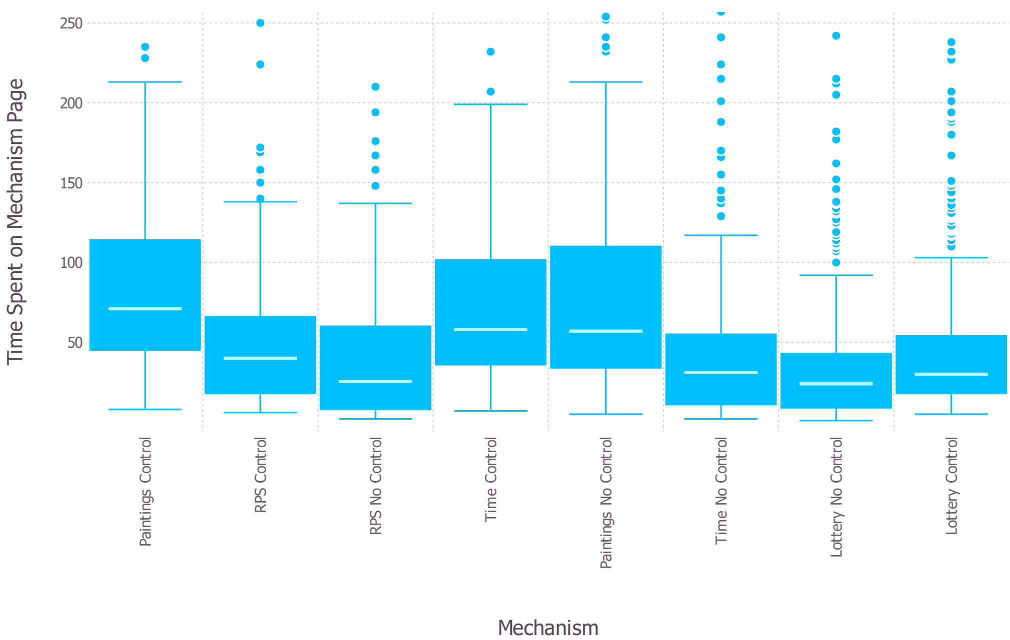


Figure 1: Time spend reading the description of a procedure and choosing actions in the control treatments. There are some outliers not shown.

#### 3.3.3 Meritocracy or Rituals of Reason?

In order to separate the role of meritocracy from the rituals of reason, we made a simple modification to our RPS procedure (with control), called *RPS loser*. When voting for the procedure (thus after tchoosing their strategies), we inform subjects that, if they choose RPS, we would reward the *bottom* half of the participants. If meritocracy is the main driver of preference for RPS, most subjects should vote against this procedure. Moreover, in the classification of Eliaz and Rubinstein (2014), RPS loser does not satisfy the property of divine providence: we expect our subjects to have chosen RPS strategies in order to win, and we inform them that only the worst-performing will be rewarded.

We find that 73% of subjects chose RPS loser when they have no control over the lottery, while 59% chose it when they have control over the lottery. These results are virtually indistinguishable from those obtained with RPS winner and control over the criteria in Table 2. In fact the p-values of both Fisher tests are close to 1. One possible explanation is that subjects value in the RPS criterion the ritual of playing it and not so much its interpretation as a meritocratic procedure. We cannot rule out, however, that our allocation rule, albeit displayed prominently in the experiment (see Online Appendix D), is so unusual and unexpected that our participants simply discarded it.

### 3.4 Regression Analysis

Finally, we gather all our variables together in a regression. We estimate in Table 7 two linear models to explain the share of subjects choosing a non-random criterion over a lottery.[21] As a benchmark, we look at the case in which the highest share of the subjects choose a criterion: when the criterion is RPS with control, and when they have no control over the lottery. We then look at the marginal impact of different factors. We have additional control variables : age, gender, attitude towards ambiguity, being a USA resident and being employed. Regression (1) looks at the main explanations identified above and keeps only significant ones. Regression (2) adds insignificant variables control variables. Additional regressions with interaction terms are available in Online Appendix B. Regressions with stated preferences are available in Appendix B.

We find that the two criteria following the rituals of reason are indistinguishable from each other, whereas the arbitrary one (Time) is less chosen (around 18 percentage points). Control matters more for the lottery than for the criteria. It increases the choice of the lottery by around 10 percentage point, and the absence of control over the criteria by 5 percentage points. For some subjects, the belief that they will win in one of the procedure and not the other drives their choices. It represents a fraction of around 20 percent of the sample. There is however no correlation between the expected relative ranking in the measures and the choice of procedure. Finally, gender matters only for the Time procedure: female are much less likely to choose Time than male, for a reason that is unclear to us.[22] It is the only demographic variable that matters. Ambiguity attitudes do not influence our results either.

## 4 The Importance of Procedures

In this section, we combine our results with those of previous experiments to show that there is a robust case for considering changes in procedures and control in order to make random (explicitly or not) tie-breaking rules socially acceptable.

[21] Non-linear estimations are in Online Appendix B.2 as a robustness check. We find the same results.
[22] We discuss this point in more details in Online Appendix C.2.

Table 7: Selection of the criteria.

| | Criteria Chosen | |
|---|---|---|
| | (1) | (2) |
| (Intercept) | 0.656**** | 0.704**** |
| | (0.044) | (0.055) |
| Time | -0.281**** | -0.277**** |
| | (0.047) | (0.048) |
| Paintings | -0.028 | -0.023 |
| | (0.032) | (0.032) |
| Control on the Lottery | -0.103**** | -0.101**** |
| | (0.027) | (0.027) |
| No Control on the Criteria | -0.053** | -0.051* |
| | (0.027) | (0.027) |
| Criteria $\geq 1/2$, Lottery $\leq 1/2$ | 0.118**** | 0.119**** |
| | (0.032) | (0.032) |
| Criteria $\leq 1/2$, Lottery $\geq 1/2$ | -0.091** | -0.092** |
| | (0.036) | (0.036) |
| Criteria $\succeq$ Lottery | | -0.004 |
| | | (0.027) |
| Male | 0.007 | 0.003 |
| | (0.033) | (0.033) |
| Time & Male | 0.150** | 0.154*** |
| | (0.059) | (0.059) |
| Ambiguity Averse/Neutral | | NS[a] |
| Age Categories[b] | | NS |
| USA Residents | | NS |
| Student[c] | | * |
| Employment Status[d] | | NS |
| Estimator | OLS | OLS |
| $N$ | 1,324 | 1,324 |
| Adjusted $R^2$ | 0.062 | 0.062 |

P-values: *: $< 0.1$, **: $< 0.05$, ***: $< 0.01$, ****: $< 0.001$

[a] NS: Not Significant.

[b] Age categories in the sample: $< 25$, 25-40, 40-55, $> 55$

[c] Being student is significant at the 10% level, but the sample is very small, as they are only 17 of them in the whole experiment. We would not give any weight to this result.

[d] Outside of student status: retired, employed, self-employed or other.

## 4.1 Procedures in this Paper

The aim of this subsection is to look at how important the effects measured in this paper are, and how likely they are to matter when the social applicability of a mechanism relies on majority voting. Hence, we want to provide back-of-the envelope calculations of the share of the population susceptible of modifying their choice when the procedure, or the level of control over the procedure, changes. An important caveat with this exercise is that we do not have representative samples, and therefore our results should be taken as indicative of the direction and approximate magnitudes, but not as exact measures of the effect in the general population, as shown by Snowberg and Yariv (2021). To do so, we compare our results with previous studies looking at either the role of procedures, or the role of control. While those studies differ widely in terms of methodology, they have two things in common. First, they allow us to compare procedures that are ex-ante payoff equivalent. Second, they give us information about the share of subjects susceptible of changing their vote in a process similar to the one described in this paper.

In our study, the smallest share of subjects voting in favour of the lottery across treatments is equal to 26%. The smallest share voting in favour of a criterion is 37%. We therefore estimate that 37% of our subjects could switch their choice between a random and a non-random mechanism depending on which non-random procedure is available to them, and on whether or not they have control over it (and on the lottery). Formally, define the *volatility* of a sample as follows: for any given set of procedures available in the experiment, if we replaced one of these procedure by another one, what share of our subjects would, at most, change their decision?

In our case, we find this volatility of 37% by starting from the treatment {{Lottery, Control}; {Time, No Control}}, and replacing the second element by {RPS, Control}. In the first treatment, 37% of our subjects choose the criterion, in the second, 74% do, a difference of 37%. We can do the same exercise by allowing for a change of procedure only (maximal difference in one column in Table 5), or a change of control only (maximal difference in one row in Table 5). In the first case, we find 28%, when there is control over the criteria and no control over the lottery, by switching the criteria from RPS to Time (first column of Table 5). In the second case we find a difference of 23%, by starting from {{RPS Control},{Lottery, No Control}} and reversing the structure of control (first row of Table 5). This definition is thus about the most that can be achieved by changing procedures, and does not discuss the fact that, for instance, the impact of control is different on different procedures.

## 4.2 Procedures in Related Papers

We summarize in Table 8 similar metrics from other studies. The second line correspond to our pilot study (see Appendix A), in which all the procedures were variations of RPS, with different framing and information on the objective probability of winning. This explains why our measure of volatility is smaller. Albeit the two experiments are different, the measure of the effect of control in the pilot is similar to the measure in the main experiment.

The third line corresponds to the survey of Eliaz and Rubinstein (2014) on preferences among different types of random procedures. They define as *emotional* subjects who exhibit a systematic ranking of procedures on the basis of their perceived fairness. They find that 30% of their subjects do have such rankings. In our setting, this corresponds to

Table 8: A typology of preferences, estimates from the literature

| | **Volatily** | **Procedure** | **Control** |
|---|---|---|---|
| Our paper | 37% | 28% | 23% |
| Pilot study | 24% | 7% | 24% |
| Eliaz and Rubinstein (2014) | 30% | 30% | |
| Bartling et al. (2014) | 83% | | 83% |
| Owens et al. (2014) | 31% | | 31% |
| Bobadilla-Suarez et al. (2017) | 44% | | 44% |
| Keren and Teigen (2010) | 18% | 18% | |

the subjects who could change their choice between lotteries and criteria depending on the framing of the former, and are therefore volatile.

Bartling et al. (2014) find that 83% of the subjects display a preference for control and 17 percent the opposite. Hence, all other things held equal, providing control over one procedure (but not the other) could switch the preferences of up to 83% of the subjects. In their experiment, control matters arguably more, as a principal has control over both her and the agent's strategy.

Owens et al. (2014) find a smaller role for control, with 31 percent of the subjects displaying a preference for control, and 18 percent the opposite. In contrast with Bartling et al. (2014), control only affects the subject's payoff. In both of these experiments, there is a lot of scope for a meritocratic interpretation of the procedures, as well as rituals of reason, in the sense that subjects have to accomplish a number of meaningful tasks.

Bobadilla-Suarez et al. (2017) find that more than 44% of the subjects reveal a preference for control by failing to delegate a task to a (virtual) advisor when it would have been beneficial to do so.

Keren and Teigen (2010) compare the stated willingness of subjects to use a lottery in an allocation problem involving other people. They find that, depending on the problem stated, between 17.2% and 35.6% of the subject opt for a coin toss (always without control), a difference of more than 18 percentage points. It should be noted however that these do not simply correspond to different framing, but also to actually different hypothetical situations (albeit all involving a life-or-death decision designed to be in a situation of equipoise).

# 5 Discussion and Conclusion

We start this section by providing three recommendations on the implementation of random tie breakers based on the results of our experiment and the previous literature. They are aimed at any designer willing to implement a random (explicitly or not) tie-breaking rule. These recommendations start from a normative standpoint: the assumption that such a tie-breaking rule is desirable. In a last subsection, we conclude by discussing limitations of the paper and avenues for future research.

## 5.1 A Random Procedure Does Not Have To Be Expensive

Among the examples of reluctance to implement lotteries detailed in the introduction, the most familiar to economists would probably be the allocation of research funds through grant applications. This procedure often consists in a ranking of projects, some deemed as clearly below the bar, some clearly above, and a share for which the ranking is not obvious. Those ties are typically broken after further meetings and debates among panel members. While tie-breaking by further peer review is largely unpredictable in practice (Cole et al., 1981; Graves et al., 2011; Pier et al., 2018), it corresponds to a ritual of reason in which subjects have control over their submitted proposals. In 2020, the Alexander von Humboldt foundation, in Germany, explained that, at the end of the first round of evaluations, some 33 applications were in that situation, deemed strong enough to be considered for funding, but not to as straightforward as the best ones. Due to the pandemic, the foundation could not organise further meetings to break the ties and chose to use a lottery instead, acknowledging that near the cut-off a random tie-breaking rule is "as just" as peer review. Aware that it might be controversial, they however decided not to publicize the lottery component, at least not directly to the recipients.[23]

In that example, the main problem is thus that the cost of keeping the rituals of reason was too high: there was no time for a further meeting. There was also a general agreement that the tie-breaking procedure was a good as random, as well as an understanding that explicitly random procedures may be unpopular. Based on the results of our experiment, we expect rituals of reasons to be preferred by a large share of the subjects, and would expect a social reluctance to implement or publicize an actual lottery. As the main problem with the procedure of writing and reviewing grant applications seems to be the opportunity cost of doing so, it may be more sensible to argue in favour of reducing this cost instead of implementing an explicit randomization device. Once everyone acknowledge that the procedure is largely random with the exception of the very best and the very worse proposals, it could make sense to drastically reduce the length of applications and the time allocated for reviewing them. This would make the procedure as close as possible to a tie-breaking lottery both in terms of opportunity cost and in terms of outcomes, while keeping the rituals of reason and leaving control to the applicants.

## 5.2 Unpredictable Rituals of Reasons as a Substitute for Lotteries

A second message from our experiment is that, when an optimal mechanism includes a random tie-breaker, there may be a more socially acceptable and payoff-equivalent alternative that follows the rituals of reason. In the United States, as in the United Kingdom and the Republic of Ireland, the high school marking standard places students in broad categories, usually in letter form. One of the reasons for such a standard is that it allows conveying information about student achievement without putting the burden on markers to provide an impossible level of precision (Schneider and Hutt, 2014). While providing more precision may be seen as unfair or arbitrary in the context of high school results, it may actually prove useful to avoid requiring an additional tie-breaking rule for the mechanisms using these results. In other words, if the difference between a 787/1000 and a 788/1000 is not meaningful and does not tell us anything about the relative abilities of two students, it nonetheless provides a criterion based on a ritual of reason on which students had some control.

[23]Bisson, Robin, "Covid-19 forces German funder to award fellowships by lottery," Research Professional News, 11 August 2020.

Without increasing the burden put on markers, simple tweaks can reduce the probability of ties. In the allocation mechanism of high school students to university in the Republic of Ireland, students are allocated a random number to break ties among those who have an identical CAO score, a metric composed of the sum, over 6 subjects, of their leaving certificate, composed of letter grades translated into marks out of 100.[24] Following complaints that random tie-breaking procedures were unfair and too frequent, the Irish government revised that metrics in 2015 with as objective to reduce the number of ties. Before the reform, possible marks on a given subject where all multiple of 5. After the reform, the possible marks are $\{0,37,46,56,66,77,88,100\}$, so that summing them up leads to much fewer ties. The government does not hide that those numbers have been chosen with the main objective of reducing ties, and have no particular meaning.[25] Other more time-consuming approaches to decrease the number of ties are the use of cover letters (in France), interviews (in the UK), or holistic applications (in the US), albeit those typically also have other objectives than breaking ties.

## 5.3 Hiding Lotteries Does Not Help, Giving Control May

If there is no obvious way to tweak criteria satisfying the rituals of reason and avoid breaking ties using an explicitly random device, our experiment give us two properties of lotteries that could make them more acceptable. First, in line with Eliaz and Rubinstein (2014) subjects need to understand what is happening, the procedure must be familiar. As shown with our arbitrary procedure, subjects do not prefer an obscure algorithm over a clear lottery. Second, it would help to let people take part to the process and have some control over their choices.

On the first point, the Washington DC "school lottery" is a prime example of a mechanism that publicly embraces the concept of lottery. It is a deferred acceptance mechanism using an individual randomly generated number as a tie-breaker (Abdulkadiroğlu et al., 2017). However, instead of trying to hide the random tie-breaker, the entire procedure goes by the name of "lottery." On the second point, the procedure leading to an actual random number being assigned to each parent is often strikingly non-transparent. For the DC lottery, as well as for the NYC and Denver deferred acceptance mechanisms, official websites simply inform parents that the algorithm generates a random number, but there is no way to actually check how the number is generated, and certainly no involvement of the subjects in the random procedure. We are not aware of any real world mechanism making the random tie-breaking rule transparent and offering some form of control to participants. Our results however suggest that doing so may increase the social acceptability of such lotteries.

## 5.4 Limitations and Future Research

We see two main limitations to our approach. The first is inherent to every incentivized study aiming at understanding major choices in life: the experimental stakes we can offer will never be sufficiently high to mimic the real life incentives. We cannot think of any experimental reward one could reasonably offer that would approach the importance for a parent of putting their child in their preferred school, or for a researcher to get a major grant. We however believe that our incentivized approach offers a step in the right

[24] Central Application Office, "Random number - How it works" on cao.ie, retrieved August, 2, 2021.

[25] Central Application Office, The New Common Point Scale, on cao.ie, retrieved August, 2, 2021.

direction, allowing to better understand individual choices than what surveys on stated preferences do.

The second is that our study allows identifying difference between treatments, but cannot measure precisely the prevalence of different type of preferences among our subjects. Some will have strict preferences for one procedure over the others, and some will be indifferent or have a preference for randomization among the different options. Further research could help identifying whether preference for certain procedures correspond to individual types, stable across different environments. It would also be interesting to see if such preferences are linked to cultural dimensions.[26] Finally, while making our subjects vote on the choice of procedures aims at replicating democratic processes, we could learn more by adding an important feature of actual democracies: debate and consensus building. Observing how smaller groups of subjects in an physical laboratory setting manage to agree on a procedure would clearly improve our understanding of individual preferences towards lotteries.

# References


Abdellaoui, M., Baillon, A., Placido, L., Wakker, P.P., 2011. The rich domain of uncertainty: Source functions and their experimental implementation. American Economic Review 101, 695–723.

Abdulkadiroğlu, A., Agarwal, N., Pathak, P.A., 2017. The welfare effects of coordinated assignment: Evidence from the new york city high school match. American Economic Review 107, 3635–89.

Abdulkadiroglu, A., Angrist, J.D., Narita, Y., Pathak, P.A., 2019. Breaking ties: Regression discontinuity design meets market design .

Agranov, M., Ortoleva, P., 2021. Ranges of randomization. mimeo .

Aldashev, G., Kirchsteiger, G., Sebald, A., 2017. Assignment procedure biases in randomised policy experiments. The Economic Journal 127, 873–895.

Avin, S., 2019. Mavericks and lotteries. Studies in History and Philosophy of Science Part A 76, 13–23.

Azrieli, Y., Chambers, C.P., Healy, P.J., 2018. Incentives in experiments: A theoretical analysis 126, 1472–1503. doi:10.1086/698136.

Barnett, A.G., 2016. Funding by lottery: political problems and research opportunities. MBio 7, e01369–16.

Bartling, B., Fehr, E., Herz, H., 2014. The intrinsic value of decision rights. Econometrica 82, 2005–2039.

Basteck, C., Klaus, B., Kübler, D., 2021. How lotteries in school choice help to level the playing field. Games and Economic Behavior .


[26] We do not find any role for those characteristics in this study (see Online Appendix D), but our data in these dimension is limited.

Bezanson, J., Edelman, A., Karpinski, S., Shah, V.B., 2017. Julia: A fresh approach to numerical computing. SIAM review 59, 65–98. doi:10.1137/141000671.

Bobadilla-Suarez, S., Sunstein, C.R., Sharot, T., 2017. The intrinsic value of choice: The propensity to under-delegate in the face of potential gains and losses. Journal of risk and uncertainty 54, 187–202.

Bolton, G.E., Brandts, J., Ockenfels, A., 2005. Fair procedures: Evidence from games involving lotteries. The Economic Journal 115, 1054–1076.

Bouacida, E., Foucart, R., 2020a. The acceptability of lotteries in allocation problems. doi:10.17635/lancaster/researchdata/395.

Bouacida, E., Foucart, R., 2020b. The acceptability of lotteries in allocation problems: a choice-based approach. Working Paper. Lancaster University, Department of Economics.

Bouacida, E., Foucart, R., 2021. The acceptability of lotteries in allocation problems. doi:10.17635/lancaster/researchdata/500.

Brezis, E.S., 2007. Focal randomisation: An optimal mechanism for the evaluation of r&d projects. Science and Public Policy 34, 691–698.

Brownback, A., Kuhn, M.A., 2019. Understanding outcome bias. Games and Economic Behavior 117, 342–360.

Budish, E., Che, Y.K., Kojima, F., Milgrom, P., 2013. Designing random allocation mechanisms: Theory and applications. American economic review 103, 585–623.

Chen, D.L., Schonger, M., Wickens, C., 2016. otree—an open-source platform for laboratory, online, and field experiments. Journal of Behavioral and Experimental Finance 9, 88 – 97. URL: http://www.sciencedirect.com/science/article/pii/S2214635016000101, doi:https://doi.org/10.1016/j.jbef.2015.12.001.

Cole, S., Simon, G.A., et al., 1981. Chance and consensus in peer review. Science 214, 881–886.

Deaton, A., Cartwright, N., 2018. Understanding and misunderstanding randomized controlled trials. Social Science & Medicine 210, 2–21.

Donovan, J.L., de Salis, I., Toerien, M., Paramasivan, S., Hamdy, F.C., Blazeby, J.M., 2014. The intellectual challenges and emotional consequences of equipoise contributed to the fragility of recruitment in six randomized controlled trials. Journal of clinical epidemiology 67, 912–920.

Dwenger, N., Kübler, D., Weizsäcker, G., 2018. Flipping a coin: Evidence from university applications. Journal of Public Economics 167, 240–250.

Eliaz, K., Rubinstein, A., 2014. On the fairness of random procedures. Economics Letters 123, 168–170.

Elster, J., 1989. Solomonic judgements: Studies in the limitation of rationality. Cambridge University Press.

Erdil, A., Ergin, H., 2008. What's the matter with tie-breaking? improving efficiency in school choice. American Economic Review 98, 669–89.

Fang, F.C., Casadevall, A., 2016. Research funding: The case for a modified lottery. mBio 7, e00422–16.

Ferreira, J.V., Hanaki, N., Tarroux, B., 2020. On the roots of the intrinsic value of decision rights: Experimental evidence. Games and Economic Behavior 119, 110–122.

Fox, C.R., Tversky, A., 1995. Ambiguity aversion and comparative ignorance 110, 585–603. URL: https://academic.oup.com/qje/article/110/3/585/1859203, doi:10.2307/2946693. publisher: Oxford Academic.

Frank, R.H., 2016. Success and luck: Good fortune and the myth of meritocracy. Princeton University Press.

Graves, N., Barnett, A.G., Clarke, P., 2011. Funding grant proposals for scientific research: retrospective analysis of scores by members of grant review panel. Bmj 343, d4797.

Greenberg, D.S., 1998. Chance and grants. The Lancet 351, 686.

Hauser, D., Paolacci, G., Chandler, J., 2019. Common concerns with mturk as a participant pool: Evidence and solutions. .

Heath, C., Tversky, A., 1991. Preference and belief: Ambiguity and competence in choice under uncertainty 4, 5–28. URL: http://link.springer.com/10.1007/BF00057884, doi:10.1007/BF00057884.

Heckman, J.J., 2020. Randomization and Social Policy Evaluation Revisited. Technical Report.

Keren, G., Teigen, K.H., 2010. Decisions by coin toss: Inappropriate but fair. Judgment and Decision Making 5, 83.

Kesten, O., Ünver, M.U., 2015. A theory of school-choice lotteries. Theoretical Economics 10, 543–595.

Kuziemko, I., Norton, M.I., Saez, E., Stantcheva, S., 2015. How elastic are preferences for redistribution? evidence from randomized survey experiments. American Economic Review 105, 1478–1508.

Langer, E.J., 1975. The illusion of control. Journal of personality and social psychology 32, 311.

Lilford, R.J., Jackson, J., 1995. Equipoise and the ethics of randomization. Journal of the Royal Society of Medicine 88, 552.

Loewenstein, G., Issacharoff, S., 1994. Source dependence in the valuation of objects. Journal of Behavioral Decision Making 7, 157–168.

Oberholzer-Gee, F., Bohnet, I., Frey, B.S., 1997. Fairness and competence in democratic decisions. Public Choice 91, 89–105.

Owens, D., Grossman, Z., Fackler, R., 2014. The control premium: A preference for payoff autonomy. American Economic Journal: Microeconomics 6, 138–61.

Pathak, P.A., Rees-Jones, A., Sönmez, T., 2020. Immigration Lottery Design: Engineered and Coincidental Consequences of H-1B Reforms. Technical Report. National Bureau of Economic Research.

Pier, E.L., Brauer, M., Filut, A., Kaatz, A., Raclaw, J., Nathan, M.J., Ford, C.E., Carnes, M., 2018. Low agreement among reviewers evaluating the same nih grant applications. Proceedings of the National Academy of Sciences 115, 2952–2957.

Rand, D.G., Greene, J.D., Nowak, M.A., 2012. Spontaneous giving and calculated greed. Nature 489, 427–430.

Roumbanis, L., 2019. Peer review or lottery? a critical analysis of two different forms of decision-making mechanisms for allocation of research grants. Science, Technology, & Human Values 44, 994–1019.

Schlag, K., Tremewan, J., 2020. Simple belief elicitation: an experimental evaluation URL: https://econpapers.repec.org/paper/pramprapa/98187.htm.

Schmidt, R.J., Trautmann, S.T., 2019. Implementing (un) fair procedures? favoritism and process fairness when inequality is inevitable .

Schneider, J., Hutt, E., 2014. Making the grade: A history of the a–f marking scheme. Journal of Curriculum Studies 46, 201–224.

Sloof, R., von Siemens, F.A., 2017. Illusion of control and the pursuit of authority. Experimental Economics 20, 556–573.

Snowberg, E., Yariv, L., 2021. Testing the waters: Behavior across participant pools. American Economic Review 111, 687–719.

Vermeule, A., 2015. Rationally arbitrary decisions in administrative law. The Journal of Legal Studies 44, S475–S507.

# Appendix

# A The Pilot Experiment

The pilot experiment results are fully described in Bouacida and Foucart (2020b). We summarize here the results of the experiment *per se*.

## A.1 Experiment

We ran two experiments. In both, individuals had to choose a mechanism used to award a reward to themselves. This contrasts with the voting procedure used in the main experiment. The two mechanisms are a lottery and a RPS, both slightly different from the main experiment. In the RPS, the set of fives actions are played against one player from a database of 2,500 players from the website Roshambo. We draw the opponent with an equiprobable distribution. The subject should win more rounds than their opponent to win the RPS game. In case of a tie, the first round which is not a tie is used to determine the winner. If they have chosen the same five actions in the same order, then the subject wins if they spent an even number of seconds on the choice screen.

The lottery in experiment 1 is a simulated coin toss (without control). In experiment 2, the lottery consists of five choices of head and tail for a repeated matching penny game (MP, with control). The subject wins the MP whenever they match three or more head and tails drawn by the computer. The computer draws head or tail with equal probability. In both cases, the *ex ante* chances of winning the lottery are 50% and are independent of the chosen strategies.

The experiments were run on Amazon Mechanical Turk on April 29, 2020 (experiment 1) and May 13-14, 2020 (experiment 2), using oTree (Chen et al., 2016). A total of 197 subjects finished experiment 1 and 89 experiment 2. The reward was \$1.40 in experiment 1 and \$0.40 in experiment 2. The show-up fee was \$0.70 in experiment 1 and \$0.40 in experiment 2. The median time spent in the experiments was 3 minutes. The median hourly wage was \$25.71 in experiment 1 and \$12.20 in experiment 2.

In both experiments, we elicit the subjects' beliefs have about their probability of winning in the RPS and the MP when these probabilities were not given. In both experiments, the belief elicitation procedure is incentivized. In experiment 1, it is done by asking the number of time they expect to win against our dataset (following Schlag and Tremewan, 2020). In experiment 2, we ask them the score they expect to reach in both RPS and MP (both scores are between -5 and 5 and are thus comparable). They are paid if their belief on their score is close enough from the real score.[27] The two belief elicitation methods are not equivalent, as the score in RPS does not exactly translate into a probability of winning, because of the tie-breaking rules.

The pilot and main experiment followed the same plan. The main difference is that in one treatment we gave them feedback on their probability of winning in the procedures before their choice of procedure.

[27]The only correct belief in experiment 2 for MP is 0. In experiment 1, subjects beliefs were considered correct if they were at within 1/2500 of the correct belief. In experiment 2, if they were within 0.1 score of their real score.

### A.1.1 Treatments

We ran four different treatments. We call the first two treatments Knowledge and Ignorance. In the knowledge treatment, we tell the subjects their average probability of winning the RPS game before choosing the mechanism. That is, we tell them the percentage of RPS games of the database they win against. In the Ignorance treatment, they receive no information about the probability of winning of their five RPS actions. We did not elicit their beliefs about the probability of winning the RPS in the Knowledge treatment.

We call the other two treatments Chance and Ability. The two treatments are two different framing of the RPS game. In the Chance treatment, we tell the subjects that "Rock-paper-scissors (RPS) is an old child play originating from ancient China," whereas, in the Ability treatment, we tell them that "Rock-paper-scissors (RPS) is an old game of strategy originating from ancient China. [...] RPS is also a well-studied game in biology, psychology, and artificial intelligence. In international competitions, some players consistently outperform the others. Computer scientists have produced algorithms able to exploit the predictable behaviour of human players and win more often against them." We are not deceiving the subjects as the two descriptions are accurate depictions of the game.

The repartition of subjects in the different treatments of Experiment 1 is shown in Table 9. In experiment 2, subjects were primed to see the RPS as a game of chance and did not know their probability of winning. So experiment 2 was run using only Ignorance and Chance treatments.

Table 9: Number of subjects in each treatment in Experiment 1.

| | Ability | Chance |
|---|---|---|
| Knowledge | 50 | 51 |
| Ignorance | 55 | 41 |

### A.1.2 RPS Data

We have collected the 2,500 strategies of the opponent in the RPS game from actual games played on the website https://roshambo.me. An overview of the data set is available in Online Appendix A, as well as a comparison with the observed choices in our two experiments. The average probability of winning against our dataset is always between 44.80% and 55.00%. The only information about their chances of winning in the RPS game – in the Knowledge treatment – is the probability of winning against our dataset. Moreover, no (pure) strategy gives an exact probability of winning of 50%.[28] As a consequence, a player valuing expected monetary payoff only should never be indifferent in the knowledge treatment.

[28]The Nash equilibrium mixed strategy gives an ex-ante probability of winning of 50%. However, in the Knowledge treatment, subjects receive the probability of winning of the strategies chosen after any randomization by the subject, and this probability is always different from 50%.

## A.2 Results

Most subjects choose the RPS in experiment 1 (66% overall treatments), whereas most of them choose the MP in experiment 2 (58%). The proportion of subjects choosing the RPS is significantly higher in experiment 1.[29] The proportion of subjects choosing the RPS is significantly different from 50% in experiment 1 but not in experiment 2.[30] In the next subsections, we explore possible explanations for these differences.

### A.2.1 Experiment 1

Table 10 shows that the different treatments do not seem to influence the choice of the allocation mechanism. What this similarity implies is that the observed preference for RPS cannot be explained by overconfidence. Overall, subjects in the Knowledge treatment are equally likely to choose it than those in the Ignorance treatment. There is, however, a role for ambiguity aversion. We find that – as expected – ambiguity averse subjects are less likely to choose RPS than the others.

Table 10: Proportion of subjects choosing the RPS mechanism, depending on the treatment.

| | Ability | Chance | Overall |
|---|---|---|---|
| Knowledge | 64% | 71% | 67% |
| Ignorance | 66% | 65% | 66% |
| Overall | 65% | 68% | 66% |

Note: the proportions are not significantly different between the different treatment cells (using a Fisher exact test).

We now look at the probability of winning the RPS, comparing the treatment where it is known by the subject (Knowledge) and those where it is not known (Ignorance). In expected monetary payoff, the lottery is more valuable if the probability of winning the RPS is below 50%, and less valuable otherwise. When we introduce the 50% probability of winning the RPS threshold, it matters in the Knowledge treatment. In the Ignorance treatment, however, what should influence the subjects' choice is their belief that they are above or below 50%. Table 11 shows that it is indeed the case. When subjects know their probability of winning the 50% threshold makes a significant difference. It is also true when they believe it is the case. The real probability, however, does not matter when they do not know it.

[29] The p-value of the Fisher exact test is $< 0.001$ between the two experiments. When restricting the sample of experiment 1 to the same treatments as experiment 2, the p-value is 0.023.

[30] The p-values of the one-sample two-sided t-test are $< 0.001$ and 0.14, respectively.

Table 11: Effect of knowing or believing that the chances of winning the RPS are above or below 50% in pilot experiment 1.

| | $q < 0.5$ | $q > 0.5$ | P-value[1] |
|---|---|---|---|
| Knowledge | 58% | 82% | 0.020 |
| Ignorance & Really better[2] | 63% | 70% | 0.59 |
| Ignorance & Believed better[3] | 55% | 80% | 0.014 |

[1] P-value of the Fisher exact test;
[2] Really better means that the real probability of winning the RPS is above 50%;
[3] Believed better means that subjects believe their probability of winning the RPS is above 50%.

To summarize, a first variable that seems to matter to understand the aggregate choices in experiment 1 is whether they believe or know that their probability of winning the RPS is higher or lower than 50%. However, it matters only for a fraction of subjects, as 58% choose the RPS even when they know they have a lower probability of winning it, and 18% of them choose the CT when they have a higher probability of winning the RPS.

Table 12 report the result of a regression analysis. The variable $RPS \succeq CT$ is a dummy that takes value 1 when subjects know or believe that their chance of winning in the RPS is above 50%. Out of 100 subjects, an average of 25 choose RPS instead of CT when their probability of winning the RPS is above the threshold of 1/2. While it clearly shows that the probability of winning matters, it also implies that a majority of the subjects are not deciding only on their expected monetary payoff. As a rough breakdown, the estimation tells us that around 55% of the subjects would choose RPS even if it gives odds lower than 1/2, around 25% pick the device that gives them the highest probability of winning, and 20% choose CT even if the RPS has better odds. Note that the estimation only looks at the exact 50% threshold. When looking at the influence of the distance with the 50% threshold, instead of a dummy, the explanatory power is similar.

### A.2.2 Experiment 2

We observe a similar phenomenon in experiment 2. Figure 2 shows a wide range of incorrect beliefs. In particular, only 4.5% of the subjects report the only correct belief of a 1/2 probability of winning in the MP (the horizontal line at MP Belief = 0). Regarding the choices made, it is clear that the decision of a large number of subjects is not only based on their beliefs on the expected monetary payoff of the two mechanisms. If it were the case, the choices of allocation mechanisms would be cleanly divided between the top-left and bottom-right by the red line.

The beliefs may have some influence on the choices made by subjects, as shown in Table 13. Simply being above or below the red line does not significantly influence the choice made, but being further away from this line may, as shown by the first regression. The explanatory power of these regressions is very small, however. It suggests that beliefs are not the main reason why subjects chose one mechanism over the other in experiment 2.

Table 12: Regression on the probability of choosing the RPS mechanism.

| | Full Sample | | Knowledge | | Ignorance | |
|---|---|---|---|---|---|---|
| | Mechanism | | | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| (Intercept) | 0.564 | 0.539 | 0.581 | 0.626 | 0.545 | 0.538 |
| | (0.046) | (0.081) | (0.063) | (0.074) | (0.068) | (0.089) |
| $RPS \succeq CT$ | 0.248**** | 0.258**** | 0.240*** | 0.256*** | 0.259*** | 0.260*** |
| | (0.064) | (0.063) | (0.089) | (0.087) | (0.092) | (0.093) |
| Knowledge | | 0.086 | | | | |
| | | (0.095) | | | | |
| Ability | | 0.012 | | -0.103 | | 0.013 |
| | | (0.096) | | (0.090) | | (0.096) |
| Knowledge & Ability | | -0.116 | | | | |
| | | (0.132) | | | | |
| Estimator | OLS | OLS | OLS | OLS | OLS | OLS |
| $N$ | 197 | 197 | 101 | 101 | 96 | 96 |
| Adjusted $R^2$ | 0.062 | 0.055 | 0.053 | 0.055 | 0.063 | 0.053 |

(* p<0.1, ** p<0.05, *** p<0.01, **** p<0.001)

# B Stated preferences

We have asked subjects at the end of the experiment to tell us why they chose one procedure over the other. We have encoded the comments made by the subjects in two categories: subjects justifying their choice by the preference for a procedure over the other, or those justifying their choices by a higher probability of winning in one procedure or the other. Not all comments can be classified with these two categories, and some belong to both. Overall, 23% of subjects said that preference drove their choices, and 28% said that probability drove their choices, with 4% of the sample belonging to both categories.

Table 14 shows that when subjects have a preference, they tend to prefer the criteria, in particular when the criteria follow the rituals of reason. It means that in general, stated preferences points to subjects having at least a relative preference for the criteria, compared to the lottery, in particular with RPS.

Table 15 shows that when subjects state that probability do matter for which procedure they choose, they behave acccording to their beliefs. If they believe that they are going to win in the criteria, but not in the lottery, they choose the criteria 86% of the time. On the other hand, if they believe that they will win in lottery, but not in the criteria, they choose the lottery 65% of the time. The pattern is true no matter what the criteria is. The sample is of course restricted to the 324 subjects who said that probabilty matters, so the closest to expected value maximizers we have in the data. Similar to what we have shown in Table 3, subjects who say that probability matter do not have correct beliefs. The Kendall correlation between their ranks in the criteria and their belief to be in the first half in the criteria is -0.03. The similar figure for the lottery is 0.02. In other words, subjects who say that probability matters are not brighter than the others in forecasting their real performance.

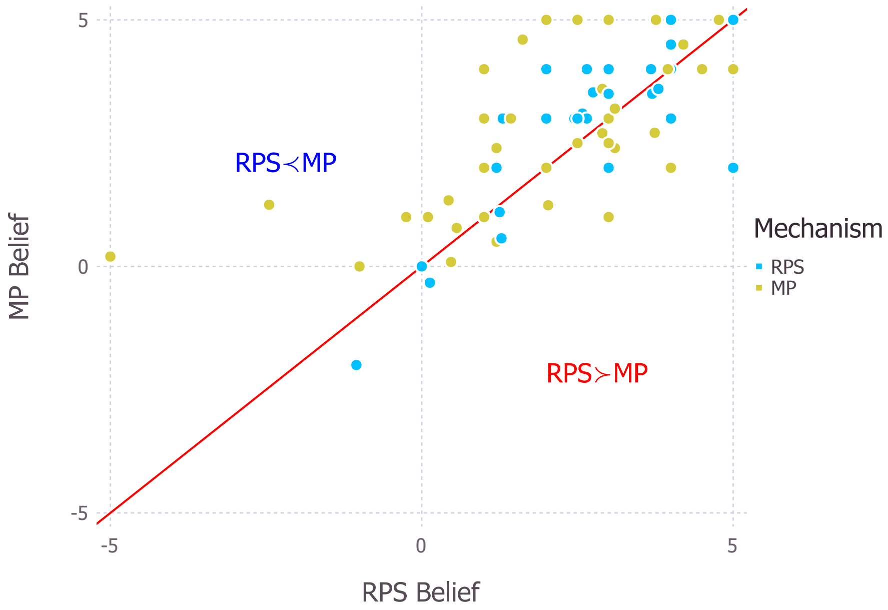


Figure 2: Beliefs and chosen allocation mechanism in Experiment 2.

Table 13: Influence of the beliefs on the choice of the allocation mechanism in pilot experiment 2.

| | Mechanism | |
|---|---|---|
| | (1) | (2) |
| (Intercept) | 0.450 | 0.413 |
| | (0.056) | (0.063) |
| Beliefs $(RPS - MP)$ | 0.058* | |
| | (0.035) | |
| Beliefs $(RPS \succeq MP)$ | | 0.049 |
| | | (0.117) |
| Estimator | OLS | OLS |
| $N$ | 89 | 89 |
| Adjusted $R^2$ | 0.010 | -0.009 |

(* p<0.1, ** p<0.05, *** p<0.01, **** p<0.001)

Table 14: Choices of criteria depending on whether preference are given as a reason for the choice.

| **Criteria** | **Preference** | **Not Preference** | **P-value** |
|---|---|---|---|
| RPS | 78.6% | 55.6% | <0.001 |
| Paintings | 64.4% | 58.3% | 0.322 |
| Time | 48.5% | 42.5% | 0.443 |
| Aggregate | 67.1% | 52.0% | <0.001 |

Table 15: Choices of criteria when subjects say that the probability matters (324 subjects), depending on the belief.

| Criteria | Believed Better Criteria | Lottery | P-value |
|---|---|---|---|
| RPS | 88.1% | 34.5% | <0.001 |
| Paintings | 81.5% | 38.9% | 0.055 |
| Time | 90.9% | 28.6% | <0.001 |
| Aggregate | 86.2% | 35.2% | <0.001 |

Finally, we can summarize the hints we got from the previous tables in a regression. Table 16 shows these regressions. Regression (1) is the regression (1) from Table 7, to provide a baseline. Regression (3) adds relevant demographic variables and interaction terms, while regression (2) removes the demographic variables. Overall, the biggest difference is the effect of beliefs. Beliefs do not matter anymore alone. They matter only for subjects who stated that probability do matter: Probability matters in interaction with beliefs, but not on their own, which is quite sensible. Stated preferences are strongly positively significant, which means that subjects who are in RPS and state a preference do state a preference for RPS on aggregate. The effect is also positive for paintings, but much smaller than for RPS. With the interaction term, we have to add the coefficients for paintings, preferences and the interaction term to assess the full effect.

To conclude this section, stated preferences go in the direction we would expect. Subjects who state they have a preference tend to have a preference for meritocratic criteria. Subjects who say that probability matters do act on their beliefs, in aggregate.

# C Demographics

## C.1 Descriptive Statistics

We have asked participants four main demographic questions: gender, country of residence, employment status and age. The questions are self-reported and not incentivized. The majority of our subjects comes from the United States (79.3%). The second highest country of residence is India, which represents 12.2% of the sample. 75.5% of the sample declare themselves to be employed. The majority are between 25 and 40 years old (64.6%). The second most represented age category are participants aged 40 to 55 (23.2%). The proportion of employed residents of the USA between 25 and 40 years old is 41.6%, which represent 551 individuals in total. The relative homogeneity of the sample does not allow us to investigate the effect of demographic characteristics on choices.

The only demographic characteristic which is sufficiently heterogeneous to allow for an investigation is gender, and we report it in the next section.

## C.2 Gender

Table 18 presents the probability of choosing RPS in the two experiments depending on the self-reported gender of the participants. 61.18% of the sample said they are male, while 38.44% said they are female. Only one subject picked the option not to do declare

Table 16: Regression analysis taking into account stated preferences and when probability matters.

| | Criteria Chosen | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| (Intercept) | 0.656**** | 0.582**** | 0.635**** |
| | (0.044) | (0.052) | (0.061) |
| Time | -0.281**** | -0.216**** | -0.210**** |
| | (0.047) | (0.053) | (0.054) |
| Paintings | -0.028 | 0.022 | 0.025 |
| | (0.032) | (0.038) | (0.038) |
| Control on the Lottery | -0.103**** | -0.087*** | -0.085*** |
| | (0.027) | (0.027) | (0.027) |
| No Control on the Criteria | -0.053** | -0.048* | -0.047* |
| | (0.027) | (0.027) | (0.027) |
| Criteria $\geq 1/2$, Lottery $\leq 1/2$ | 0.118**** | 0.056 | 0.057 |
| | (0.032) | (0.039) | (0.039) |
| Criteria $\leq 1/2$, Lottery $\geq 1/2$ | -0.091** | -0.053 | -0.052 |
| | (0.036) | (0.041) | (0.041) |
| Criteria $\succeq$ Lottery | | | -0.019 |
| | | | (0.032) |
| Stated Preference | | 0.243**** | 0.251**** |
| | | (0.060) | (0.061) |
| Paintings & Stated Preference | | -0.160** | -0.151** |
| | | (0.069) | (0.069) |
| Time & Stated Preference | | -0.162 | -0.158 |
| | | (0.141) | (0.144) |
| Probability Matters | | -0.017 | -0.031 |
| | | (0.038) | (0.047) |
| Criteria $\geq 1/2$, Lottery $\leq 1/2$ & Probability Matters | | 0.241**** | 0.241**** |
| | | (0.064) | (0.065) |
| Criteria $\leq 1/2$, Lottery $\geq 1/2$ & Probability Matters | | -0.133 | -0.144* |
| | | (0.082) | (0.083) |
| Male | 0.007 | 0.013 | 0.010 |
| | (0.033) | (0.039) | (0.039) |
| Time & Male | 0.150** | 0.141** | 0.146** |
| | (0.059) | (0.065) | (0.066) |
| Ambiguity Averse/Neutral | | | NS[a] |
| Age Categories[b] | | | NS |
| USA Residents | | | NS |
| Student[c] | | | * |
| Employment Status[d] | | | NS |
| Estimator | OLS | OLS | OLS |
| $N$ | 1,324 | 1,324 | 1,324 |
| Adjusted $R^2$ | 0.062 | 0.082 | 0.084 |

P-values: *: $< 0.1$, **: $< 0.05$, ***: $< 0.01$, ****: $< 0.001$

[a] NS: Not Significant.

[b] Age categories in the sample: $< 25$, 25-40, 40-55, $> 55$

[c] Being student is significant at the 10% level, but the sample is very small, as they are only 17 of them in the whole experiment. We would not give any weight to this result.

[d] Outside of student status: retired, employed, self-employed or other.

Table 17: Median time spent on the criteria page, depending on the control over the criteria, and p-value of the Mann-Whitney U Test

| **Control** | **Yes** | | | **No** | | |
|---|---|---|---|---|---|---|
| | Women | Men | P-Value[a] | Women | Men | P-Value[a] |
| RPS | 44.0 | 36.0 | 0.099 | 21.0 | 30.0 | 0.402 |
| Paintings | 84.0 | 65.0 | <0.001 | 52.0 | 59.0 | 0.336 |
| Time | 61.5 | 55.0 | 0.497 | 32.0 | 30.0 | 0.976 |
| Aggregate | 62.0 | 53.0 | 0.009 | 36.0 | 41.0 | 0.47 |

[a] P-value of the Approximate Mann-Whitney U-test of the sample being drawn from the same population.

Table 18: Choices of criteria by gender.

| **Criteria** | **Male** | **Female** | **P-value** |
|---|---|---|---|
| RPS | 60.9% | 62.8% | 0.748 |
| Paintings | 61.8% | 57.1% | 0.401 |
| Time | 49.6% | 33.1% | 0.001 |
| Aggregate | 57.7% | 51.9% | 0.045 |

the gender. The subject is excluded from this analysis. We observe some differences between declared genders in Table 18. Men choose the Time criteria more often than women do. We do therefore include in the regression the interaction term between time and gender (a dummy for male in this instance).

To investigate further the role of gender in our results, we look at the time spent on each criteria page. That is, the time spent reading the description of each criteria, and choosing the strategies for the criteria, if needed. The results are displayed in Table 17. Women spend significantly more time when they have control than men on the description of the criteria that follow the rituals of reasons, but not on the arbitrary one. One possible explanation for this is that women recognize relatively more than men which criteria they might have more ability to understand and influence. Consequently, they spend more time on them, while they discard criteria on which they have less agency. This explanation is consistent with the fact that they spend relatively less time than men when they have no control.

# D Culture

One missing variable likely to influence our results is culture. The acceptability of lotteries is likely to depend on the cultural background of the participant. We cannot test for this in our sample. Our use of AMT means that we have selected subjects who are more sensitive to payments than the rest of the population. It should mitigate at least some cultural effect.

As a small glimpse into the subject, we run the same analysis as in Appendix B, but on a restricted sample of USA residents between the age of 25 and 55. They represent

921 individuals, which is 69.6% of the total sample. The results (Table 19) are consistent with the results on the unrestricted sample. The direction of the effects is always the same, and the magnitude are similar and not significantly different between the restricted and unrestricted samples.

Estimates are not more precise on the restricted sample as well. If we thought that culture yield drastically different behaviours, then we would expect a more culturally homogeneous sample to have a less noisy behaviour and therefore more precise estimate. It is not the case here, as the adjusted $R^2$ is slightly smaller, and the standard deviation broadly the same. Of course, the reasoning is limited by the fact that we do not capture heterogeneity among USA residents between the age of 25 and 55. If one thought that ethnicity, for instance, influenced our results, then our restrictions do not address this issue and our sample is not more homogeneous.

In the end, our results are robust to restricting our sample to be theoretically more homogeneous among some dimensions. They should not be interpreted as culture having no impact, but as the cultural variables with used in our study having no impact.

## D.1 Data Treatment and Reproducibility

The code for the experiment as well as the treatment for the data is available in Bouacida and Foucart (2021). The dataset for the pilot is given in Bouacida and Foucart (2020a). The experiment and the pilot are built using OTree Chen et al. (2016). The data analysis is in Julia Bezanson et al. (2017), version 1.6. A readme is available with the dataset to reproduce all the computations made for the paper. Two GitHub repositories contain all the code and steps we have made t treat the data and write the experiments code. Access to them is available upon request to the authors.

Table 19: Regression analysis taking into account stated preferences and when probability matters, comparison when restricting to USA residents only, aged between 25 and 55.

| | Criteria Chosen | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| (Intercept) | 0.582**** | 0.575**** | 0.635**** | 0.603**** |
| | (0.052) | (0.064) | (0.061) | (0.068) |
| Time | -0.216**** | -0.276**** | -0.210**** | -0.283**** |
| | (0.053) | (0.061) | (0.054) | (0.061) |
| Paintings | 0.022 | -0.012 | 0.025 | -0.013 |
| | (0.038) | (0.047) | (0.038) | (0.047) |
| Control on the Lottery | -0.087*** | -0.085*** | -0.085*** | -0.083*** |
| | (0.027) | (0.032) | (0.027) | (0.032) |
| No Control on the Criteria | -0.048* | -0.014 | -0.047* | -0.014 |
| | (0.027) | (0.032) | (0.027) | (0.032) |
| Criteria $\geq 1/2$, Lottery $\leq 1/2$ | 0.056 | 0.074* | 0.057 | 0.071 |
| | (0.039) | (0.045) | (0.039) | (0.045) |
| Criteria $\leq 1/2$, Lottery $\geq 1/2$ | -0.053 | -0.017 | -0.052 | -0.018 |
| | (0.041) | (0.050) | (0.041) | (0.050) |
| Criteria $\succeq$ Lottery | | | -0.019 | -0.016 |
| | | | (0.032) | (0.038) |
| Stated Preference | 0.243**** | 0.189*** | 0.251**** | 0.187** |
| | (0.060) | (0.072) | (0.061) | (0.074) |
| Paintings & Stated Preference | -0.160** | -0.139 | -0.151** | -0.132 |
| | (0.069) | (0.088) | (0.069) | (0.088) |
| Time & Stated Preference | -0.162 | -0.140 | -0.158 | -0.122 |
| | (0.141) | (0.173) | (0.144) | (0.176) |
| Probability Matters | -0.017 | 0.000 | -0.031 | 0.005 |
| | (0.038) | (0.047) | (0.047) | (0.059) |
| Criteria $\geq 1/2$, Lottery $\leq 1/2$ & Probability Matters | 0.241**** | 0.252**** | 0.241**** | 0.250*** |
| | (0.064) | (0.076) | (0.065) | (0.078) |
| Criteria $\leq 1/2$, Lottery $\geq 1/2$ & Probability Matters | -0.133 | -0.120 | -0.144* | -0.125 |
| | (0.082) | (0.104) | (0.083) | (0.105) |
| Male | 0.013 | -0.011 | 0.010 | -0.014 |
| | (0.039) | (0.047) | (0.039) | (0.048) |
| Time & Male | 0.141** | 0.162** | 0.146** | 0.172** |
| | (0.065) | (0.075) | (0.066) | (0.076) |
| Ambiguity Averse/Neutral | NS[a] | * | NS | * |
| Age Categories[b] | | | NS | NS |
| USA Residents | | | NS | |
| Student[c] | | | ** | **** |
| Employment Status[d] | | | NS | NS |
| Estimator | OLS | OLS | OLS | OLS |
| $N$ | 1,324 | 921 | 1,324 | 921 |
| Adjusted $R^2$ | 0.082 | 0.075 | 0.084 | 0.076 |

P-values: *: $< 0.1$, **: $< 0.05$, ***: $< 0.01$, ****: $< 0.001$

[a] NS: Not Significant.

[b] Age categories in the total sample: $< 25$, 25-40, 40-55, $> 55$, in the restricted sample: $25 - 40, 40 - 55$.

[c] Being student is significant, but the sample is very small, as they are only 17 of them in the whole experiment. We would not give any weight to this result.

[d] Outside of student status: retired, employed, self-employed or other.

# Online Appendix

## A RPS Strategies and Roshambo Sample

In this Appendix, we aim to understand if the participants in our experiment who played RPS are different from the general public while playing RPS.

Table 20 shows the strategies used by a sample of 44,442 human players who participated in RPS on the Roshambo website (https://roshambo.me). Players on Roshambo choose one opponent to play with and a version of RPS. Either a one-shot version or one with at least 3 or 5 repetitions (first to 3 or 5 respectively). We restrict ourselves to participants who played at least five times against each other in the same game.

There is some differences between players on Roshambo and our players. First, they observed the result of their first action before playing the second, of the second before playing the third, and so on. It allows some well-known biases in RPS to kick in, but also corrective actions. Second, Roshambo players self-select and self-select their opponent: they went on the website to play RPS, which is not the case in our experiment. It is therefore likely that the sample has more experience and maybe more aware of the Nash strategies. We nevertheless observe in the Table 20 that they do not play the Nash strategies in aggregate.

Table 20: RPS strategies played by the Roshambo sample.

| | **Paper** | **Scissors** | **Rock** |
|---|---|---|---|
| Round 1 | 35.61% | 29.08% | 35.32% |
| Round 2 | 33.45% | 32.71% | 33.83% |
| Round 3 | 33.66% | 31.82% | 34.52% |
| Round 4 | 32.72% | 32.34% | 34.95% |
| Round 5 | 34.29% | 31.52% | 34.19% |

Finally, we can look at the best and the worst strategies against the Roshambo sample. That is, when drawing with equal probability one opponent from the data, which strategy has the highest odds. The best strategy is to play Paper five time. The chances of winning are then 53.79%. The worst strategy is to play Rock, Rock, Scissors, Scissors, Rock, its chances of winning are 46.30%. The spread between the best and the worst strategy is not very high, but there is still room to exploit human biases here.

We can compare those results to the data of RPS players in our experiment. Table 21 shows the strategies used in our experiment. 489 participants played in the various RPS treatments. The figures of the different actions played is different from the Roshambo sample. It has in particular a lot more spread, which might be explained by the smaller sample of less experienced players. The best and worst strategies are Paper, Paper, Paper, Paper, Rock and Rock five times, with respective probabilities of winning 62.47% and 39.78%. There is a lot more room in our experiment to win against our participants, but the strategies and the data exhibit a similar pattern, which suggest to us that our players were not particularly good in RPS.

Overall, the results suggest that we have not a selected sample of participants in RPS.

Table 21: RPS strategies played by the experimental sample.

| | **Paper** | **Scissors** | **Rock** |
|---|---|---|---|
| Round 1 | 32.52% | 23.52% | 43.97% |
| Round 2 | 37.63% | 28.02% | 34.36% |
| Round 3 | 32.72% | 32.72% | 34.56% |
| Round 4 | 34.56% | 31.49% | 33.95% |
| Round 5 | 33.74% | 29.65% | 36.61% |

# B Robustness of the Regressions

## B.1 Long Models

We have in Table 22 the long models of the regressions. The first regression only interacts the different characteristics of the criteria and choices, while the second adds the demographic characteristics of the participants in the experiment. Generally speaking, not much can be drawn from these regressions. There seems to be too many variable for any of them to explain the data. We therefore chose to restrict to fewer interaction terms in the core of the paper.

Table 22: Selection of the criteria.

| | Criteria Chosen | |
|---|---|---|
| | (1) | (2) |
| (Intercept) | 0.678**** | 0.734**** |
| | (0.080) | (0.119) |
| Criteria $\geq 1/2$, Lottery $\leq 1/2$ | 0.108 | -0.174 |
| | (0.103) | (0.287) |
| Criteria $\leq 1/2$, Lottery $\geq 1/2$ | -0.044 | 0.026 |
| | (0.118) | (0.164) |
| Criteria $\succeq$ Lottery | 0.044 | 0.141 |
| | (0.089) | (0.140) |
| Paintings | -0.037 | -0.101 |
| | (0.111) | (0.209) |
| Time | -0.202* | -0.280 |
| | (0.113) | (0.189) |
| Control on the Lottery | -0.205* | -0.164 |
| | (0.114) | (0.183) |
| No Control on the Criteria | -0.061 | -0.081 |
| | (0.097) | (0.157) |
| Averse to Ambiguity | 0.044 | 0.040 |
| | (0.032) | (0.032) |
| Criteria $\geq 1/2$, Lottery $\leq 1/2$ & Paintings | -0.075 | 0.366 |
| | (0.150) | (0.354) |
| Criteria $\geq 1/2$, Lottery $\leq 1/2$ & Time | 0.147 | 0.266 |
| | (0.166) | (0.362) |
| Criteria $\leq 1/2$, Lottery $\geq 1/2$ & Paintings | -0.053 | 0.228 |

Table 22: Selection of the criteria.

| | Criteria Chosen | |
|---|---|---|
| | (1) | (2) |
| | (0.185) | (0.268) |
| Criteria $\leq 1/2$, Lottery $\geq 1/2$ & Time | 0.134 | 0.045 |
| | (0.171) | (0.264) |
| Criteria $\succeq$ Lottery & Paintings | -0.061 | -0.167 |
| | (0.133) | (0.232) |
| Criteria $\succeq$ Lottery & Time | -0.209 | -0.204 |
| | (0.134) | (0.231) |
| Criteria $\geq 1/2$, Lottery $\leq 1/2$ & Control on the Lottery | 0.018 | 0.367 |
| | (0.166) | (0.342) |
| Criteria $\leq 1/2$, Lottery $\geq 1/2$ & Control on the Lottery | -0.078 | -0.094 |
| | (0.175) | (0.291) |
| Criteria $\succeq$ Lottery & Control on the Lottery | 0.114 | 0.067 |
| | (0.135) | (0.218) |
| Paintings & Control on the Lottery | 0.151 | 0.093 |
| | (0.161) | (0.296) |
| Time & Control on the Lottery | 0.057 | -0.074 |
| | (0.159) | (0.258) |
| Criteria $\geq 1/2$, Lottery $\leq 1/2$ & No Control on the Criteria | 0.084 | 0.308 |
| | (0.132) | (0.316) |
| Criteria $\leq 1/2$, Lottery $\geq 1/2$ & No Control on the Criteria | -0.142 | -0.156 |
| | (0.158) | (0.227) |
| Criteria $\succeq$ Lottery & No Control on the Criteria | -0.066 | -0.103 |
| | (0.118) | (0.185) |
| Paintings & No Control on the Criteria | 0.012 | 0.067 |
| | (0.146) | (0.261) |
| Time & No Control on the Criteria | 0.086 | 0.004 |
| | (0.158) | (0.274) |
| Control on the Lottery & No Control on the Criteria | -0.030 | -0.172 |
| | (0.147) | (0.242) |
| Criteria $\geq 1/2$, Lottery $\leq 1/2$ & Paintings & Control on the Lottery | -0.109 | -0.811* |
| | (0.238) | (0.468) |
| Criteria $\geq 1/2$, Lottery $\leq 1/2$ & Time & Control on the Lottery | -0.017 | -0.142 |
| | (0.247) | (0.455) |
| Criteria $\leq 1/2$, Lottery $\geq 1/2$ & Paintings & Control on the Lottery | -0.065 | -0.264 |
| | (0.294) | (0.466) |
| Criteria $\leq 1/2$, Lottery $\geq 1/2$ & Time & Control on the Lottery | -0.015 | -0.212 |
| | (0.249) | (0.371) |
| Criteria $\succeq$ Lottery & Paintings & Control on the Lottery | -0.078 | 0.122 |
| | (0.198) | (0.349) |

Table 22: Selection of the criteria.

| | Criteria Chosen | |
|---|---|---|
| | (1) | (2) |
| Criteria $\succeq$ Lottery & Time & Control on the Lottery | 0.080 | 0.088 |
| | (0.193) | (0.317) |
| Criteria $\geq 1/2$, Lottery $\leq 1/2$ & Paintings & No Control on the Criteria | -0.022 | -0.112 |
| | (0.200) | (0.394) |
| Criteria $\geq 1/2$, Lottery $\leq 1/2$ & Time & No Control on the Criteria | -0.413* | -0.328 |
| | (0.220) | (0.439) |
| Criteria $\leq 1/2$, Lottery $\geq 1/2$ & Paintings & No Control on the Criteria | 0.121 | -0.236 |
| | (0.248) | (0.377) |
| Criteria $\leq 1/2$, Lottery $\geq 1/2$ & Time & No Control on the Criteria | 0.073 | 0.291 |
| | (0.238) | (0.363) |
| Criteria $\succeq$ Lottery & Paintings & No Control on the Criteria | 0.075 | 0.019 |
| | (0.178) | (0.293) |
| Criteria $\succeq$ Lottery & Time & No Control on the Criteria | 0.156 | 0.100 |
| | (0.184) | (0.309) |
| Criteria $\geq 1/2$, Lottery $\leq 1/2$ & Control on the Lottery & No Control on the Criteria | 0.076 | -0.091 |
| | (0.207) | (0.397) |
| Criteria $\leq 1/2$, Lottery $\geq 1/2$ & Control on the Lottery & No Control on the Criteria | 0.033 | 0.080 |
| | (0.225) | (0.368) |
| Criteria $\succeq$ Lottery & Control on the Lottery & No Control on the Criteria | 0.068 | 0.167 |
| | (0.174) | (0.282) |
| Paintings & Control on the Lottery & No Control on the Criteria | 0.079 | 0.150 |
| | (0.216) | (0.372) |
| Time & Control on the Lottery & No Control on the Criteria | 0.042 | 0.393 |
| | (0.223) | (0.366) |
| Criteria $\geq 1/2$, Lottery $\leq 1/2$ & Paintings & Control on the Lottery & No Control on the Criteria | 0.012 | 0.140 |
| | (0.311) | (0.564) |
| Criteria $\geq 1/2$, Lottery $\leq 1/2$ & Time & Control on the Lottery & No Control on the Criteria | 0.053 | -0.065 |
| | (0.325) | (0.575) |
| Criteria $\leq 1/2$, Lottery $\geq 1/2$ & Paintings & Control on the Lottery & No Control on the Criteria | 0.012 | 0.222 |
| | (0.367) | (0.593) |
| Criteria $\leq 1/2$, Lottery $\geq 1/2$ & Time & Control on the Lottery & No Control on the Criteria | 0.084 | -0.147 |

Table 22: Selection of the criteria.

| | Criteria Chosen | |
|---|---|---|
| | (1) | (2) |
| | (0.347) | (0.499) |
| Criteria $\succeq$ Lottery & Paintings & Control on the Lottery & No Control on the Criteria | -0.209 | -0.227 |
| | (0.261) | (0.443) |
| Criteria $\succeq$ Lottery & Time & Control on the Lottery & No Control on the Criteria | -0.276 | -0.352 |
| | (0.264) | (0.426) |
| Male | | 0.000 |
| | | (0.147) |
| age: 40-55 | | -0.062* |
| | | (0.034) |
| age: <25 | | -0.020 |
| | | (0.075) |
| age: >55 | | -0.041 |
| | | (0.055) |
| age: Prefer not to say | | 0.058 |
| | | (0.227) |
| USA | | -0.041 |
| | | (0.036) |
| employment: Other | | 0.080 |
| | | (0.083) |
| employment: Retired | | -0.039 |
| | | (0.116) |
| employment: Self-employed | | -0.046 |
| | | (0.037) |
| Student | | 0.205* |
| | | (0.116) |
| Criteria $\geq 1/2$, Lottery $\leq 1/2$ & Male | | 0.339 |
| | | (0.311) |
| Criteria $\leq 1/2$, Lottery $\geq 1/2$ & Male | | -0.136 |
| | | (0.224) |
| Criteria $\succeq$ Lottery & Male | | -0.130 |
| | | (0.178) |
| Paintings & Male | | 0.089 |
| | | (0.250) |
| Time & Male | | 0.106 |
| | | (0.240) |
| Male & Control on the Lottery | | -0.071 |
| | | (0.237) |
| Male & No Control on the Criteria | | 0.032 |
| | | (0.200) |
| Criteria $\geq 1/2$, Lottery $\leq 1/2$ & Paintings & Male | | -0.562 |
| | | (0.399) |

Table 22: Selection of the criteria.

| | Criteria Chosen | |
|---|---|---|
| | (1) | (2) |
| Criteria $\geq 1/2$, Lottery $\leq 1/2$ & Time & Male | | -0.126 |
| | | (0.414) |
| Criteria $\leq 1/2$, Lottery $\geq 1/2$ & Paintings & Male | | -0.415 |
| | | (0.359) |
| Criteria $\leq 1/2$, Lottery $\geq 1/2$ & Time & Male | | 0.171 |
| | | (0.341) |
| Criteria $\succeq$ Lottery & Paintings & Male | | 0.156 |
| | | (0.288) |
| Criteria $\succeq$ Lottery & Time & Male | | 0.006 |
| | | (0.286) |
| Criteria $\geq 1/2$, Lottery $\leq 1/2$ & Male & Control on the Lottery | | -0.498 |
| | | (0.411) |
| Criteria $\leq 1/2$, Lottery $\geq 1/2$ & Male & Control on the Lottery | | 0.085 |
| | | (0.362) |
| Criteria $\succeq$ Lottery & Male & Control on the Lottery | | 0.070 |
| | | (0.276) |
| Paintings & Male & Control on the Lottery | | 0.101 |
| | | (0.358) |
| Time & Male & Control on the Lottery | | 0.224 |
| | | (0.332) |
| Criteria $\geq 1/2$, Lottery $\leq 1/2$ & Male & No Control on the Criteria | | -0.230 |
| | | (0.358) |
| Criteria $\leq 1/2$, Lottery $\geq 1/2$ & Male & No Control on the Criteria | | 0.021 |
| | | (0.311) |
| Criteria $\succeq$ Lottery & Male & No Control on the Criteria | | 0.007 |
| | | (0.241) |
| Paintings & Male & No Control on the Criteria | | -0.019 |
| | | (0.321) |
| Time & Male & No Control on the Criteria | | 0.147 |
| | | (0.338) |
| Male & Control on the Lottery & No Control on the Criteria | | 0.231 |
| | | (0.308) |
| Criteria $\geq 1/2$, Lottery $\leq 1/2$ & Paintings & Male & Control on the Lottery | | 0.971* |
| | | (0.556) |
| Criteria $\geq 1/2$, Lottery $\leq 1/2$ & Time & Male & Control on the Lottery | | 0.223 |
| | | (0.559) |

Table 22: Selection of the criteria.

| | Criteria Chosen | |
|---|---|---|
| | (1) | (2) |
| Criteria $\leq 1/2$, Lottery $\geq 1/2$ & Paintings & Male & Control on the Lottery | | 0.207 |
| | | (0.599) |
| Criteria $\leq 1/2$, Lottery $\geq 1/2$ & Time & Male & Control on the Lottery | | 0.334 |
| | | (0.487) |
| Criteria $\succeq$ Lottery & Paintings & Male & Control on the Lottery | | -0.324 |
| | | (0.428) |
| Criteria $\succeq$ Lottery & Time & Male & Control on the Lottery | | -0.045 |
| | | (0.400) |
| Criteria $\geq 1/2$, Lottery $\leq 1/2$ & Paintings & Male & No Control on the Criteria | | -0.074 |
| | | (0.474) |
| Criteria $\geq 1/2$, Lottery $\leq 1/2$ & Time & Male & No Control on the Criteria | | -0.216 |
| | | (0.517) |
| Criteria $\leq 1/2$, Lottery $\geq 1/2$ & Paintings & Male & No Control on the Criteria | | 0.513 |
| | | (0.497) |
| Criteria $\leq 1/2$, Lottery $\geq 1/2$ & Time & Male & No Control on the Criteria | | -0.369 |
| | | (0.486) |
| Criteria $\succeq$ Lottery & Paintings & Male & No Control on the Criteria | | 0.167 |
| | | (0.376) |
| Criteria $\succeq$ Lottery & Time & Male & No Control on the Criteria | | 0.111 |
| | | (0.390) |
| Criteria $\geq 1/2$, Lottery $\leq 1/2$ & Male & Control on the Lottery & No Control on the Criteria | | 0.200 |
| | | (0.491) |
| Criteria $\leq 1/2$, Lottery $\geq 1/2$ & Male & Control on the Lottery & No Control on the Criteria | | -0.119 |
| | | (0.469) |
| Criteria $\succeq$ Lottery & Male & Control on the Lottery & No Control on the Criteria | | -0.122 |
| | | (0.363) |
| Paintings & Male & Control on the Lottery & No Control on the Criteria | | -0.122 |
| | | (0.464) |
| Time & Male & Control on the Lottery & No Control on the Criteria | | -0.542 |

Table 22: Selection of the criteria.

| | Criteria Chosen | |
|---|---|---|
| | (1) | (2) |
| | | (0.463) |
| Criteria $\geq 1/2$, Lottery $\leq 1/2$ & Paintings & Male & Control on the Lottery & No Control on the Criteria | | -0.015 |
| | | (0.695) |
| Criteria $\geq 1/2$, Lottery $\leq 1/2$ & Time & Male & Control on the Lottery & No Control on the Criteria | | 0.212 |
| | | (0.718) |
| Criteria $\leq 1/2$, Lottery $\geq 1/2$ & Paintings & Male & Control on the Lottery & No Control on the Criteria | | -0.219 |
| | | (0.761) |
| Criteria $\leq 1/2$, Lottery $\geq 1/2$ & Time & Male & Control on the Lottery & No Control on the Criteria | | 0.531 |
| | | (0.680) |
| Criteria $\succeq$ Lottery & Paintings & Male & Control on the Lottery & No Control on the Criteria | | -0.086 |
| | | (0.557) |
| Criteria $\succeq$ Lottery & Time & Male & Control on the Lottery & No Control on the Criteria | | 0.057 |
| | | (0.547) |
| Estimator | OLS | OLS |
| $N$ | 1,324 | 1,324 |
| Adjusted $R^2$ | 0.053 | 0.057 |

P-values: *: $< 0.1$, **: $< 0.05$, ***: $< 0.01$, ****: $< 0.001$

## B.2 Non Linear Regressions

We have used a linear OLS regression in all our specifications so far, but our dependent variable is a binary one (choosing the criteria or the lottery). In fact, the prediction made by the regression models is a probability (or proportion of the population) that chooses the criteria over the lottery. In the linear model, the dependent variable is not bounded, so that we could predict proportions lower than 0 or higher than 1. The canonical models to work with discrete choices are probit and logit.

We did not find proportions that are out of bounds with the linear estimates, so that we only run probit and logit models to check that we have the same significant variables in the main ones, corresponding to Table 7 and 16. In Tables 23 and 24, we find the same significant variables in linear and non-linear regressions, and the same direction for the effects. As expected, the coefficients estimations are different. To compare them, we need to compare the impact of each of them with predictions everything else staying equal. We obtain roughly the same results by doing so (details available upon request). To conclude, our results do not depend on the specification of the regressions.

Table 23: Regression analysis comparing OLS, Probit and Logit models in the model without control.

| | Criteria Chosen | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| (Intercept) | 0.656**** | 0.409**** | 0.658**** |
| | (0.044) | (0.119) | (0.193) |
| Criteria $\geq 1/2$, Lottery $\leq 1/2$ | 0.118**** | 0.322**** | 0.524**** |
| | (0.032) | (0.090) | (0.147) |
| Criteria $\leq 1/2$, Lottery $\geq 1/2$ | -0.091** | -0.237** | -0.384** |
| | (0.036) | (0.093) | (0.151) |
| Paintings | -0.028 | -0.079 | -0.121 |
| | (0.032) | (0.086) | (0.139) |
| Time | -0.281**** | -0.746**** | -1.200**** |
| | (0.047) | (0.131) | (0.215) |
| Male | 0.007 | 0.021 | 0.035 |
| | (0.033) | (0.088) | (0.142) |
| Control on the Lottery | -0.103**** | -0.275**** | -0.444**** |
| | (0.027) | (0.071) | (0.115) |
| No Control on the Criteria | -0.053** | -0.139* | -0.230** |
| | (0.027) | (0.072) | (0.117) |
| Averse to Ambiguity | 0.042 | 0.113 | 0.181 |
| | (0.031) | (0.082) | (0.133) |
| Time & Male | 0.150** | 0.399** | 0.642** |
| | (0.059) | (0.158) | (0.257) |
| Estimator | OLS | Probit | Logit |
| $N$ | 1,324 | 1,324 | 1,324 |
| Adjusted $R^2$ | 0.062 | | |

P-values: *: $< 0.1$, **: $< 0.05$, ***: $< 0.01$, ****: $< 0.001$

Table 24: Regression analysis comparing OLS, Probit and Logit models in the model with stated preferences.

| | Criteria Chosen | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| (Intercept) | 0.582**** | 0.214 | 0.347 |
| | (0.052) | (0.135) | (0.220) |
| Time | -0.216**** | -0.569**** | -0.916**** |
| | (0.053) | (0.144) | (0.236) |
| Paintings | 0.022 | 0.055 | 0.098 |
| | (0.038) | (0.099) | (0.161) |
| Control on the Lottery | -0.087*** | -0.238**** | -0.387**** |
| | (0.027) | (0.072) | (0.117) |
| No Control on the Criteria | -0.048* | -0.133* | -0.216* |
| | (0.027) | (0.073) | (0.119) |
| Criteria $\geq 1/2$, Lottery $\leq 1/2$ | 0.056 | 0.151 | 0.245 |
| | (0.039) | (0.105) | (0.171) |
| Criteria $\leq 1/2$, Lottery $\geq 1/2$ | -0.053 | -0.141 | -0.227 |
| | (0.041) | (0.108) | (0.174) |
| Stated Preference | 0.243**** | 0.734**** | 1.200**** |
| | (0.060) | (0.194) | (0.328) |
| Paintings & Stated Preference | -0.160** | -0.491** | -0.820** |
| | (0.069) | (0.204) | (0.340) |
| Time & Stated Preference | -0.162 | -0.512 | -0.832 |
| | (0.141) | (0.397) | (0.649) |
| Probability Matters | -0.017 | -0.042 | -0.069 |
| | (0.038) | (0.101) | (0.164) |
| Criteria $\geq 1/2$, Lottery $\leq 1/2$ & Probability Matters | 0.241**** | 0.793**** | 1.363**** |
| | (0.064) | (0.224) | (0.394) |
| Criteria $\leq 1/2$, Lottery $\geq 1/2$ & Probability Matters | -0.133 | -0.352 | -0.565 |
| | (0.082) | (0.224) | (0.364) |
| Male | 0.013 | 0.035 | 0.058 |
| | (0.039) | (0.102) | (0.165) |
| Time & Male | 0.141** | 0.375** | 0.599** |
| | (0.065) | (0.174) | (0.284) |
| Estimator | OLS | Probit | Logit |
| $N$ | 1,324 | 1,324 | 1,324 |
| Adjusted $R^2$ | 0.082 | | |

P-values: *: $< 0.1$, **: $< 0.05$, ***: $< 0.01$, ****: $< 0.001$

Table 25: Choices of criteria when subjects have no control over the lottery, by the number of Even in the sequence given to them.

| | Evens in the sequence | | | | | |
|---|---|---|---|---|---|---|
| **Criteria** | 0 | 1 | 2 | 3 | 4 | 5 |
| RPS | 40.0% | 63.6% | 65.0% | 77.5% | 76.9% | 50.0% |
| Paintings | 66.7% | 63.4% | 70.0% | 62.2% | 52.2% | 75.0% |
| Time | 28.6% | 39.3% | 49.2% | 54.3% | 34.4% | 62.5% |
| Aggregate | 43.5% | 57.5% | 61.6% | 64.6% | 56.4% | 63.6% |

Table 26: Choices of criteria when subjects have no control over the lottery, by the number of even or odd in the sequence given to them. So the sequences 5 Even and 5 Odd are both under 0.

| | Evens or odds in the sequence | | |
|---|---|---|---|
| **Criteria** | 0 | 1 | 2 |
| RPS | 43.8% | 69.9% | 70.9% |
| Paintings | 71.4% | 59.4% | 65.5% |
| Time | 46.7% | 36.7% | 51.9% |
| Aggregate | 53.3% | 57.0% | 63.1% |

# C Influence of the Given Sequence

It is possible that participants are influenced in their choices between procedures by the strategies they have in each procedure. It is particularly true when they have no control over the sequence. For instance, they may believe that the sequence Even, Even, Even, Even, Even or Odd, Odd, Odd, Odd, Odd is less likely to happen than the sequence Even, Odd, Even, Odd, Even. This incorrect understanding of probabilities is more likely to happen when facing the lottery, as there is no particular order for the paintings or for the Time procedure. In RPS, subjects would have to be sophisticated in their strategies to form a belief about it, and it does not appear to be the case.

To investigate this issue, we restrict ourselves to the sample of subjects without control on the lotteries, which represents 678 participants in total. We first count the number of Even in the sequence we gave them. Table 25 shows the proportions of subjects choosing each criteria depending on the number of Even in their sequence. The sample sizes vary widely: subjects are much more likely to have a sequence with 3 Even than 5 (even if any given sequence is equally likely).

Frist we find no difference between the treatment of Odd and Even. There is no significant difference between the share of subjects choosing the criteria when awarded a sequence of 0 or 5 Even. The same holds between 1 and 4 or 2 and 3, respectively, according to a Fisher test of equal proportions. So in Table 25, Odd and Even are treated the same by participants, which is reassuring.

We can therefore group up sequences of 0 and 5 Even together, as well as sequences of 1 and 4 and 2 and 3. It yields to Table 26. The proportions of each criteria being chosen are never significantly different between 0 and 1 or 2 Even in the sequence, according to a Fisher test of equal proportions.

Finally, we run the regression with stated preferences, but restricted to subjects without control on the lottery. We add dummy for getting 0 or 1 even (or odd) and use as a baseline getting 2 or 3 evens. The results given by Table 27 shows that the dummy are not significant. The significant variables also do not change, as well as their magnitudes. It means that despite some anecdotal evidence, how the sequence looks may only be marginally taken into account by our participants in their choices of mechanism.

# D Screenshots and Supplementary Material

You can choose to participate in the experiment by following this link: https://bouacida-foucart.herokuapp.com/room/readers. It is short, around 5 minutes. Contrary to the original experiment, you can choose which treatment you participate in. If you want to do several treatments, due to otree anti-repetition methods, you will need to open the experiment in a private browser window, or in another browser. It may take a few seconds for the webpage to appear, this is normal. The results were shown to the subjects the next day on the following website: https://wp.lancs.ac.uk/lexelresults/.

## D.1 The Arbitrary Algorithm (Time)

We show here the control version of the arbitrary algorithm.

We will ask you to choose a time (in 24 hours format). For each player we will record the five last digits, that we denote as your "code". For instance, 10 hours, 26 minutes and 31 seconds become “02631”.

We have developed an algorithm ranking all participants based on their code (We expect around 100 participants today). Among others, it takes into account whether your code is above or below the median. You can read the details of the algorithm by clicking on this button.

*The details below are hidden by default, but can be revealed by clicking on a button.*

We use the following algorithm to rank the codes:

1. We will count for all the players in the experiment the number n of odd digits of the code, with 0 counting as even. In the example, the number of odd digits is n = 2.

2. We will then rank everyone according to the number n (a higher n yield a higher rank). We call this rank your "code rank".

3. For the tied players with the same number n of odd digits, we rank them by the statistical frequency of the first digit of the "code". We will give a higher rank to those with the lowest frequency, then to the second lowest one, until there are no more number left (tied frequencies are bundled together). If there is still a tie, we repeat the same procedure with the second digit of the code. And so till the last digit.

4. We then determine the winner as follows:

   - CASE 1: If strictly more players have n < 2.5 than n > 2.5 (i.e., if the median is below 2.5): your award rank is the same as your code rank. All the 50% higher ranked players in the award rank win the reward.

Table 27: Selection of the criteria when subjects have no control on the lottery.

| | Criteria Chosen | |
|---|---|---|
| | (1) | (2) |
| (Intercept) | 0.521**** | 0.540**** |
| | (0.070) | (0.076) |
| Time | -0.204*** | -0.205*** |
| | (0.076) | (0.076) |
| Paintings | -0.014 | -0.011 |
| | (0.053) | (0.053) |
| No Control on the Criteria | -0.038 | -0.035 |
| | (0.037) | (0.037) |
| Criteria $\geq 1/2$, Lottery $\leq 1/2$ | 0.056 | 0.050 |
| | (0.055) | (0.055) |
| Criteria $\leq 1/2$, Lottery $\geq 1/2$ | 0.021 | 0.019 |
| | (0.056) | (0.057) |
| Stated Preference | 0.280**** | 0.290**** |
| | (0.075) | (0.076) |
| Paintings & Stated Preference | -0.094 | -0.115 |
| | (0.086) | (0.087) |
| Time & Stated Preference | 0.161 | 0.141 |
| | (0.196) | (0.199) |
| Probability Matters | 0.085 | 0.134** |
| | (0.054) | (0.063) |
| Criteria $\geq 1/2$, Lottery $\leq 1/2$ & Probability Matters | 0.151* | 0.157* |
| | (0.085) | (0.086) |
| Criteria $\leq 1/2$, Lottery $\geq 1/2$ & Probability Matters | -0.253** | -0.244** |
| | (0.115) | (0.118) |
| Male | 0.006 | 0.002 |
| | (0.055) | (0.055) |
| Time & Male | 0.131 | 0.136 |
| | (0.092) | (0.092) |
| Sum of Odds or Evens = 0 | | -0.073 |
| | | (0.076) |
| Sum of Odds or Evens = 1 | | -0.049 |
| | | (0.040) |
| Estimator | OLS | OLS |
| $N$ | 678 | 678 |
| Adjusted $R^2$ | 0.088 | 0.089 |

P-values: *: $< 0.1$, **: $< 0.05$, ***: $< 0.01$, ****: $< 0.001$

- CASE 2: If strictly fewer players have n $<$ 2.5 than n $>$ 2.5 (i.e., if the median is above 2.5): your award rank is the revert of the code rank. If there are N participants and your code rank was j, then your award rank is N+1-j. All the 50% higher ranked players in the award rank win the reward.

5. In the unlikely event that a tie remains at the end of the procedure, exactly at the 50% mark, all tied players will receive a reward.