When to stop? A Theoretical and experimental investigation of an individual search task

Imen Bouhlel, Michela Chessa, Agnès Festré, Eric Guerci

WHEN TO STOP? A THEORETICAL AND EXPERIMENTAL INVESTIGATION OF AN INDIVIDUAL SEARCH TASK

Documents de travail GREDEG
GREDEG Working Papers Series

Imen Bouhlel
Michela Chessa
Agnès Festré
Eric Guerci

GREDEG WP No. 2019-40
https://ideas.repec.org/s/gre/wpaper.html

When to stop?
A theoretical and experimental investigation of an individual search task

Imen Bouhlel[a,*], Michela Chessa[a], Agnès Festré[a], Eric Guerci[a]

[a] *Université Côte d'Azur, CNRS,*
Groupe de Recheche en Droit, Economie et Gestion, Nice, France

GREDEG Working Paper No. 2019–40

Abstract

Information search and opinion formation are central aspects of decision making in consumers choices. Indeed, before taking a decision, the alternatives among which the *rational* choice will be made should be clearly valued. In standard economic theory, the search dynamics is generally neglected because the process is assumed to be carried out without any cost or without spending time. However, whenever only a significant collection of experience can provide the bulk of relevant information to make the best choice, as it is the case for experience goods (Nelson, 1970), some engendered costs in collecting such information might be considered. Our paper lies on a conceptual framework for the analysis of an individual sequential search task among a finite set of alternatives. This framework is inspired by both the *Secretary problem* (Ferguson et al., 1989) and the *multi-armed bandit problem* (Robbins, 1952). We present a model where an individual is willing to locate the best choice among a set of alternatives. The total amount of time for searching is finite and the individual aims at maximizing the expected payoff given by an exploration-exploitation trade-off: a first phase for exploring the value of new alternatives, and a second phase for exploiting her past collected experience. The task involves an iterative exploitation – i.e., where the final payoff does not only depend on the value of the chosen alternative, but also on the remaining time that has not been dedicated to exploration –. Given the finite horizon of time, the optimal stopping strategy can be assimilated to a *satisficing* behavior (Simon, 1956). We manipulate the degree of certainty of information, and we find that the optimal stopping time is later under the uncertain information condition. We experimentally test the model's predictions and find a tendency to oversearch when exploration is costly, and a tendency to undersearch when exploration is relatively cheap. We also find under the certain information condition that participants learn to converge towards the optimal stopping time, but this learning effect is less present under the uncertain information condition. Regret and anticipation lead to more exploration under both information conditions. A gender effect is also exhibited with women tending to explore more than men.

*Corresponding author
Email address: imen.bouhlel@gredeg.cnrs.fr (Imen Bouhlel)

Keywords: Optimal-stopping, Exploration-exploitation trade-off, Regret, Experiment

1. Introduction

When a firm disposes of a capital that it wants to invest in research and development (R&D) in a strategic domain, where multiple options are possible, a challenging issue is to know when and in which option to invest. The firm's resources are limited and can only be engaged in one direction at a time. But, without having any knowledge about any of the options, one cannot make a choice, unless a random one. Uncertainty is intrinsic to R&D and plays a fundamental role as for the decision to invest. Czarnitzki and Toole (2011) find for example that firm-level R&D investment falls in response to higher levels of uncertainty as perceived through revenue volatility. Before making a choice, the firm needs then to collect some information about the potential of the different options in order to reduce its uncertainty and guide its choice. By doing this, it delays its investment. Waiting has the advantage to possibly avoid large losses for the firm if it discovers some unfavorable information about that option, in which case it would forgo the investment. But this information collection is costly in terms of time, and possibly also in terms of resources and experts consulting, making only a sequential study of the different options possible. R&D investment has moreover the particularity of being considered as an irreversible form of investment. For example, a large proportion of R&D covers the salaries of research personnel that cannot be recouped, as well as other types of sunk costs. The ability to delay investment decisions can be even more valuable in this case. Models from real options theory integrating irreversible investment are very helpful in order to understand how uncertainty influences R&D investment. They predict that greater uncertainty about market revenues reduces investment in irreversible capital by increasing the value of waiting to invest (Pindyck et al., 1991; Dixit, 1992; Dixit et al., 1994). Moreover, Goel and Ram (2001) found "a much sharper adverse effect of uncertainty on R&D investments, which are likely to be highly irreversible, than on non-R&D (and aggregate) investments."

However, considering a finite time horizon, waiting too long reduces the actual investment duration, and consequently decreases the total amount of returns on investment. Thus, a trade-off is faced here, as for how to optimally allocate the available time between information collection and actual investment. There may exist too many options, and collecting information about all of them does not leave enough time for investment, and may even make it unfeasible. Thus, the firm should consider only a subset of these options. In addition, an option can be explored to different degrees. Thus, a second trade-off is faced, as for how many options to study, and to what extent to explore each of these options: is it better to concentrate on a small number of options, or to roughly study a big number?

We are here facing what is referred in the literature as an optimal stopping problem with an uncertainty component. Optimal stopping problems entail an exploration-exploitation trade-off (March, 1991): a trade-off between exploiting a safe known option, but possibly sub-optimal, and exploring a new risky one, which may be much better, but may also be very unprofitable. This type of problems can be faced in many other economic and real-world situations, that have the characteristic of an investment, such as searching for an apartment, a job or a mate[1]. In addition to the optimal stopping theory, the exploration-exploitation trade-off has also been investigated in a separate and rich literature: the *multi-armed bandits* problems (Robbins, 1952)[2]. Unlike in the standard optimal stopping problems, this class of problems accounts for uncertainty. However, two main differences with optimal stopping problems are that the number of alternatives is lower than the number of trials[3], and that switching between alternatives during the exploration phase is possible.

The most famous optimal stopping problem studied in economics is the *Secretary problem* (Ferguson et al., 1989). It describes the situation where a firm needs to find the best applicant for a secretarial job. The firm can see each of the applicants sequentially. After seeing an applicant, the firm can rank her quality, and must decide to hire or reject her, without being able to return to a rejected applicant, or predict the quality of future applicants[4]. The basic problem (Rapoport and Tversky, 1970; Ferguson et al., 1989; Seale and Rapoport, 1997; Bearden et al., 2006) as well as many of its extensions have been studied both theoretically and experimentally, embedding for example explicit search costs (Rapoport and Tversky, 1970; Zwick et al., 2003), unknown population size (Seale and Rapoport, 2000), cardinal instead of ordinal alternatives (Lee, 2006), cognitive load (Caplin et al., 2011), possibility of recall (Rapoport and Tversky, 1966; Hey, 1987; Zwick et al., 2003), iterative exploitation (Eriksson and Strimling, 2009, 2010; Sang et al., 2018), depleting alternatives (Sang, 2017), dynamic environments (Navarro et al., 2016), or mutual search (Todd and Miller, 1999; Beckage et al., 2009). However, none of theses studies accounts for uncertainty at the level of the alternatives. Indeed, one trial is enough to observe the value (or the rank) of an alternative with certainty.

The existing experimental studies considering *à la Secretary problem* settings lead to the conclusion that the participants tend to undersearch compared to the theoretical prediction (Hey, 1987; Seale and Rapoport, 1997; Sonnemans, 1998; Seale and Rapoport, 2000; Bearden et al., 2006; Schunk and Winter, 2009; Oprea et al., 2009; Costa and Averbeck, 2013), while some other studies observed the opposite effect where participants tended

[1] In the mate search or "revisited" marriage example, waiting has a value, since marriage is a life risky decision, and one does not necessarily want to gamble and marry a random person. Instead, one may be searching for the perfect mate, but searching too long leaves less time to enjoy marriage, build a family etc. Thus, one may engage into dating few potential partners (one at a time). The longer the dating period, the better one gets to know the other person. However, dating is costly (courtship costs). Also, it may be not possible to jump back and forth between the different relationships. Though, marriage example has few main differences with the one of R&D investment: it involves a risk in recalling an old explored alternative (it is possible that that ex is no more free, or does not agree to get back after "break-up"). Second, marriage is a two-sided decision, thus the selected potential partner may have his own first choice.

[2] See Bubeck et al. (2012) for a review.

[3] See however the *infinitely many armed-bandits* (Berry et al., 1997).

[4] This problem has mainly attracted attention due to its elegant theoretical solution which consists at interviewing and rejecting the first $1/e$ ($\approx 37\%$) of the applicants, considering the "value" of the best of these rejected applicants as a threshold, then continuing interviewing and hiring the first applicant who is above this threshold

to oversearch (Eriksson and Strimling, 2010; Sandri et al., 2010; Juni et al., 2016; Sang et al., 2018). The settings used in these papers present a main difference compared to *à la Secretary problems*: the payoff does not only depend on the value of the chosen alternative, but also on the remaining time that was not used for exploration. Indeed, the goal is not only to find the best possible alternative, but also to enjoy it for long enough. Some studies interestingly found the two opposite effects in a single setting when exploration costs were manipulated (Zwick et al., 2003; Descamps et al., 2016). In Zwick et al. (2003), the authors find that participants inspect too many alternatives when inspection is costly, and too few when it is free. Confirming that result, Descamps et al. (2016) find that when sampling is relatively expensive, participants oversample; on the other hand, when it is relatively cheap, they undersample. Also Juni et al. (2016) found in an estimation task based on information sampling in which they manipulated the relative cost per cue that participants oversampled only when this cost was high, but did not show any significant over or under sampling when this cost was low.

When only the ranks of the alternatives are available, the optimal search strategy involves observing a fixed number of values, then choosing the first subsequently maximal one (Gilbert and Mosteller, 1966; Rapoport and Tversky, 1970; Zwick et al., 2003). When the values of the alternatives are available, the optimal search strategy is described in terms of a decreasing threshold strategy, where an alternative should be accepted if its value exceeds the threshold for the current position in the sequence (Lee, 2006; Sang et al., 2018). These thresholds can naturally be assimilated to the aspiration levels in Simon's theory of *satisficing* (Simon, 1955, 1982). Several studies were interested in the learning dynamics towards the optimal stopping behavior. Contrary results are also found here. While many studies (Seale and Rapoport, 1997; Zwick et al., 2003; Oprea et al., 2009; Sang et al., 2018) conclude that participants are able to learn to converge towards the optimal solution, some others find evidence for a lack of learning (Campbell and Lee, 2006; Lee, 2006). Descamps et al. (2016) find a mixed result where the observation of learning highly depends on the sampling costs. Indeed, when sampling is relatively costly, they find that participants tend to learn over time to improve their search strategy. However, when sampling is "cheap", they fail to observe any improvement over time. Thus, learning seems to happen only under particular settings and particular exploration costs.

Regret has been shown to play a certain role in the stopping behavior dynamics (Zwick et al., 2003; Oprea et al., 2009; Viefers and Strack, 2014), where local features of the observed sequence of inspected alternatives, which should have been ignored by a fully rational agent, seem to influence the participants' length of search. Zwick et al. (2003) considered for example measures of anticipation and regret[5] and showed that the search behavior of the participants was sensitive to these measures. Viefers and Strack (2014) also found that participants' stopping behavior was largely determined by the anticipation of and aversion to regret. As for in Oprea et al. (2009), the authors find an impact of foregone earnings on reservation levels where forgone earnings from previous rounds tend

[5]The authors consider two measures: the AROCA (*Average Rate Of Candidates Arrival*) as a measure of anticipation, and the PSLC (*Periods Since Last Candidate*) as a measure of regret.

to affect choices in later rounds.

The experimental literature highlights large individual differences and heterogeneity in the dynamics of search behavior (Lee, 2006; Schunk, 2009). Counter-intuitively, several studies have found that risk aversion does not play any role in different settings of optimal stopping problems (Sonnemans, 1998; Oprea et al., 2009; Schunk and Winter, 2009; Schunk, 2009). Surprisingly, Eriksson and Strimling (2009) and Eriksson and Strimling (2010) both find a gender effect where women tend to explore more than men, possibly suggesting triggered cognitive gender differences. Also, von Helversen and Mata (2012) find evidence for an age effect, where older adults tend to perform poorly compared to younger ones. Relatedly, Sandri et al. (2010) suggest that high levels of positive affect can lead to insufficient search in sequential decision making in elderly persons.

We consider a setting which, to our knowledge has not yet been investigated in the optimal stopping literature since it involves an uncertainty component at the level of the alternatives' values. In our setting, we consider a iterative exploitation – i.e., where the payoff does not only depend on the value of the chosen alternative, but also on the remaining time that has not been used for exploration –, a payoff function that accounts only for the exploitation phase (unlike in Sang et al. 2018, where the authors also consider iterative exploitation but where the payoff function accounts for both the exploration and the exploitation phases), and a possible recall of the explored alternatives.

Our paper is organized as follows: First we theoretically model our situation as a class of sequential stopping rule problems, where the quality (or value) of the alternatives is given by some random variable whose joint distribution is known and that the individual can sequentially observe. Depending on the condition, the quality of each alternative is observed with certainty or with uncertainty. In both conditions, given the past draws, the individual has to choose whether to stop exploring and select one of the observed alternatives, or to continue observing in order to maximize her utility. Moreover, if the quality is uncertain, the observed values are sampled from a normal distribution around the exact alternative's quality as a generative mean, and with a known standard deviation. Thus, the individual faces an additional quantity-accuracy trade-off during the exploration phase, between the quantity of explored alternatives and the accuracy of the collected information about each of these alternatives.

Second, we present the experimental implementation of our model. The basic protocol is a repeated game, which consists in sequentially drawing some alternatives from a total of n alternatives with hidden payoffs. The player can click m or up to m times (depending on the treatment) on the same alternative, where the total number of trials is equal to $m \times n$. By clicking on an alternative, the player discovers the information about its payoff. Depending on the treatment, this information can be certain or uncertain. By choosing an alternative, the player "exploits" it and obtains its payoff for the remaining trials. Her objective is to maximize the sum of the stream of payoffs from the exploitation phase. The players repeat the game 60 rounds, alternating sequences of both treatments (within-subjects). The order in which these sequences are encountered is also manipulated in a between-subjects design. The main task is followed by the *Bomb Risk Elicitation Task* (Crosetto and Filippin, 2013) to measure the individual risk preferences, and the *Sustainted Attention to Response Task* (Robertson et al., 1997) to control for the impulsivity level.

Then, we compare in terms of performance and exploration amount the theoretical predictions to the experimentally observed behavior and we study the learning dynamics towards those predictions. We find in line with the literature, a tendency to overseach when exploration is costly, and a tendency to undersearch when exploration is relatively cheap, as well as a more pronounced learning effect under certain information treatment, i.e. when the exploration of a new alternative is more costly. We also model the stopping decision using a survival analysis in order to account for behavioral measures of regret and anticipation. Our main findings are that both regret and anticipation lead to more exploration, and that women tend to explore more than men.

2. Theoretical model

2.1. Sequential stopping rule problem

The theory of optimal stopping is concerned with the problem of choosing a time to take a given action, based on sequentially observed random variables and in order to maximize an expected payoff. Formally, in the standard literature, adopting the well known definition given by Ferguson (Ferguson et al., 1989), *stopping rule problems* are defined by two objects:

(i) a *sequence of random variables* $X_1, \dots, X_n$, whose joint distribution is assumed known,

(ii) a *sequence of real-valued reward functions*

$$u_0, u_1(x_1), u_2(x_1, x_2), \dots, u_\infty(x_1, x_2, \dots).$$

In the general stopping rule problem, given these two objects, an agent may observe the sequence $X_1, X_2, \dots$ for as long as she wishes. For each $t = 1, 2, \dots$, after observing $X_1 = x_1, X_2 = x_2, \dots, X_t = x_t$, she may stop and receive the known payoff $u_t(x_1, x_2, \dots, x_t)$, or she may continue and observe the realization of X_{t+1}. In our model, we consider instead the variation of this stopping rule problem with finite horizon of time. A stopping rule problem has a *finite horizon* T if there is a known upper bound on the number of stages at which one may stop. This modification is useful to model many real-world applications, in which the time for investigation is limited. A stopping rule problem with finite horizon may be obtained as a special case of the general problem described below by (i) and (ii), by setting $u_{T+1} = \dots = y_\infty = -\infty$. Such problems may be solved by the method of *backward induction.*

2.2. Our theoretical model under certain observations

In our basic model, that theoretically describes our first treatment, we suppose that the horizon of time and the number of choices are finite and coincide, namely, $n = T$. Then, from now on, we will use the variable n to denote both quantities. Suppose that the sequence in (i) is such that the X_i, with $i = 1, \dots, n$, are i.i.d., $X_i \sim \mathcal{U}(\{1, \dots, N\})$. In our interpretation, the observed value x_i represents the intrinsic quality of good i, which is uniformly distributed between a minimum of 1 and a maximum of N, and which does not depend on the quality of the other ones. This intrinsic value is fixed and may be observed with certainty by the agent. For each $t = 1, \ldots, n-1$, the agent has already observed the quality of the first t goods and decides whether to exploit one of them for the remaining time, or to explore a new alternative. When $t = 0$, the exploration has not started yet, and then there is no alternative available for future exploitation. When $t = n$, time is over and the agent does not have the possibility of exploiting anymore any of the alternatives. We consider a reward function with recall factor equal to 1, meaning that it is always possible to recall a past alternative, and with memory factor equal to 0, meaning that we do not have any memory of the past utilities experienced during the exploration phase, and we can start enjoying a strictly positive utility only when exploiting. Such assumption is typical of situations, such as the secretary problem, in which the exploration phase does not really correspond to a real enjoyment of the good. Moreover, our problem describes a classical exploration-exploitation trade-off in which, once an agent has decided to start exploiting, she cannot go back exploring. The decision is then definitive for the remaining steps in the time horizon. Then, at each time t, if an agent decides to stop, her reward function is defined as

$$\begin{cases} u_0 = 0 \\ u_t(x_1, \ldots, x_t) = (n-t)\max_{i=1,\ldots,t}\{x_i\}, \quad t = 1, \ldots, n \end{cases} \tag{1}$$

i.e., it is given by the exploitation for the remaining horizon of time $n - t$ of the best alternative found during the exploration phase. We observe that, in a more general setting and how assumed in our experiments, an agent may be allowed to choose an alternative which is not the best one, up to her knowledge. However, such a strategy will always be dominated by the choice of exploiting the maximum known value and then, without any loss of generality, we can exclude this eventuality by the theoretical analysis of the model.

2.3. Theoretical results under certain observations

The optimal exploration-exploitation trade-off can be investigated by the method of backward induction. First, we observe that, as shown in (1), at $t = 0$ by exploiting the agent would have a utility equal to zero, as she does not have any information about any of the alternatives and then it is optimal for her to explore. Second, we observe that, when $t = n$, the agent has always utility equal to zero, as she has completed her exploration but she has no time left for exploiting any of the past alternatives. Then, for each other step $t = 1, \ldots, n-1$, we define the optimal threshold $\bar{m}_t$ below which it is optimal for her, in expected value, to keep on exploring, and above which it is better to exploit.

Step n-1: the agent has already drawn $n-1$ balls. Let M_{n-1} be the random variable maximum of the first $n-1$ drawings. The agent knows its realization m_{n-1}, as she knows the realization x_i of each random variable X_i with $i \in \{1, \ldots, n-1\}$, results of all the past drawings. If she decides to stop and exploit, she can get a certain amount equal to $m_{n-1} > 0$, while if she decides to continue and explore, she gets 0. As obvious, at step $n-1$ the optimal choice is always to exploit.Then, the threshold m at step $n-1$ is $\bar{m}_{n-1} = 1$.

Step n-2: the agent has already drawn $n-2$ balls. Let m_{n-2} be the maximum of the first $n-2$ drawings. If she decides to exploit, she gets a certain amount equal to $2m_{n-2}$, while if she decides to explore, she gets an expected utility

$$\mathsf{E}(M_{n-1}|M_{n-2}=m_{n-2}),$$

that is equal to

$$\begin{aligned}
&\frac{m_{n-2}^2}{N}+\sum_{k=m_{n-2}+1}^{N}\frac{k}{N}\\
=&\frac{m_{n-2}^2}{N}+\frac{N(N+1)}{2N}-\frac{m_{n-2}(m_{n-2}+1)}{2N}\\
=&\frac{2m_{n-2}^2+N(N+1)-m_{n-2}(m_{n-2}+1)}{2N}\\
=&\frac{m_{n-2}^2-m_{n-2}+N(N+1)}{2N}
\end{aligned}$$

This quantity is bigger than $2m_{n-2}$ if and only if

$$m_{n-2}^2-m_{n-2}+N(N+1)>4Nm_{n-2}$$

and the inequality is verified if

$$m_{n-2}<\frac{1+4N-\sqrt{12N^2+4N+1}}{2}\bigcup m_{n-2}>\frac{1+4N+\sqrt{12N^2+4N+1}}{2}.$$

The quantity on the right is bigger than N, then the inequality is never verified. We obtain that is it convenient to draw if

$$m_{n-2}<\frac{1+4N-\sqrt{12N^2+4N+1}}{2}$$

and then the threshold m at step $n-2$ is $\bar{m}_{n-2}=\frac{1+4N-\sqrt{12N^2+4N+1}}{2}$. When N is "big'", we may simplify and write the threshold as

$$\bar{m}_{n-2}\sim(2-\sqrt{3})N.$$

Step t: we observe that if $m_t\le\bar{m}_{t+1}$, it is always advantageous to explore. Then, we suppose now that the agent observes $m_t>\bar{m}_{t+1}$. It trivially follows that at the following step $t+1$, $m_{t+1}\ge m_t>\bar{m}_{t+1}$ and then it will be advantageous to exploit at time $t+1$, if the agent has not started exploring yet at step t. Then, with the same reasoning than before, at step t, it is convenient to draw if

$$m_t<\frac{(n-t-1)+2N(n-t)-\sqrt{(n-t-1)^2+4N^2(2n-2t-1)+4N(n-t-1)}}{2(n-t-1)}.$$

When N is "big", we may simplify and write the threshold as

$$\bar{m}_t\sim\frac{n-t-\sqrt{2n-2t-1}}{n-t-1}N. \tag{2}$$

We define $f(y) := \frac{y-\sqrt{2y-1}}{y-1}$ and we observe that this function is decreasing and concave in y. With $y := n - t$ and up to the multiplicative constant N, this function defines the threshold $\bar{m}_t$ below which it is convenient to keep on exploring and above which is convenient to start exploiting, when exactly y steps are missing in our finite horizon of time and when N is large. These threshold under certainty observations are represented in Figure 1.

2.4. Our theoretical model under uncertain observations

In our second treatment, we present a similar optimal stopping rule problem, in which the intrinsic value of a good is still fixed. Differently from before, such a value may now be observed by the agent with certainty, spending, as before, the necessary amount of time, or with more uncertainty, with the advantage that some time is saved for exploring other alternatives or for exploiting for a longer time. In many real-world examples, this additional variable represents well the fact that, while exploring an alternative, we can get a fast idea of its value, but it could take a long time to be sure about it and, in general, the more the time we dedicate to explore the alternative, the more accurate the information we get about it. The agent faces now two different trade-offs: the first one, as before, is about choosing the right moment for switching from the exploration to the exploitation; in the second one, instead, the agent faces the problem, for each alternative, to decide how precise she wants her exploration to be. She could decide, for example, to get a more accurate information, reducing the time for exploitation but limiting the risk of getting a lower utility than expected, or she could prefer, instead, to trust a less accurate information, letting more time for exploring other alternatives or for exploiting an expected value.

Formally, in our model the agent has still a maximum number of alternatives to explore equal to n and an horizon of time of $T = n$ steps, but she is not obliged as before to spend a complete step to see the exact realization of a random variable $X_i = x_i$. Each step is divided in s substeps and in each substep she observes a certain number of normally distributed values $y_{ij} = x_i + \epsilon_{ij}$, where ϵ_i are realizations of some iid zero-mean random variables with finite variance σ^2. After a total of k observations of alternative i, the agent can estimate its intrinsic value implementing the mean $\sum_{i=1}^{k} \frac{y_i}{k}$ with standard deviation $\sigma_M = \frac{\sigma}{\sqrt{k}}$. In particular, in our implementation we suppose that each substep of exploration of an alternative is twice more informative than the previous one exploring the same alternative. Formally, the first substep deciding to explore good i the agent gets the information about 2 observations, the second step about 4 (getting a total of 6), third step about 8 observations (getting a total of 14), fourth step about 16 observations (getting a total of 30), and so on at step h getting the information about 2^h observations for a total of $H := \sum_{j=1}^{h} 2^j$. In such a model, the variance is divided by two at each step till the moment in which, at substep s, the information is given by a total of $S := \sum_{j=1}^{s} 2^j$ observations. All the parameters in our second treatment are calibrated in order to get the almost certainty if the agent decides to use the complete number of substeps for investigating an alternative. As in the previous treatment, we consider a reward function with recall factor equal to 1 and memory factor equal to 0, the last assumption meaning that a deeper exploration significantly reduces the time for exploitation and then, the time for enjoying the utility. Moreover, we assume that an agent can decide how long she wishes to explore an alternative, but she can never go back keep on exploring a past alternative she has already partially explored. Under uncertainty, if the agent decides to start exploiting alternative l after having observed t alternatives, each alternative $i = 1, \dots, t$ for a number h_i of substeps (i.e., for a total of $H_i = \sum_{j=1}^{h_i} 2^j$ observations each), she will have an expected utility equal to:

$$\mathbb{E}[u_{t,h_1,\dots,h_t}(y_{11}, \dots, y_{1h_1}, \dots, y_{t1}, \dots, y_{th_t}, l)] = \frac{ns - \sum_{i=1}^{t} h_i}{s} \frac{\sum_{j=1}^{H_l} y_{lj}}{H_l}. \quad (3)$$

As for the previous case, it is always possible, in our experimental design and in our theoretical model, to decide to explore an alternative with a lower mean of the observed variables.

2.5. Theoretical results under uncertain observations

At first we observe that, whenever deciding to explore an alternative l after the exploration of t alternatives for $h_1, \dots, h_t$ substeps respectively, with $h_i > 0$ for at least one $i = 1, \dots, t$, such a strategy is dominated by the choice of exploiting alternative l after the exploration of t alternative for 1 substep each. Then, a priori for the agent it is optimal to decide to spend the minimum amount of time in exploring each alternative. Then, after exploring t alternatives (one substep each), if an agent decides to stop, her reward function in (3) is now given by

$$u_{t,1,\dots,1}(1, \dots, 1) = \frac{ns - t}{s} \max_{i=1,\dots,t} \left\{ \frac{y_{i1} + y_{i2}}{2} \right\}. \quad (4)$$

For each $i \in N$, we denote $\bar{x}_i = \frac{y_{i1}+y_{i2}}{2}$, the mean of the first two observations of an alternative, i.e., the mean of the observations provided in the first substep.

Analogously to before and by backward induction, at step t when the agent already knows the observations $\bar{x}_1, \dots, \bar{x}_t$ and, consequently, the value of their maximum $\bar{m}_t$, is she decides to exploit, she gets a certain utility equal to $\bar{m}_t \frac{ns-t}{s}$, while if she decides to explore the last alternative, she gets an expected utility equal to

$$\left[\frac{\bar{m}_t^2}{N} + \sum_{k=\bar{m}_t+1}^{N} \frac{k}{N} \right] \frac{ns - t - 1}{s}.$$

Consequently, when N is “big” the threshold below which it is optimal to draw is

$$\bar{m}_t \sim \frac{ns - t - \sqrt{2ns - 2t - 11}}{ns - t - 1} N, \quad (5)$$

which is the analogous of the threshold in (2), when the agent has already consumed t over a total of ns substeps. These threshold under uncertainty observations are represented in Figure 1.

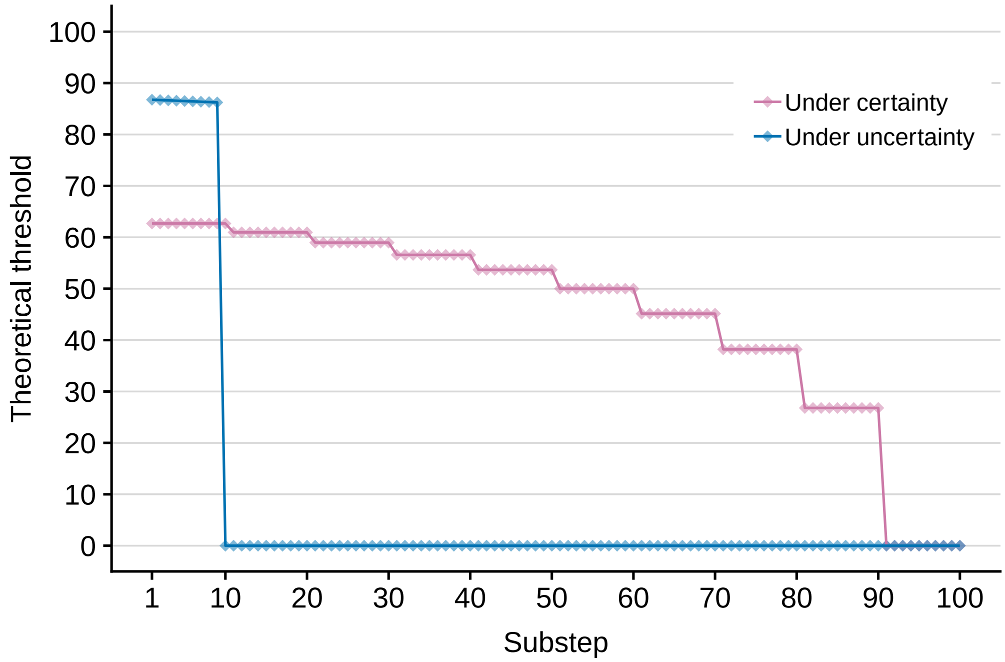


Figure 1: Threshold $\bar{m}_t$ when $n = T = 10$, $N = 100$

Figure 1 illustrates the theoretical thresholds for particular parameters values ($n = T = 10$, $N = 100$) under both certain and uncertain observations' conditions.

3. Experiment

3.1. The experimental protocol

The timeline of the experiment is described below in Figure 2. The participants played the search task after listening to the instructions and answering a comprehension questionnaire (see Supplementary materials). The search task is the implementation of the theoretical model described above. We used two treatments. In Treatment C, we implemented our model under Certainty, and in Treatment U, our model under Uncertainty (see subsection 3.1.1 for more details). The search task was followed by two control tests and a short demographic questionnaire asking for the gender and the age class of the participant.

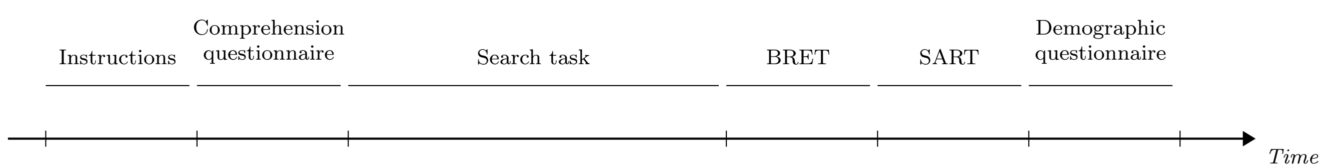


Figure 2: Timeline of the experiment

The players repeated the search task for 60 rounds, alternating sequences of both treatments (within-subjects). The order in which these sequences are encountered is also manipulated in a between-subjects design (see Table 1). In both treatments orders, the main task is preceded by a comprehension questionnaire of five questions, to which the correct answers are given after answering the full questionnaire.

	10 rounds	**10 rounds**	**10 rounds**	**10 rounds**	**20 rounds**
Treatments order CU:	Treatment C	Treatment U	Treatment C	Treatment U	Random combination
Treatments order UC:	Treatment U	Treatment C	Treatment U	Treatment C	Random combination

Table 1: Treatments orders

3.1.1. Main task

The main task is the sequential search task described in Section 2. Figure S1 in the supplementary materials shows the computer screen of the participant at the beginning of the game. The player faces $n = 10$ alternatives. The values of these alternatives are discrete and drawn from a uniform distribution $\mathcal{U}(1, N)$ where $N = 100$. The values of the alternatives are hidden and can be discovered sequentially, from the left one to the right one. Each alternative is represented by a graduation, at the bottom of which are placed two buttons: "Voir" ("See") and "Choisir" ("Choose").

The player disposes of a total of 100 trials that she can freely[6] distribute between the *exploration* of as many alternatives as she wants, and the *exploitation* of the selected alternative.

Exploring alternatives

1) To *explore* or discover an alternative, the player can click m or up to m times on its "See" button, depending on the treatment. m corresponds to the maximum number of trials the player can use to explore a given alternative, and is equal to 10. In Treatment C, the player is forced to click exactly 10 times on an explored alternative, thus is forced to discover its true value (see next point). In Treatment U she is free to click as many times as she wishes between 1 and 10 on an explored alternative. She thus can stop the discovery process when the true value of the alternative is still uncertain. Table 2) reports these settings.

2) By clicking on the "See" button of an alternative, $\sum_{i=1}^{k} 2^i$ values, where k is the click number on that alternative, are sampled from a Normal distribution with as mean the value of alternative, and with a standard deviation of $\sqrt{\sum_{i=1}^{m} 2^i}$, thus a standard deviation of around 45.23. We used this generating process in order to guarantee the discovery of the exact value of the alternative after 10 clicks. Inddeed, at the 10th click, the standard error of the alternative's value is of around 1[7]. The player then sees the mean of the sampled values represented by a dot, and an interval around that dot representing

[6]At the first trial, the player has only one possible action: "see" the first alternative.

[7]The players were simply told that the values where sampled from a normal distribution, that at each click, the sample size increased, and at the last click, the interval inside which the value of the alternative was with about 95% certainty becomes very small.

the [mean − 2 mean standard error, mean + 2 mean standard error] interval[8] (see Figures S2 and S3 in Supplementary materials).

3) It is not possible to jump back and forth between already explored alternatives in order to continue their exploration. Once an alternative is *left*, it is no more possible to go back to *explore* it again, but only to select for *exploitation*.

	Treatment C	Treatment U
Alternatives	10	10
Trials	100	100
Minimum trials by alternative	10	1
Maximum trials by alternative	10	10

Table 2: Treatments

Exploiting an alternative

The player can select one single alternative to *exploit*. When she decides to *exploit* an alternative and thus to stop *exploring*, she needs to click on the "Choose" button. A non explored alternative so far cannot be chosen. When choosing an alternative, the payoff is:

$$\frac{\text{Number of remaining trials} \times \text{Value of the selected alternative}}{10}$$

If the player spends all the trials on the *exploration* phase (i.e. does not choose any alternative), then her payoff is equal to zero.

3.1.2. Control tasks

The main task is followed by two control tasks as well as a demographic questionnaire asking for the gender and the age class of the participant. The first control task is the *Bomb Risk Elicitation Task* (Crosetto and Filippin, 2013). We translated the program from Holzmeister and Pfurtscheller (2016). The BRET is a visual real-time risk elicitation task (see Figure S4a in Supplementary materials). It consists of a two-dimensional grid in which each cell represents a box. Only one of the boxes contains a bomb. The player can select as many boxes as she wants, where each empty box yields points. Once the player decides to stop, the collected boxes are all uncovered at the same time. If the bomb has been collected, it destroys all the player's points. Comparatively to other commonly used tests (Holt and Laury, 2002; Eckel and Grossman, 2002), this task is simple to understand, and allows to classify subjects precisely, given that the choices vary almost in the continuum (Crosetto and Filippin, 2016). Moreover, it neither suffers from the impact of loss aversion – as it does not provide any endogenous reference point – nor from inconsistent decisions as it entails a unique choice (Crosetto and Filippin, 2016). In addition, no gender gap in risk aversion can be observed in this task (Crosetto and Filippin, 2013).

[8] Players were informed that if the upper bound of the interval was greater than 100, it was replaced by 100. If the lower bound of the interval was lower than 1, it was replaced by 1.

The type of sequential optimal stopping problems we are studying entails indeed an exploration-exploitation trade-off that is by definition a trade-off between exploiting a safe option and exploring a risky one. Nevertheless, while one could naturally conceive that risk aversion would be linked to this type of problems, several studies did not find any correlation with it. We consider the BRET risk elicitation task in particular as it can be viewed, contrarily to the tasks used in these studies, as a "sequential" risk task (in which risk taking is rewarded up to a point beyond which taking risk is likely to result in diminishing returns and increasing potential losses), and we want to test whether individual risk preferences elicited by this task are linked to more search and exploration.

The second control task is the *Sustained Attention to Response Task* (Robertson et al., 1997). The SART measures executive functioning – specifically, a person's ability to monitor her performance for errors and to inhibit incorrect responses. The task consists a total of 225 digits (25 for each of the nine digits) displayed on screen, in a black rectangle background (see Figure S4b in Supplementary materials). The digits are displayed in a random order. Each digit is displayed during 250 milliseconds, and followed by a crossed circle symbol (⨂) during 900 milliseconds. The player is required to respond[9] each time a digit is displayed on screen, unless this digit is a 3. The player is asked to be as fast as accurate. Despite its name, it has been suggested that the SART may be a better measure of impulsivity (i.e. failure of response inhibition) than it is of sustained attention (Dockree et al., 2004; Helton, 2009; O'Connell et al., 2009; Carter et al., 2013). Moreover, Roebuck et al. (2015) suggested that in this task omission errors (i.e. missing a GO target, here any digit except 3) may be the better indicator of inattention in this task, whilst the commission errors (i.e. failure to withhold response to a NO-GO target, here a 3) are a better measure of impulsivity. We use this task in order to investigate whether impulsivity is linked to exploration and oversearch. This effect migth be stronger in Treatment U in which exploration is less costly.

In order to incentivize the task, we gave the players the following payoff function:

$$\pi_{SART} = \begin{cases} 0 & \text{if } n = 0 \\ max(n_{correct} - 8 \times n_{error}, 0) + 200 \times \frac{T - t_r}{T} & \text{otherwise.} \end{cases}$$

where:
n is the total number of responses made by the player,
$n_{correct}$ is the number of correct responses,
n_{error} is the number of commission error responses,
t_r is the average response time in milliseconds,
T is the time in milliseconds during which a digit and its following symbol are displayed, and is equal to 1150.

[9]by pressing a key, or clicking inside the black rectangle in our case.

3.2. Participants

We recruited 77 participants among which 46 were female, 48 were in the age group of 20 to 25 years old. Participants were paid €5 for showing up to the experiment and a performance contingent bonus of €18 in average after spending about one hour and half in the lab. Participants were recruited using ORSEE (Greiner, 2015). Four experimental sessions were run with 15 to 23 subjects, and took place on September 26th, October 24th and October 26th 2018 at the Laboratory of Experimental Economics of Nice (LEEN - NiceLab)[10]. The experiment was implemented using the framework oTree (Chen et al., 2016).

3.2.1. Behavioral hypotheses

Treatment effects

In Treatment C (certain information), the exploration of a new alternative is much more costly compared to Treatment U (uncertain information). Indeed, while it requires 10 trials in Treatment C (which are then lost for the exploitation phase), the exploration can be realized by less trials in Treatment U, though at the price of a higher uncertainty. As discussed above, literature has shown that when sampling is relatively expensive, participants oversample and tend to learn over time to improve their search strategy; on the other hand, when it is relatively cheap, they undersample and fail to improve over time. Moreover, We expect to observe the same behavioral pattern in our data: a tendency to oversearch in Treatment C and to undersearch in Treatment U, and a stronger learning in Treatment C compared to Treatment U.

H1: A tendency to oversearch in Treatment C.

H2: A tendency to undersearch in Treatment U.

H3: A learning effect in Treatment C.

H4: A weaker learning effect in Treatment U compared to Treatment C.

Search determinants

We consider the number of candidates[11] met at a given position in the sequence as a measure of anticipation, and expect to find, as in Zwick et al. (2003), a positive effect of this measure on the search amount. The intuition behind this hypothesis is that the abundance of candidates would lead the participants to the erroneous optimistic belief that such a trend will continue ("*hot hand fallacy*"), thus increase the probability that they explore a new alternative.

H5: Anticipation increases the search amount.

Moreover, we consider the number of alternatives since the last candidate (ASLC) as in Zwick et al. (2003)[12] as a measure of regret, and we expect to find a role played by this measure in the search behavior dynamics.

[10] http://leen.unice.fr

[11] An alternative is a candidate if it is the highest alternative observed so far.

[12] In which it is called the number of periods since the last candidate (PSLC).

H6: Regret plays a role in the search behavior dynamics.

We expect also a tendency for women to explore more than men since this pattern already appeared in the literature (Eriksson and Strimling, 2009, 2010).

H7: Women tend to explore more than men.

4. Results

In this section, we will systematically conduct both parametric (Two-sided paired two-samples t-tests when paired observations and two-sided unpaired two-samples t-test when non-paired observations) and non-parametric (Two-tailed Wilcoxon signed rank test when paired observations and two-tailed Mann-Whitney U test when non paired observations) statistical comparison tests. When consistent, we will report only parametric tests results. When non consistent, we will report the non parametric tests results in footnote. The adopted significance threshold is 5%.

4.1. Theoretical results

For each of the series used in the experiment[13], we have computed the predictions based on our theoretical model, as well as simulation-based predictions for both treatments. The simulation-based results are presented in the Supplementary materials. These results support the optimal prescription of the theoretical model under uncertainty to spend the minimum time (i.e. one trial) on each explored alternative.

Under Treatment C, to compute the theoretical predictions, we refer to the theoretical thresholds given by equation 2. At each step (or explored alternative), the corresponding theoretical threshold is calculated. Using the same series as in the experiment, we test at each step whether the highest observed alternative so far is higher or equal to the theoretical threshold, in which case the search is stopped and this alternative is selected. Otherwise, the exploration continues. If at the penultimate step (i.e. the 9th step), none of the observed alternatives was higher than the current threshold, then, the highest one is exploited at the last step.

Under Treatment U, we apply a similar algorithm to the one under Treatment C. However, we consider subsets instead of steps, i.e., only the mean at the first click of each alternative is observed. The theoretical thresholds are here given by equation 5. Moreover, the value of the last alternative can be explored unlike in Treatment C. In both treatments, the payoff and the number of explored alternatives are computed.

Table 3 shows the median and average theoretically optimal stopping times (i.e. the total number of explored alternatives) in both treatments, as well as their standard deviations. These results show that stopping should occur later under uncertainty.

[13] In total 2315 series in Treatment C and 2305 series in Treatment U.

	Treatment C (Certainty)	**Treatment U (Uncertainty)**
Median	2	5
Mean	2.35	5.37
Standard deviation	1.48	3.23

Table 3: Mean, median and standard deviation of the optimal stopping times

4.2. Experimental results

Observations from few participants were discarded from the analysis based on two exclusion rules. The first rule is having less than two correct answers out of five at the comprehension questionnaire preceding the main task. The second rule is not recalling and exploiting a "non-candidate" [14] more than 10% of the time in Treatment C (see Figure S15). We excluded in total two participants based on the first rule and nine participants based on the second rule, ending up with sample size of 34 subjects (out of 38) in treatments order CU, and 32 subjects (out of 39) in treatments order UC.

We first compare the stopping times (or number of explored alternatives) observed experimentally between the two treatments C and U. Table 4 shows that the average stopping time is higher in Treatment U. A two-tailed t-test shows that this difference is significant ($t = -19.12$, $p < 0.001$). Figure 3) shows also the comparison with the theoretical prediction. Under both treatments, the average experimentally observed stopping time is significantly different from the theoretical one (Treatment C: $t = 3.05$, $p = 0.002$; Treatment U: $t = -16.89$, $p < 0.001$, Two-tailed paired t-test).

	Treatment C (Certainty)	**Treatment U (Uncertainty)**
Median	2	3
Mean	2.46	3.95
Standard deviation	1.73	3

Table 4: Mean, median and standard deviation of the stopping times from the experimental results

[14] An alternative is a candidate if it is the highest alternative observed so far.

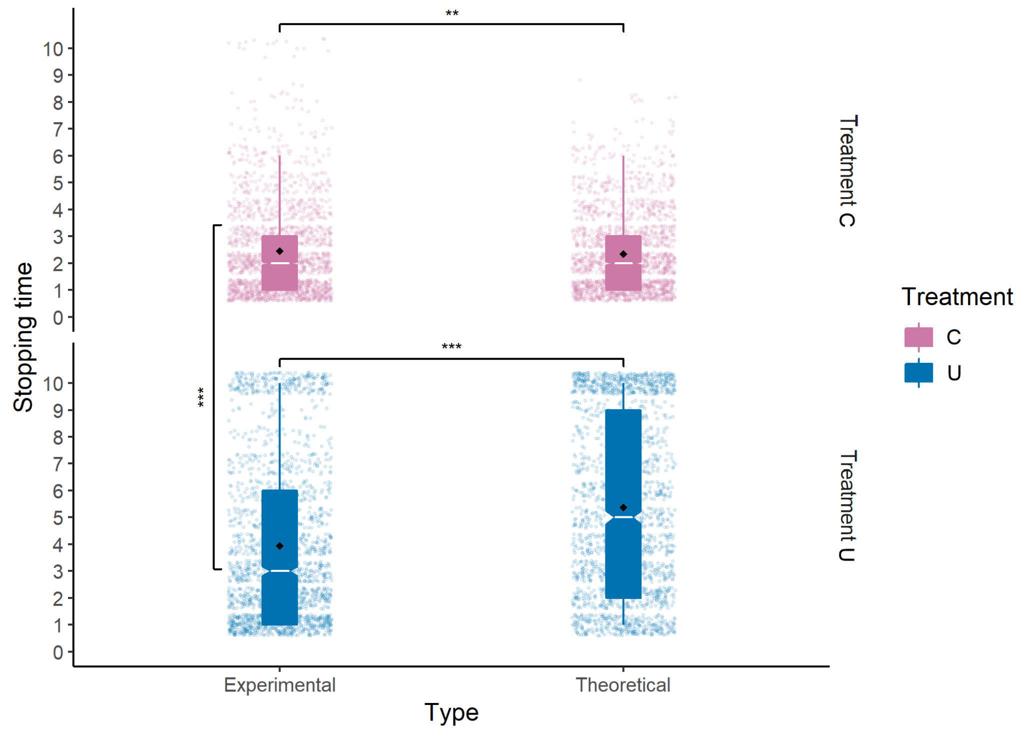


Figure 3: Comparison of the stopping times between the two treatments C and U (Boxplots, Black diamonds: means, Comparisons: Two-tailed paired t-test)
*Note: *p<0.05; **p<0.01; ***p<0.001*

In the following analysis, two main indicators are considered, namely, the *payoff gap* – therefore PG – and the *stopping time gap* – therefore STG. For a given participant i at a given round r, they are defined as follows:

$$PG_{ri} = \frac{\text{Experimentally observed payoff}_{ri}}{\text{Theoretically predicted payoff}_{ri}}$$

$$STG_{ri} = \frac{\text{Experimentally observed stopping time}_{ri}}{\text{Theoretically predicted stopping time}_{ri}}$$

The closer the PG or the STG are to 1, the closer the experimental result of the participant is to the optimal strategy. Values of STG higher than 1 are interpreted as oversearch while values lower than 1 are interpreted as undersearch.

We consider the geometric mean by participant over each of the treatments as well as each of the parts of each treatment. The different parts correspond to the different sequences of rounds: The first part of one Treatment corresponds to the first sequence of 10 consecutive rounds of that treatment, the second part to the second sequence of 10 rounds of that treatment and the last part to the rounds of that treatment in the last 20 rounds. For example, for a participant starting with Treatment C (i.e. playing the treatments order CU, see Table 1):

$$PG_i^{\text{C, part 1}} = \sqrt[10]{\prod_{r=1}^{10} PG_{ri}} \quad ; \quad PG_i^{\text{C, part 2}} = \sqrt[10]{\prod_{r=21}^{30} PG_{ri}} \quad ; \quad PG_i^{\text{C, part 3}} = \sqrt[|F1|]{\prod_{r \in F1} PG_{ri}}$$

$$PG_i^{\text{U, part 1}} = \sqrt[10]{\prod_{r=11}^{20} PG_{ri}} \quad ; \quad PG_i^{\text{U, part 2}} = \sqrt[10]{\prod_{r=31}^{40} PG_{ri}} \quad ; \quad PG_i^{\text{U, part 3}} = \sqrt[|F2|]{\prod_{r \in F2} PG_{ri}}$$

$$PG_i^{\text{C}} = \sqrt[|E1|]{\prod_{r \in E1} PG_{ri}} \quad ; \quad PG_i^{\text{U}} = \sqrt[|E2|]{\prod_{r \in E2} PG_{ri}}$$

where:
$F1 = \{41 \leq r \leq 60, \quad r \in \text{Treatment C}\}$,
$F2 = \{41 \leq r \leq 60, \quad r \in \text{Treatment U}\}$,
$E1 = \{1 \leq r \leq 60, \quad r \in \text{Treatment C}\}$,
$E2 = \{1 \leq r \leq 60, \quad r \in \text{Treatment U}\}$.

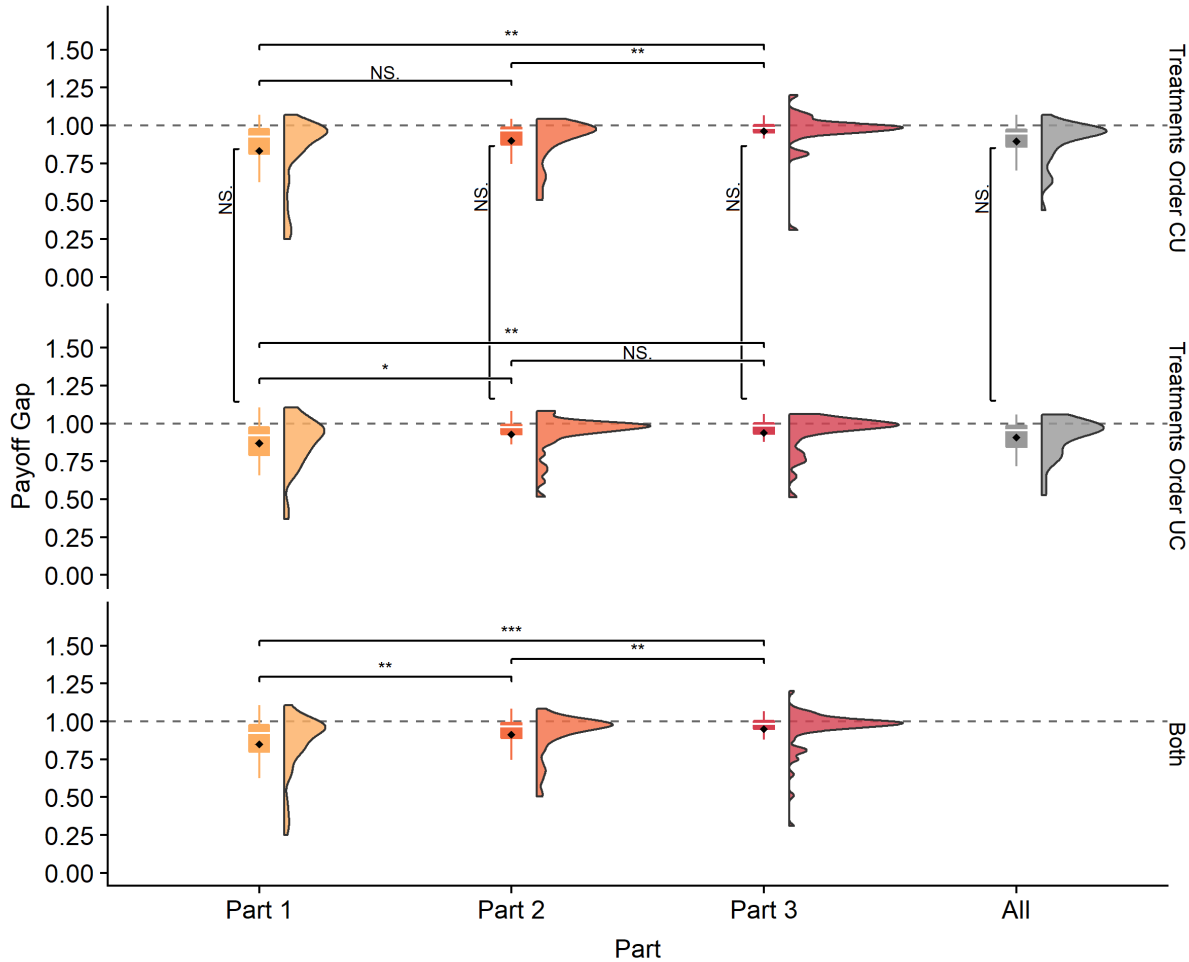
4.2.1. Experimental results of Treatment C

First, we want to know whether participants oversearch in Treatment C (**H1**). To do so, we pool all the rounds of Treatment C at the subject level, and test whether the mean stopping time gap of all the participants is greater than 1. We perform a two-tailed, one-sample t-test and we find that this value is significantly greater than 1 (*Mean*= 1.11, $t = 2.77$, $p = 0.007$). This confirms our hypothesis **H1**.

Learning dynamics

Figure 4 shows the distribution of the participants' payoff gap in Treatment C. The first three columns report the results for each part (part 1, part 2 and part 3), for each of the treatments order (CU and UC) and for both of them together, in one row each. The last column reports the aggregate results of each of the two treatments orders. Paired t-tests are performed between the different parts for the same groups, while non paired t-tests are performed between the same parts of the two treatments orders. These comparisons show a strong significant difference between the first and the last part, where in the last part, the mean payoff gap is higher. This significant difference is observed both when considering the two treatments orders together or separately. This can be interpreted as a learning effect, where participants' performance gets closer to the optimal theoretical prediction through the game. This learning seems faster in treatments order UC, since a significant difference is already observed between the first two parts, while this is not the case in treatments order CU. This difference can be interpreted as the result of a longer practice. Indeed, those who started with Treatment U have an advantage since they have practiced 10 rounds more (from the other treatment though) than those who started with Treatment C.

Figure 5 reports equivalently the stopping time gap. This gap is significantly higher in the last part compared to the first part in the aggregate results of both treatment orders (last row). Again, this can be interpreted as a learning effect where the exploration amount of the participants gets closer to the optimal prediction as the game progresses. However, when considering the two treatment orders separately, we do not observe this learning effect when the participants start with treatment C. In general, these results confirm our hypothesis **H3**.



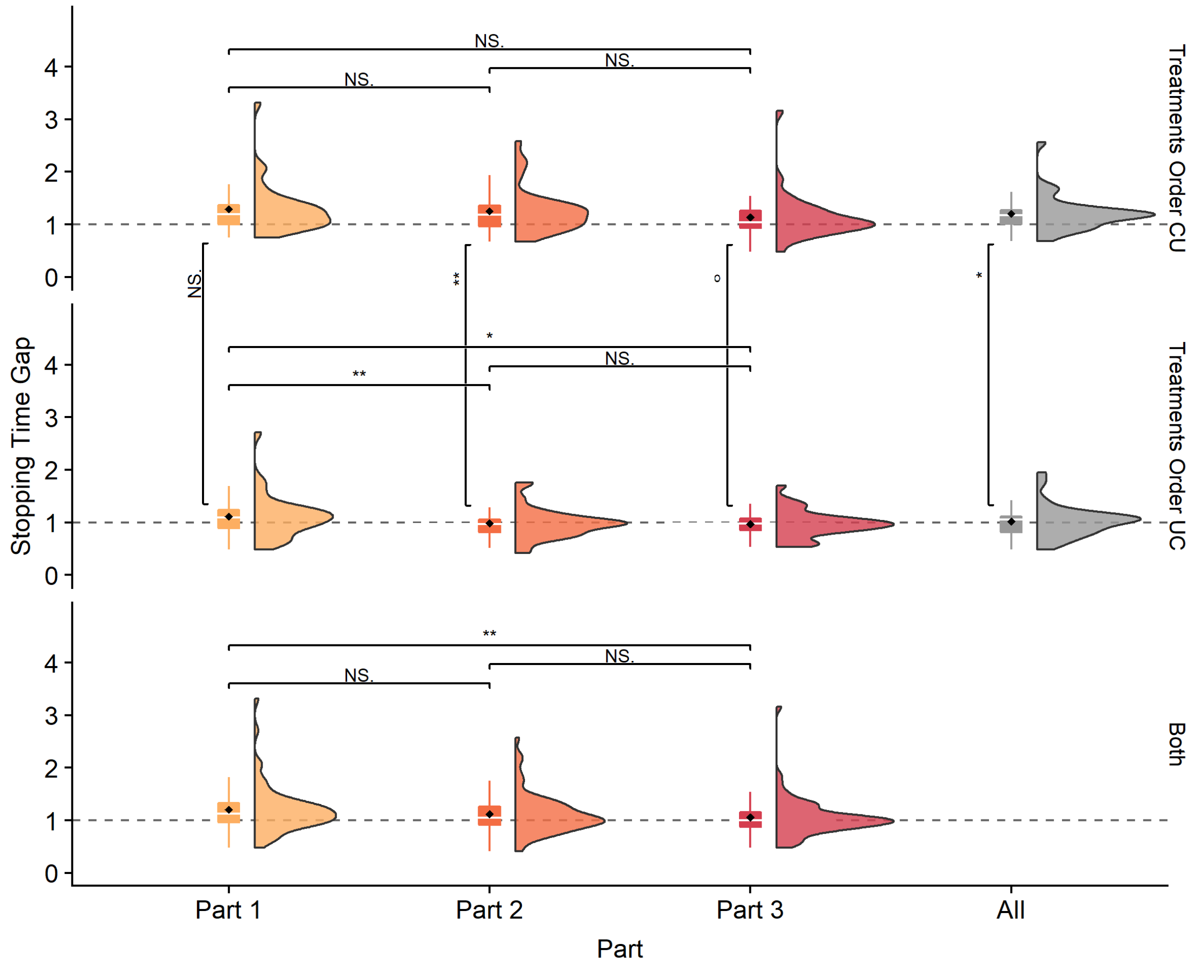
		Part 1	Part 2	Part 3	All
Treatments order CU	*Median*	0.926	0.966	0.985	0.947
	Mean	0.832	0.900	0.962	0.895
Treatments order UC	*Median*	0.921	0.976	0.986	0.954
	Mean	0.870	0.928	0.938	0.909
Both	*Median*	0.924	0.968	0.985	
	Mean	0.850	0.913	0.950	

Figure 4: Payoff gap's evolution through the different parts in Treatment C by treatments order

(a). Boxplots (Black diamonds: mean) and Kernel probability density plots. Horizontal comparisons: Two-tailed paired t-test; Vertical comparisons: Two-tailed unpaired t-test

Note: $^{NS.}p \geq 0.1$; $^{\circ}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$

(b). Medians and means



		Part 1	Part 2	Part 3	All
Treatments order CU	*Median*	1.197	1.184	1.042	1.172
	Mean	1.290	1.252	1.132	1.206
Treatments order UC	*Median*	1.090	0.964	0.973	1.052
	Mean	1.111	0.982	0.972	1.019
Both	*Median*	1.122	1.045	1.000	
	Mean	1.203	1.121	1.054	

Figure 5: Stopping time gap's evolution through the different parts in Treatment C by treatments order
(a). Boxplots (Black diamonds: mean) and Kernel probability density plots. Horizontal comparisons: Two-tailed paired t-test; Vertical comparisons: Two-tailed unpaired t-test
Note: $^{NS.}p \geq 0.1$; $^{\circ}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$
(b). Medians and means

We next fit a generalized linear mixed-effects model by maximum likelihood to explain the individual payoff gap and the individual stopping time gap in Treatment C. We fit the models separately for each treatment. We consider the following models:

Model 1a :

$$\begin{aligned} PG_{ri} = \beta_0 + u_{0i} + (\beta_{Part} + u_{Part\,i})Part_{ri} + \beta_{UC}UC_i \\ +\beta_{BRET}BRET_i + \beta_{SART}SART_i + \epsilon_{ri} \end{aligned}$$

Model 1b :

$$\begin{aligned} PG_{ri} = \beta_0 + u_{0i} + (\beta_{Part} + u_{Part\,i})Part_{ri} + \beta_{UC}UC_i \\ +\beta_{BRET}BRET_i + \beta_{SART}SART_i \\ +\beta_{Age}Age_i + \beta_{Male}Male_i + \epsilon_{ri} \end{aligned}$$

where:
PG_{ri} is the payoff gap at round r of player i,
$Part$ is a categorical variable corresponding to the part in the treatment (first as reference, second or third),
UC is a dummy variable being 1 if the Treatment U is played first,
$BRET$ is the percentage of boxes collected at the BRET (number of boxes collected divided by the total number of boxes) ,
$SART$ is the percentage of commission errors made at the SART (number of errors made divided by the maximum possible number of errors),
Age is the age class
$Male$ is a dummy variable corresponding to being a male player,
ϵ_{ri} is the error term,
the random effects u_{0i} and $u_{Part\,i}$.

We also consider the models 2a and 2b, similar to 1a and 1b but where the explained variable is STG_{ri}, which represents the stopping time gap at round r of player i.

The results of the regressions are presented in Table 5. The results show a positive and highly significant effect of the part number on the payoff gap, while a negative significant effect on the stopping time gap. This supports the learning effect suggested above: As the game progresses, the relative performance of the players compared to the theoretically optimal one is higher, while the *oversearch* decreases. This again confirms our hypothesis **H3**. The treatments order does not have any significant effect on the payoff gap, but does have a negative significant effect on the stopping time gap. Moreover, the results show a positive significant effect on the payoff gap of the percentage of boxes collected at the BRET. This means that risk loving participants play closer to the theoretical prediction. Finally, being a male has also a positive significant effect on the payoff gap.

Generalized Linear Mixed Model	*Dependent variable:*							
	Payoff Gap				Stopping Time Gap			
	(1a)		(1b)		(2a)		(2b)	
	coef	p-value	coef	p-value	coef	p-value	coef	p-value
Playing the second part	**0.062**	< 0.0001	**0.062**	< 0.0001	**−0.119**	0.0194	**−0.119**	0.0194
	(0.015)	***	(0.015)	***	(0.051)	*	(0.051)	*
Playing the third part	**0.096**	< 0.0001	**0.096**	< 0.0001	**−0.228**	0.0006	**−0.227**	0.0006
	(0.022)	***	(0.022)	***	(0.063)	***	(0.063)	***
Playing the Treatment U first	0.002	0.9471	0.004	0.8914	**−0.239**	0.0124	**−0.240**	0.0104
	(0.027)		(0.026)		(0.093)	*	(0.091)	*
% Boxes collected at BRET	0.140	0.0762	**0.173**	0.0265	−0.005	0.9850	−0.075	0.7779
	(0.078)	°	(0.076)	*	(0.264)		(0.264)	
% Errors at SART	−0.031	0.5201	−0.016	0.7265	0.190	0.2527	0.149	0.3614
	(0.048)		(0.047)		(0.164)		(0.162)	
Age class			−0.019	0.2292			0.055	0.3219
				(0.016)			(0.055)	
Being a male			**0.065**	0.0220			−0.152	0.1199
			(0.028)	*			(0.096)	
Constant	**0.831**	< 0.0001	**0.821**	< 0.0001	**1.360**	< 0.0001	**1.363**	< 0.0001
	(0.037)	***	(0.049)	***	(0.124)	***	(0.169)	***
Observations	1,989		1,989		1,989		1,989	
Subjects	66		66		66		66	
Log Likelihood	−14.124		−11.213		−2,528.430		−2,527.060	
Akaike Inf. Crit.	48.249		46.426		5,076.860		5,078.119	
Bayesian Inf. Crit.	104.202		113.570		5,132.814		5,145.264	

Note: °$p<0.1$; *$p<0.05$; **$p<0.01$; ***$p<0.001$

Table 5: Generalized linear mixed model of the payoff gap and the stopping time gap in Treatment C
(standard errors in brackets)

Behavioral analysis

To go further, we perform a survival analysis in order to study the determinants of the "exploit" decision. We do not consider all the trials, but only the first trial for each alternative (except the very first trial). Indeed, the player does not have the choice between exploring and exploiting for her first trial: she can only start exploring the first alternative. Moreover, in Treatment C, once the player starts exploring an alternative, she is forced to spend a totality of 10 trials on it. Thus, the decision (between exploring a new alternative and exploiting one) can only be made at trials 11, 21, 31, 41, 51, 61, 71, 81 and 91, provided that the player did not exploit.

We include also measures of anticipation and regret: the number of encountered candidates (where a candidate is an alternative that is the highest one observed so far) as a measure of anticipation, and the number of explored alternatives since the last candidate was encountered – that we will call "*alternatives since last candidate*" ($ASLC$) – as a measure of regret. These covariates are computed at each trial. We also consider the payoff of the previous round. Finally, we control for the individual characteristics (risk aversion level, impulsivity level, gender and age).

We fit the following Cox proportional hazards model:

$$\begin{aligned}&\lambda\left(t, C_{ri}(t), ASLC_{ri}(t), Time_{ri}, \pi_{ri}, Part_{ri}, UC_{ri}, SART_i, BRET_i, Male_i, Age_i\right)\\=\;&\lambda_0\left(t\right) exp\left(t, C_{ri}(t), ASLC_{ri}(t), Time_{ri}, \pi_{ri}, Part_{ri}, UC_{ri}, SART_i, BRET_i, Male_i, Age_i\right)\end{aligned}$$

where:
$C_{ri}(t)$ is the number of candidate encountered by player i at round r, at trial t,
$ASLC_{ri}(t)$ is the $ASLC$ of player i at round r, at trial t,
$Time_{ri}$ is the logarithm of the average time by click player i at round r,
π_{ri} is the payoff of player i at round $r-1$ [15],
$Part_{ri}$ is for player i at round r, a categorical variable corresponding to the part in the treatment (first as reference, second or third),
UC_{ri} is for player i at round r, a dummy variable being 1 when the Treatment U is played first,
$SART_i$ is the percentage of commission errors made at the SART of player i,
$BRET_i$ is the percentage of boxes collected at the BRET of player i,
$Male_i$ is whether player i is a male player,
Age_i is the age class of player i.

The results are presented in Table 6 and are illustrated by a forest plot in Figure 6. A hazard ratio above 1 indicates a covariate that is positively associated with the "exploit" decision probability, and thus negatively associated with the duration of exploration phase. They show a negative and strongly significant effect of regret and anticipation measures on the probability of "exploit" decision, leading to a longer exploration. Contrarily, this probability is significantly increased when being at the second or the last part of the treatment, and when starting with Treatment U. The impulsivity level, given by the percentage of commission errors made at the SART has a negative effect on the probability to exploit, but which is only significant at the 10% level. In line with this result, faster clicks, which could be interpreted as engaging less deliberation, are significantly related to more exploration. We also find a positive effect of being a male on the probability to "exploit", meaning that women tend to explore more than men, but which is only significant at the 10% level. The risk aversion level does not show any significant effect. Finally, we were expecting that previous round's payoff would contribute to reinforce learning, where search strategies yielding high payoffs would tend to be replicated and *reinforced*. Counter-intuitively, we find that the previous round payoff significantly increases the exploration amount. To summarize, these results allow us to confirm the hypotheses **H5**, **H6** and **H7**.

[15] We fit the model only for $r > 1$

Cox regression model	*Time to "exploit" decision*			
	coef	HR = exp(coef)	95% CI	p-value
Number of candidates	-1.64 (0.0984)	**0.19**	[0.16, 0.24]	< 0.0001 ***
ASLC	-1.26 (0.0942)	**0.28**	[0.24, 0.34]	< 0.0001 ***
Log(average time by click)	0.43 (0.0461)	**1.53**	[1.40, 1.68]	< 0.0001 ***
Previous round payoff (normalized)	-0.35 (0.1283)	**0.70**	[0.55, 0.90]	0.0060 **
Playing the second part	0.30 (0.0642)	**1.35**	[1.18, 1.53]	< 0.0001 ***
Playing the third part	0.27 (0.0612)	**1.31**	[1.17, 1.48]	< 0.0001 ***
Playing Treatment U first	0.13 (0.0623)	**1.14**	[1.01, 1.29]	0.0381 *
% Errors at the SART	−0.18 (0.1032)	0.84	[0.68, 1.02]	0.0829 °
% Boxes collected in the BRET	0.09 (0.1564)	1.09	[0.81, 1.49]	0.5651
Being a male	0.11 (0.0652)	1.12	[0.98, 1.27]	0.0867 °
Age class	-0.04 (0.0322)	0.96	[0.91, 1.03]	0.2544
n	4779			
number of events	1940			
number of subjects	66			
R^2	0.318			
Max. Possible R^2	0.995			
Log Likelihood	−11,922.950			
Concordance	0.952	se = 0.005		
Likelihood ratio Test	1828	p-value< 0.0001	(df=11)	
Wald Test	1129	p-value< 0.0001	(df=11)	
Score (Logrank) Test	1813	p-value< 0.0001	(df=11)	
Robust	64.63	p-value< 0.0001		

Note: $^{\circ}p<0.1$; $^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

Table 6: Cox regression model of time to exploit in Treatment C with robust standard errors (in brackets).

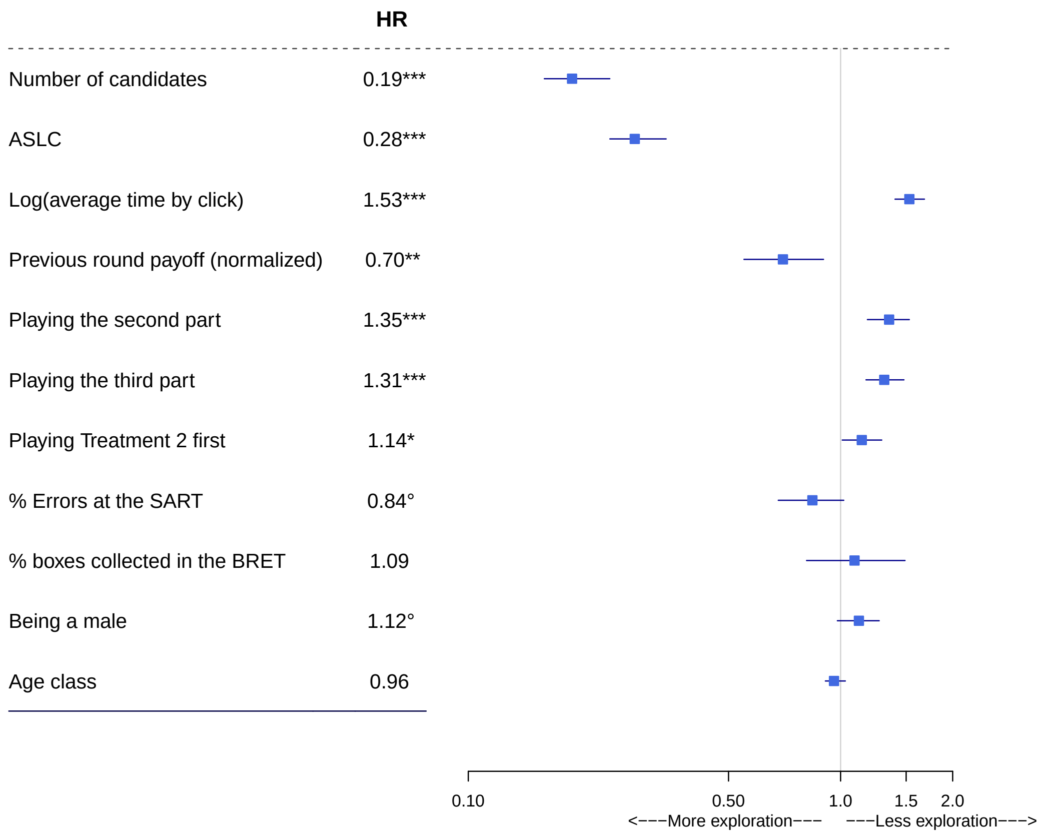


Figure 6: Forest plot of Cox regression model of time to exploit in Treatment C with robust standard errors
Note: $^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

Discussion

As predicted, we observe in Treatment C (certain information) a tendency of the players to oversearch. Though, participants learn over time to search less and get closer to the optimal search amount. Additionally, the results support the hypothesis that abundance of candidates interpreted as a measure of anticipation tends to increase the exploration. They also support the hypothesis that regret has a role in the search behavior dynamics. Similarly to anticipation, they show that it increases exploration. The results fail however to confirm the intuition that the previous round's payoff would play a reinforcing effect, helping the participants getting closer to the optimal search amount. Indeed, we observe the reverse effect, that is increasing exploration and consequently oversearch. Also as expected, we find a gender effect where women tend to explore more than men.

Regarding the effect of risk aversion level, our results are mixed. Indeed, we find that lower risk aversion increases the relative performance compared to the optimal theoretical prediction, but do not have a strong significant effect on relative stopping time or on the probability to explore a new alternative. For impulsivity, it does not show any strong significant effect on any of these.

4.2.2. Experimental results of Treatment U

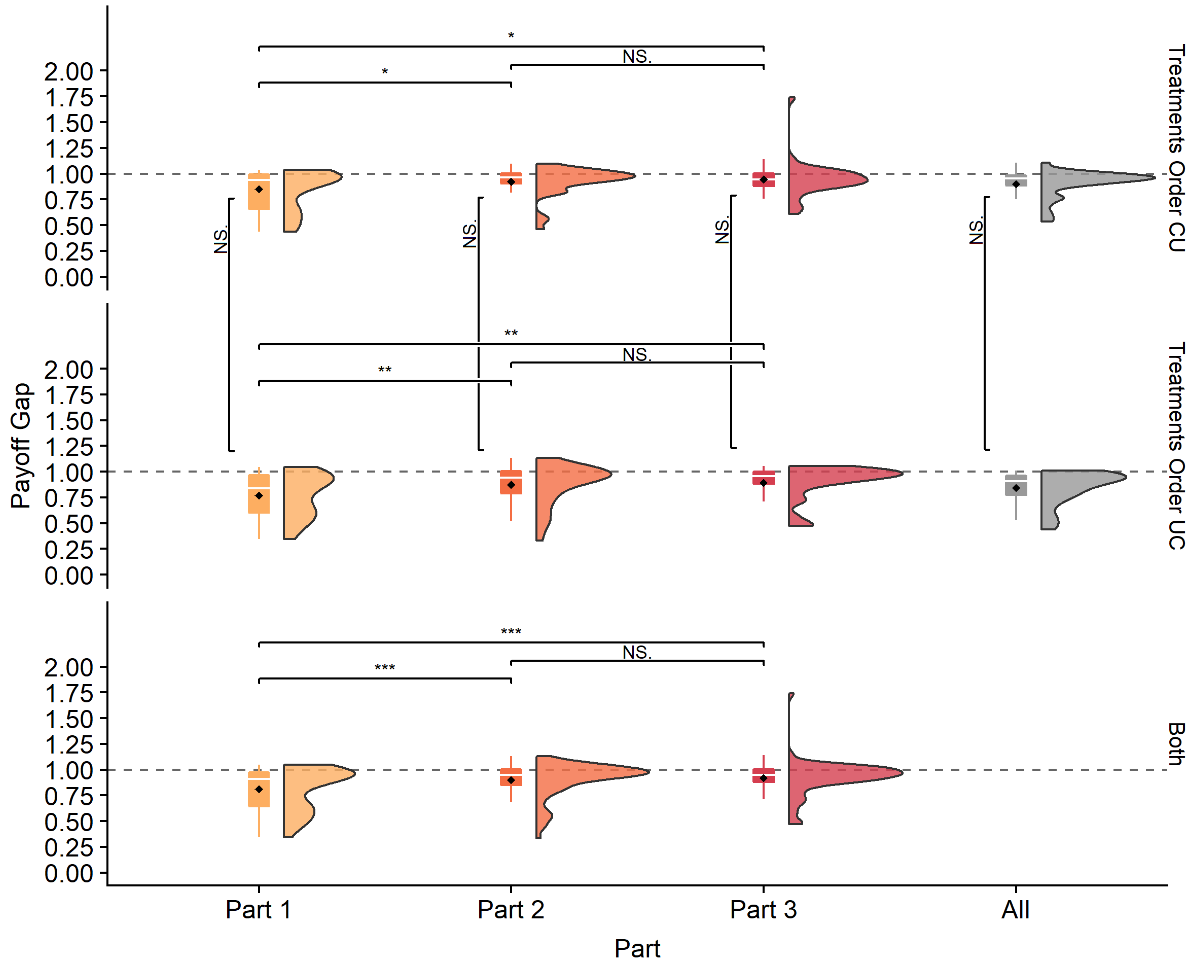
We first want to know whether participants undersearch in Treatment U (**H2**). To do so, we pool all the rounds of Treatment U at the subject level, and test whether the mean stopping time gap of all the participants is lower than 1. We perform a two-tailed, one-sample t-test and we find that this value is significantly lower than 1 ($Mean = 0.886$, $t = -2.404$, $p = 0.019$). We then confirm **H2**.

Learning dynamics

Like for Treatment C, we plot the distributions of the participants' payoff gap (see Figure 7) and stopping time gap (see Figure 8) in Treatment U and we perform within (paired t-tests) and between (non paired t-tests) comparisons.

As in Treatment C, we find an overall learning effect where participants average relative payoff compared to the theoretical prediction is significantly higher in the last part of the treatment compared to its beginning (see Figure 7). We do not observe any order effect.

As for the stopping time gap (see Figure 8), results show a tendency to undersearch compared to the optimal search amount. As in Treatment C, significant differences are observed when considering both treatments orders together. However, the participants average search amount moves away from the optimal one, since the average stopping time gap gets further from 1. As previously noticed in Treatment C, when considering the two treatment orders separately, no significant difference is found in treatment order CU. These results confirm our hypothesis **H4**. Indeed, participants' relative performance tend to improve across the game, but their search amount does not converge towards the optimal one.



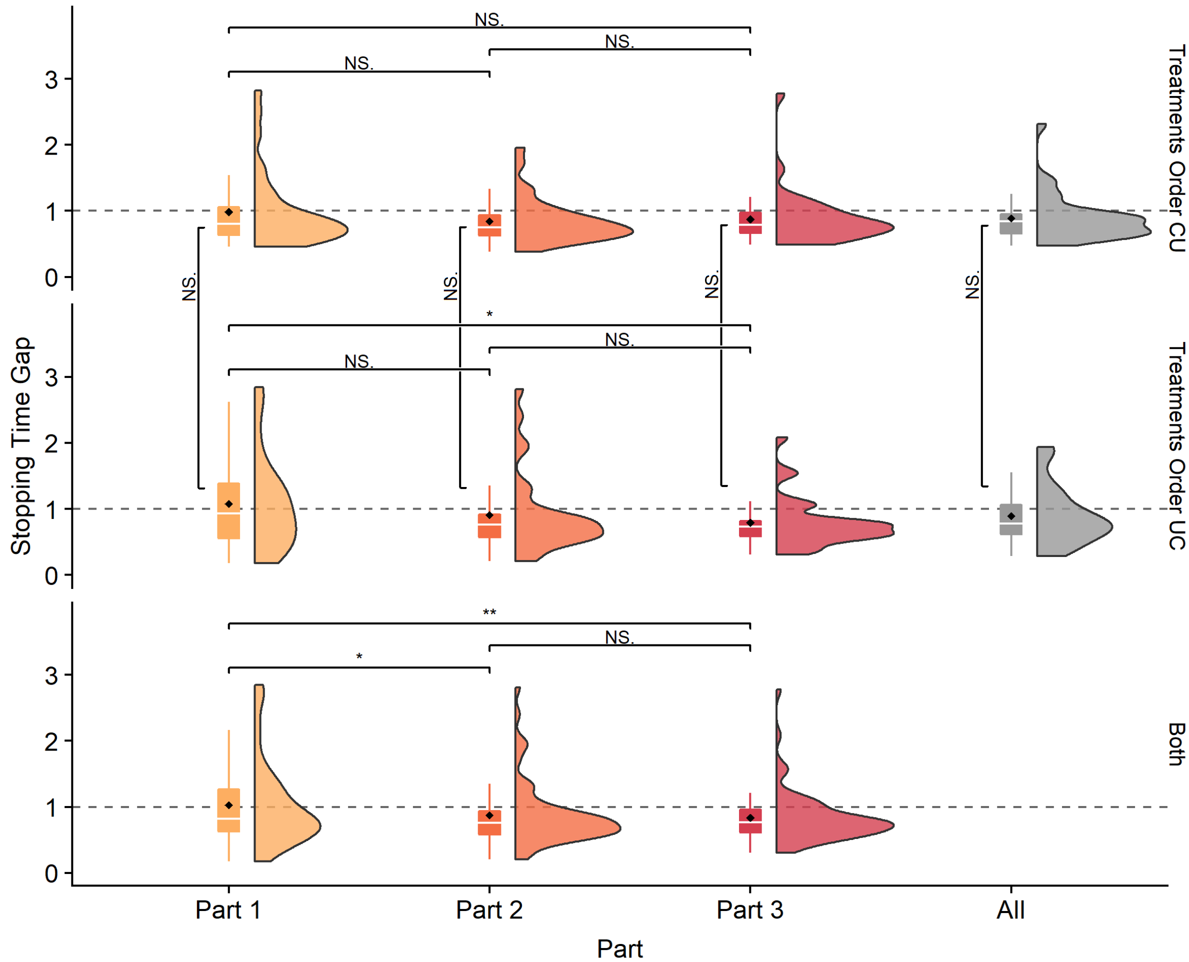
		Part 1	Part 2	Part 3	All
Treatments order CU	*Median*	0.940	0.966	0.943	0.951
	Mean	0.850	0.922	0.946	0.901
Treatments order UC	*Median*	0.837	0.943	0.954	0.907
	Mean	0.770	0.874	0.891	0.841
Both	*Median*	0.907	0.948	0.949	
	Mean	0.811	0.899	0.919	

Figure 7: Payoff gap's evolution through the different parts in Treatment U by treatments order

(a). Boxplots (Black diamonds: mean) and Kernel probability density plots. Horizontal comparisons: Two-tailed paired t-test; Vertical comparisons: Two-tailed unpaired t-test

Note: $^{NS.}p\geq 0.1$; $^{\circ}p<0.1$; $^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

(b). Medians and means



		Part 1	Part 2	Part 3	All
Treatments order CU	*Median*	0.801	0.752	0.786	0.845
	Mean	0.985	0.843	0.877	0.888
Treatments order UC	*Median*	0.928	0.760	0.728	0.780
	Mean	1.077	0.907	0.789	0.895
Both	*Median*	0.819	0.760	0.764	
	Mean	1.029	0.874	0.834	

Figure 8: Stopping time gap's evolution through the different parts in Treatment U by treatments order

(a). Boxplots (Black diamonds: mean) and Kernel probability density plots. Horizontal comparisons: Two-tailed paired t-test; Vertical comparisons: Two-tailed unpaired t-test

Note: $^{NS.}p \geq 0.1$; $^{\circ}p<0.1$; $^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

(b). Medians and means

We estimate the generalized linear mixed-effects models 1a, 1b, 2a and 2b by maximum likelihood to explain the individual payoff gap and the individual stopping time gap in Treatment U. The results are presented in Table 7. They show like in Treatment C a positive significant effect of being at the second or last part in the treatment for the payoff gap, and a negative significant effect for the stopping time gap. The first reflects a learning effect in terms of performance. The second reflects a decrease in the amount of search as the game progresses. Starting the experiment with Treatment U does not have any significant effect. When controlling for the age and the gender, the percentage of boxes collected at the BRET shows a positive effect on the payoff gap, only at the 10% significance level. The percentage of errors made in the SART shows here a positive effect on the stopping time gap, meaning that the higher the impulsivity level, the higher the observed amount of search, but this effect is only significant at the 10% level. We note also that the age has a negative significant effect on the payoff gap. Finally, being a male shows a positive effect on the payoff effect, and negative effect on the stopping time gap, but again, only significant at the 10% level.

Generalized Linear Mixed Model	*Dependent variable:*							
	Payoff Gap				Stopping Time Gap			
	(1a)		(1b)		(2a)		(2b)	
	coef	*p*-value	coef	*p*-value	coef	*p*-value	coef	*p*-value
Playing the second part	**0.078** (0.029)	0.0082 **	**0.078** (0.029)	0.0083 **	**−0.219** (0.076)	0.0042 **	**−0.219** (0.076)	0.0042 **
Playing the third part	**0.100** (0.035)	0.0056 **	**0.101** (0.035)	0.0055 **	**−0.292** (0.100)	0.0048 **	**−0.293** (0.100)	0.0046 **
Playing the Treatment U first	−0.055 (0.039)	0.1578	−0.055 (0.037)	0.1424	−0.044 (0.124)	0.7244	−0.057 (0.121)	0.6422
% Boxes collected at BRET	0.177 (0.115)	0.1176	0.205 (0.108)	0.0617 °	−0.233 (0.349)	0.5159	−0.353 (0.356)	0.3253
% Errors at SART	−0.064 (0.069)	0.3534	−0.046 (0.066)	0.4847	0.419 (0.219)	0.0602 °	0.384 (0.215)	0.0794 °
Age class			**−0.052** (0.022)	0.02217 *			−0.029 (0.072)	0.6877
Being a male			0.074 (0.039)	0.0642 °			−0.214 (0.128)	0.0995 °
Constant	**0.837** (0.052)	< 0.0001 ***	**0.886** (0.068)	< 0.0001 ***	**1.149** (0.175)	< 0.0001 ***	**1.360** (0.231)	< 0.0001 ***
Observations	1,971		1,971		1,971		1,971	
Subjects	66		66		66		66	
Log Likelihood	−1,462.496		−1,458.916		−3,242.133		−3,240.491	
Akaike Inf. Crit.	2,944.991		2,941.833		6,504.265		6,504.981	
Bayesian Inf. Crit.	3,000.854		3,008.868		6,560.128		6,572.017	

Note: °$p<0.1$; *$p<0.05$; **$p<0.01$; ***$p<0.001$

Table 7: Generalized linear mixed model of the payoff gap and stopping time gap in Treatment U
(standard errors in brackets)

We look in Figure S16 at the average number of clicks by alternative for each player, and its evolution across the different parts of the treatment. Even though the players spent relatively few trials on the explored alternatives (around 3 trials), the figure does not highlight any overall decreasing or increasing trend that could be interpreted as a learning effect.

Behavioral analysis

We also perform a survival analysis to study the determinants of the "exploit" decision. Here, we consider all the trials (except the first one for which there is no decision to make), since unlike in Treatment C, the player is not forced to spend a certain number of trials exploring an alternative.

We fit the following Cox proportional hazards model:

$$\begin{aligned}\lambda\,(t, C_{ri}(t), A_{ri}(t), ASLC_{ri}(t), Time_{ri}, MSE_{ri}(t), \pi_{ri}, Part_{ri}, UC_{ri},\\ SART_i, BRET_i, Male_i, Age_i)\\ = \lambda_0\,(t)\ exp\,(C_{ri}(t), A_{ri}(t), ASLC_{ri}(t), Time_{ri}, MSE_{ri}(t), \pi_{ri}, Part_{ri}, UC_{ri},\\ SART_i, BRET_i, Male_i, Age_i)\end{aligned}$$

where:

$C_{ri}(t)$ is the number of candidates encountered by player i at round r, at the trial t,

$A_{ri}(t)$ is the number of explored alternative by player i at round r, at the trial t[16],

$ASLC_{ri}(t)$ is the $ASLC$ of player i at round r, at the trial t,

$Time_{ri}$ is the logarithm of the average time by click player i at round r,

$MSE_{ri}(t)$ is mean standard error of the clicked alternative by player i at round r, at the trial t,

π_{ri} is the payoff of player i at round $r-1$[17],

$Part_{ri}$ is for player i at round r, a categorical variable corresponding to the part in the treatment (first as reference, second or third),

UC_{ri} is for player i at round r, a dummy variable being 1 when the Treatment U is played first,

$SART_i$ is the percentage of commission errors made at the SART of player i,

$BRET_i$ is the percentage of boxes collected at the BRET of player i,

$Male_i$ is whether player i is a male player,

Age_i is the age class of player i.

The results presented in Table 8 and Figure 9 show, as in Treatment C, a strong negative effect of both regret and anticipation measures (number of candidates and $ASLC$) on the probability to "exploit", which means that they delay exploitation. Click speed has also the same type of effect as in Treatment C, where it is positively associated with the probability to exploit. Also as in Treatment C, progress in the game (i.e. playing the second or the last part of the treatment) leads to a sooner exploitation. Though, the treatments order does not show any significant effect. We find again a positive effect of being a male on the probability to "exploit", meaning that women tend to explore more than men, as well as a a negative effect of age, meaning that older participants tend to

[16] We control for the number of explored alternative in Treatment U and not in Treatment C since in Treatment C, the number of explored alternative is equal to t (as long as the "exploit" event is not yet observed).

[17] We fit the model only for $r > 1$

explore more, but which is only significant at the 10% level. Contrarily to Treatment C, treatments order is not any more significant. The risk aversion and impulsivity variables do not show any significant effect on the duration of the exploration. Finally, we find again a counter-intuitive significant effect of the previous round's payoff, which is a positive effect on the probability to exploit, thus decreasing the search amount. To summarize, these results allow us to confirm the hypotheses **H5**, **H6** and **H7**.

Cox regression model	Time to "exploit" decision			
	coef	HR = exp(coef)	95% CI	p-value
Number of candidates	-0.09 (0.0368)	**0.91**	[0.85, 0.98]	0.0116 *
Number of explored alternatives	0.07 (0.0208)	**1.07**	[1.03, 1.12]	0.0011 **
ASLC	-0.19 (0.0314)	**0.83**	[0.78, 0.88]	< 0.0001 ***
Log(average time by click)	0.84 (0.0828)	**2.33**	[1.98, 2.73]	< 0.0001 ***
Mean standard error of the clicked alternative	-0.02 (0.0042)	**0.98**	[0.97, 0.99]	< 0.0001 ***
Previous round payoff (normalized)	0.89 (0.2090)	**2.43**	[1.61, 3.66]	< 0.0001 ***
Playing the second part	0.43 (0.0918)	**1.54**	[1.29, 1.85]	< 0.0001 ***
Playing the third part	0.56 (0.1010)	**1.75**	[1.44, 2.13]	< 0.0001 ***
Playing Treatment U first	0.09 (0.1312)	1.09	[0.84, 1.41]	0.5156
% Errors at the SART	-0.02 (0.2429)	0.98	[0.61, 1.58]	0.9491
% Boxes collected in the BRET	-0.18 (0.3009)	0.84	[0.45, 1.55]	0.5767
Being a male	0.26 (0.3009)	**1.30**	[1.01, 1.67]	0.0.381 *
Age class	-0.11 (0.3009)	0.89	[0.79, 1.01]	0.0692 °
n	18726			
number of events	1939			
number of subjects	66			
R^2	0.046			
Max. Possible R^2	0.744			
Log Likelihood	−12,306.870			
Concordance	0.716	se = 0.013		
Likelihood ratio Test	873.8	p-value< 0.0001	(df=13)	
Wald Test	289.5	p-value< 0.0001	(df=13)	
Score (Logrank) Test	880.6	p-value< 0.0001	(df=13)	
Robust	57.44	p-value< 0.0001		

Note: $^{\circ}p<0.1$; $^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

Table 8: Cox regression model of time to exploit in Treatment U with robust standard errors (in brackets).

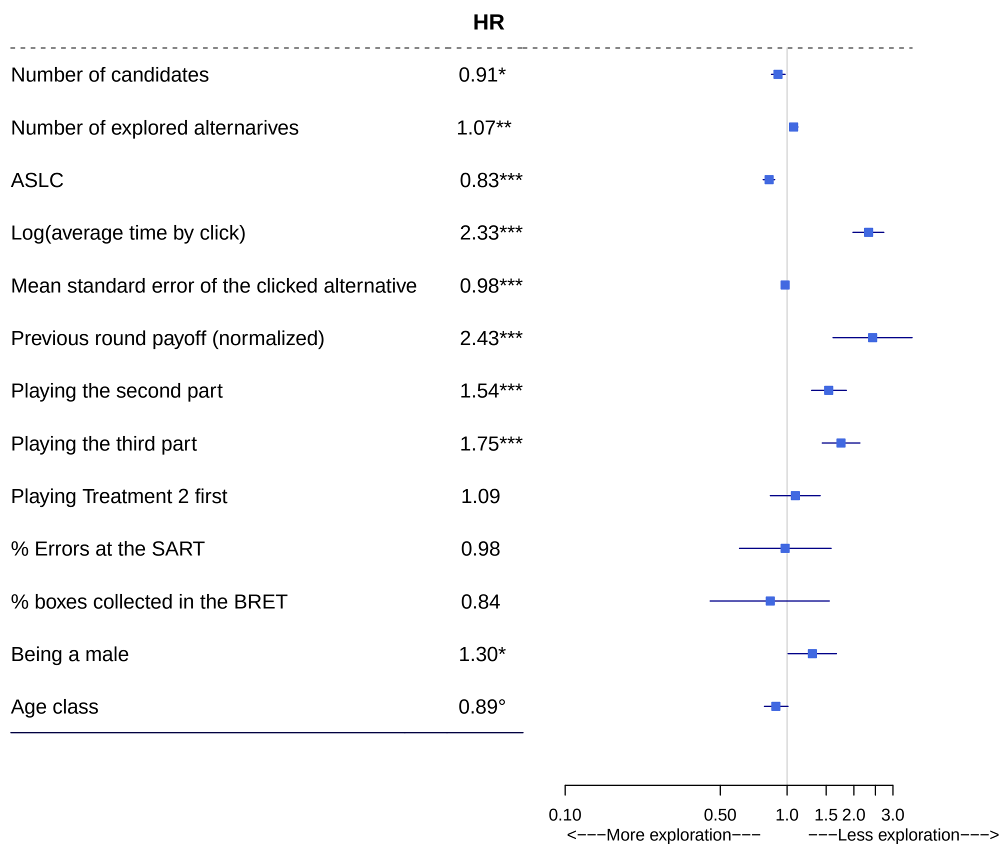


Figure 9: Forest plot of Cox regression model of time to exploit in Treatment U with robust standard errors
Note: $^{\circ}p<0.1$; $^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

Discussion

As predicted, we find evidence for a tendency to undersearch in Treatment U (uncertain information), and a much weaker or even absent learning effect compared to Treatment C. We also observe as expected a positive effect of anticipation on the amount of search, and a role of regret in the dynamics of the search behavior. This role consists in boosting the exploration. Again, as in Treatment C, we fail to find a reinforcing effect of the previous round's payoff, where higher previous payoffs would help bringing the participants closer to the optimal prediction. At the opposite, we find the reverse effect, where it decreases the exploration, thus accentuating the undersearch. In this treatment we also observe in accordance with our hypothesis a gender effect where men tend to explore less and perform better relative to the optimal strategy. Moreover, we find that younger participants tend to perform better relative to the optimal strategy. Contrarily to our intuition, we do not find any strong significant effect of risk aversion and impulsivity on the performance or on the search amount.

5. General discussion

We consider an optimal stopping problem with iterative exploitation and possible recall. We manipulate the degree of uncertainty at the level of the alternatives values. We present the analytic solution, which can be described as a decreasing threshold search strategy, meaning that the minimum value that the individual should be ready to accept decreases over time. The results of this theoretical model show that search in terms of number of explored alternatives is longer under uncertainty. When the information on the value of the alternative is certain, the exploration of a new alternative can be viewed as costly, since it requires no less than 10 trials to be subtracted from the exploitation phase, compared to the uncertain information condition, in which the exploration can be realized by less trials (though at the price of a higher uncertainty).

We next implement the model experimentally. We find, in line with the existing experimental literature on optimal stopping problems discussed earlier in this paper, that participants tend to oversearch and learn better when exploration is costly, and to undersearch when exploration is relatively cheap.

We find that anticipation tend to increase exploration and regret also plays a role in the search dynamics. Contrarily to Zwick et al. (2003) who find a negative effect of regret on exploration, the effect that we find is the opposite. Though, this negative effect was found in a framework *à la Secretary*. However, this framework and the one we are studying, which involves iterative exploration, have already experimentally exhibited several fundamental differences. Thus, it is not aberrant to observe such an opposite effect of regret under the two frameworks. As for the interpretation of such an effect in our setting, a possible explanation could be a resurrection sake strategy, where after several attempts to find a new candidate, one would want to recover her losses and would *gamble for resurrection*[18]. Moreover, in our setting, the possibility of recall makes exploration somehow much safer than it would be in other settings.

We observe consistently with Eriksson and Strimling (2009) and Eriksson and Strimling (2010) a gender effect where women tend to explore more than men, which however does not seem to be predicted by any known cognitive gender differences. A possible justification could lie in a difference in loss aversion levels (that would have been masked by the gender). Indeed, the literature has shown that women tend to be more loss averse that men (Crosetto and Filippin, 2016). At the same time, some studies on optimal stopping search have found that loss aversion plays a role (Schunk and Winter, 2009; Schunk, 2009). For example, Schunk and Winter (2009) find that search heuristics are not related to measures of risk aversion, but to measures of loss aversion. However, contradictory results are observed as for the direction of this effect. Indeed, Schunk and Winter (2009)

[18]The gamble for resurrection strategy generally refers to the strategies pursued by firms that are close to bankruptcy, and that continue their activity in the aim to remain alive, based on the hope that a fortunate event will occur or that trends will reverse, and make them avoid liquidation.

find that the more loss averse a subject is, the earlier she stops the search process, while Schunk (2009) find that the higher the loss aversion, the higher the reservation price. Thus, loss aversion effect seems to be sensitive to the studied framework. We think this should be investigated more thoroughly.

References

Bearden, J.N., Rapoport, A., Murphy, R.O., 2006. Sequential observation and selection with rank-dependent payoffs: An experimental study. Management Science 52, 1437–1449.

Beckage, N., Todd, P.M., Penke, L., Asendorpf, J., 2009. Testing sequential patterns in human mate choice using speed dating, in: Proceedings of the 2009 cognitive science conference, Cognitive Science Society. pp. 2365–2370.

Berry, D.A., Chen, R.W., Zame, A., Heath, D.C., Shepp, L.A., et al., 1997. Bandit problems with infinitely many arms. The Annals of Statistics 25, 2103–2116.

Bubeck, S., Cesa-Bianchi, N., et al., 2012. Regret analysis of stochastic and nonstochastic multi-armed bandit problems. Foundations and Trends® in Machine Learning 5, 1–122.

Campbell, J., Lee, M.D., 2006. The effect of feedback and financial reward on human performance solving'secretary'problems, in: Proceedings of the Annual Meeting of the Cognitive Science Society.

Caplin, A., Dean, M., Martin, D., 2011. Search and satisficing. American Economic Review 101, 2899–2922.

Carter, L., Russell, P.N., Helton, W.S., 2013. Target predictability, sustained attention, and response inhibition. Brain and Cognition 82, 35–42.

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.

Costa, V.D., Averbeck, B.B., 2013. Frontal–parietal and limbic-striatal activity underlies information sampling in the best choice problem. Cerebral cortex 25, 972–982.

Crosetto, P., Filippin, A., 2013. The "bomb" risk elicitation task. Journal of Risk and Uncertainty 47, 31–65.

Crosetto, P., Filippin, A., 2016. A theoretical and experimental appraisal of four risk elicitation methods. Experimental Economics 19, 613–641.

Czarnitzki, D., Toole, A.A., 2011. Patent protection, market uncertainty, and r&d investment. The Review of Economics and Statistics 93, 147–159.

Descamps, A., Massoni, S., Page, L., 2016. Knowing when to stop and make a choice, an experiment on optimal sequential sampling. Available at SSRN 2883847 .

Dixit, A., 1992. Investment and hysteresis. Journal of economic perspectives 6, 107–132.

Dixit, A.K., Dixit, R.K., Pindyck, R.S., 1994. Investment under uncertainty. Princeton university press.

Dockree, P.M., Kelly, S.P., Roche, R.A., Hogan, M.J., Reilly, R.B., Robertson, I.H., 2004. Behavioural and physiological impairments of sustained attention after traumatic brain injury. Cognitive Brain Research 20, 403–414.

Eckel, C.C., Grossman, P.J., 2002. Sex differences and statistical stereotyping in attitudes toward financial risk. Evolution and human behavior 23, 281–295.

Eriksson, K., Strimling, P., 2009. Biases for acquiring information individually rather than socially. Journal of Evolutionary Psychology 7, 309–329.

Eriksson, K., Strimling, P., 2010. The devil is in the details: Incorrect intuitions in optimal search. Journal of Economic Behavior & Organization 75, 338–347.

Ferguson, T.S., et al., 1989. Who solved the secretary problem? Statistical science 4, 282–289.

Gilbert, J.P., Mosteller, F., 1966. Recognizing the maximum of a sequence. Journal of the American Statistical Association 61, 35–73.

Goel, R.K., Ram, R., 2001. Irreversibility of r&d investment and the adverse effect of uncertainty: Evidence from the oecd countries. Economics Letters 71, 287–291.

Greiner, B., 2015. Subject pool recruitment procedures: organizing experiments with orsee. Journal of the Economic Science Association 1, 114–125.

Helton, W.S., 2009. Impulsive responding and the sustained attention to response task. Journal of Clinical and Experimental Neuropsychology 31, 39–47.

von Helversen, B., Mata, R., 2012. Losing a dime with a satisfied mind: Positive affect predicts less search in sequential decision making. Psychology and aging 27, 825.

Hey, J.D., 1987. Still searching. Journal of Economic Behavior & Organization 8, 137–144.

Holt, C.A., Laury, S.K., 2002. Risk aversion and incentive effects. American economic review 92, 1644–1655.

Holzmeister, F., Pfurtscheller, A., 2016. otree: The "bomb" risk elicitation task. Journal of Behavioral and Experimental Finance 10, 105–108.

Juni, M.Z., Gureckis, T.M., Maloney, L.T., 2016. Information sampling behavior with explicit sampling costs. Decision 3, 147.

Lee, M.D., 2006. A hierarchical bayesian model of human decision-making on an optimal stopping problem. Cognitive Science 30, 1–26.

March, J.G., 1991. Exploration and exploitation in organizational learning. Organization science 2, 71–87.

Navarro, D.J., Newell, B.R., Schulze, C., 2016. Learning and choosing in an uncertain world: An investigation of the explore–exploit dilemma in static and dynamic environments. Cognitive psychology 85, 43–77.

Nelson, P., 1970. Information and consumer behavior. Journal of political economy 78, 311–329.

O'Connell, R.G., Dockree, P.M., Bellgrove, M.A., Turin, A., Ward, S., Foxe, J.J., Robertson, I.H., 2009. Two types of action error: electrophysiological evidence for separable inhibitory and sustained attention neural mechanisms producing error on go/no-go tasks. Journal of Cognitive Neuroscience 21, 93–104.

Oprea, R., Friedman, D., Anderson, S.T., 2009. Learning to wait: A laboratory investigation. The Review of Economic Studies 76, 1103–1124.

Pindyck, R.S., et al., 1991. Irreversibility, uncertainty, and investment. Journal of Economic Literature 29, 1110–1148.

Rapoport, A., Tversky, A., 1966. Cost and accessibility of offers as determinants of optional stopping. Psychonomic Science 4, 145–146.

Rapoport, A., Tversky, A., 1970. Choice behavior in an optional stopping task. Organizational Behavior and Human Performance 5, 105–120.

Robbins, H., 1952. Some aspects of the sequential design of experiments. Bulletin of the American Mathematical Society 58, 527–535.

Robertson, I.H., Manly, T., Andrade, J., Baddeley, B.T., Yiend, J., 1997. Oops!': performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. Neuropsychologia 35, 747–758.

Roebuck, H., Guo, K., Bourke, P., 2015. Attending at a low intensity increases impulsivity in an auditory sustained attention to response task. Perception 44, 1371–1382.

Sandri, S., Schade, C., Musshoff, O., Odening, M., 2010. Holding on for too long? an experimental study on inertia in entrepreneurs' and non-entrepreneurs' disinvestment choices. Journal of economic behavior & organization 76, 30–44.

Sang, K., 2017. Modeling exploration/exploitation behavior and the effect of individual differences. Ph.D. thesis. Indiana University.

Sang, K., Todd, P.M., Goldstone, R., Hills, T.T., 2018. Explore/exploit tradeoff strategies in a resource accumulation search task doi:https://doi.org/10.31234/osf.io/zw3s8.

Schunk, D., 2009. Behavioral heterogeneity in dynamic search situations: Theory and experimental evidence. Journal of Economic Dynamics and Control 33, 1719–1738.

Schunk, D., Winter, J., 2009. The relationship between risk attitudes and heuristics in search tasks: A laboratory experiment. Journal of Economic Behavior & Organization 71, 347–360.

Seale, D.A., Rapoport, A., 1997. Sequential decision making with relative ranks: An experimental investigation of the" secretary problem". Organizational behavior and human decision processes 69, 221–236.

Seale, D.A., Rapoport, A., 2000. Optimal stopping behavior with relative ranks: The secretary problem with unknown population size. Journal of Behavioral Decision Making 13, 391–411.

Simon, H.A., 1955. A behavioral model of rational choice. The quarterly journal of economics 69, 99–118.

Simon, H.A., 1956. Rational choice and the structure of the environment. Psychological review 63, 129.

Simon, H.A., 1982. Models of bounded rationality: Behavioral economics and business organizations. Cambridge. MA: MIT Press. 32, 337–351.

Sonnemans, J., 1998. Strategies of search. Journal of economic behavior & organization 35, 309–332.

Todd, P.M., Miller, G.F., 1999. From Pride to Prejudice to Persuasion: Satisficing in Mate Search. Oxford University Press, New York. Evolution and Cognition, pp. 287–308.

Viefers, P., Strack, P., 2014. Too proud to stop: Regret in dynamic decisions. Working paper .

Zwick, R., Rapoport, A., Lo, A.K.C., Muthukrishnan, A., 2003. Consumer sequential search: Not enough or too much? Marketing Science 22, 503–519.

Supplementary Materials

Experimental interface

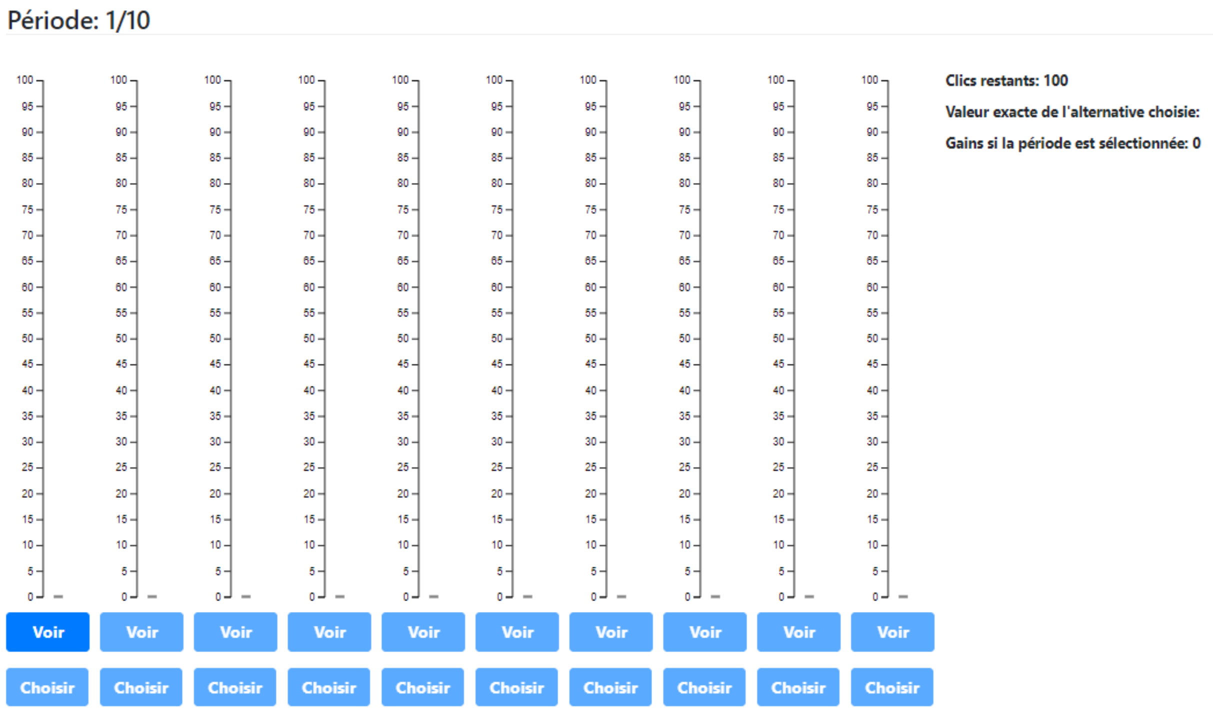


Figure S1: Main task interface at the beginning of the round (for both treatments)

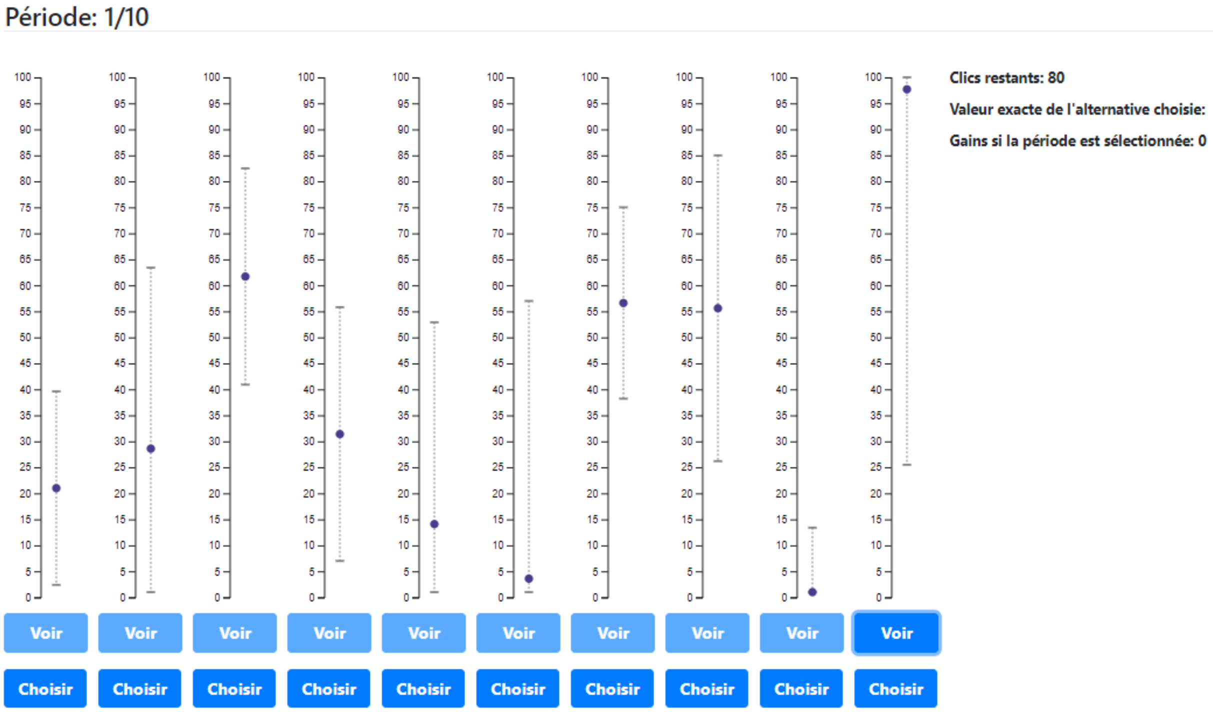


Figure S2: Main task interface after few clicks on each alternative (in Treatment U)

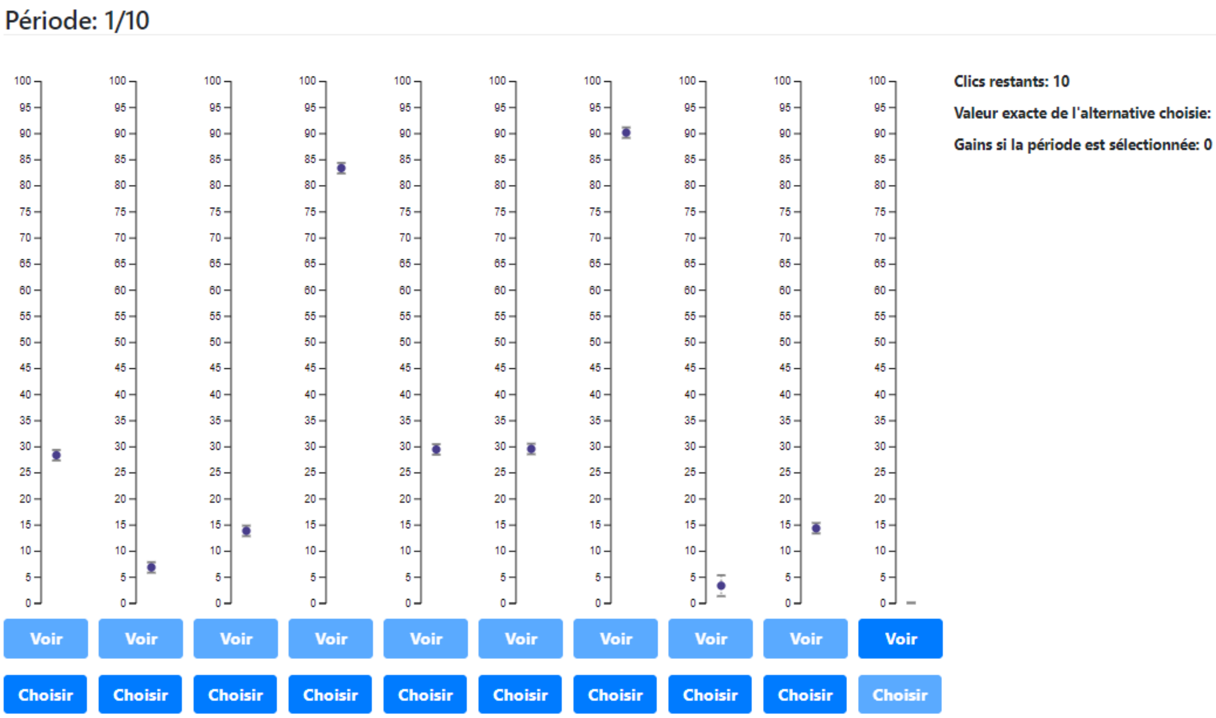
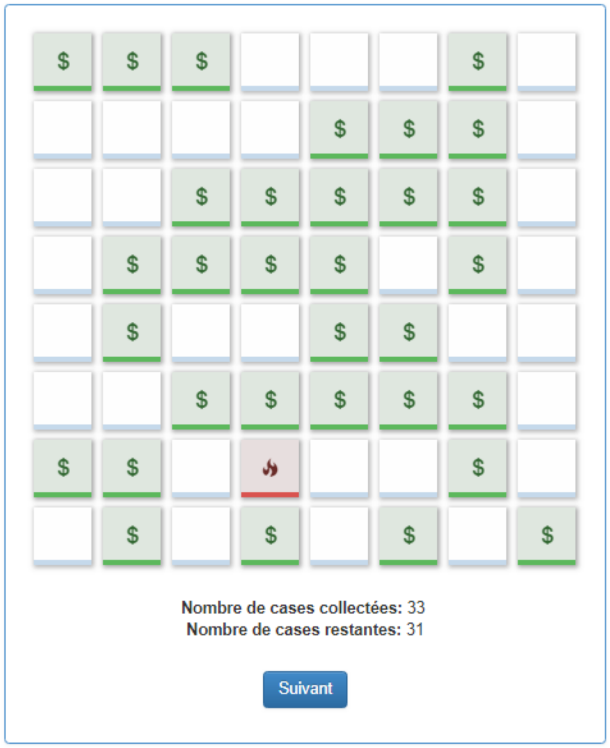
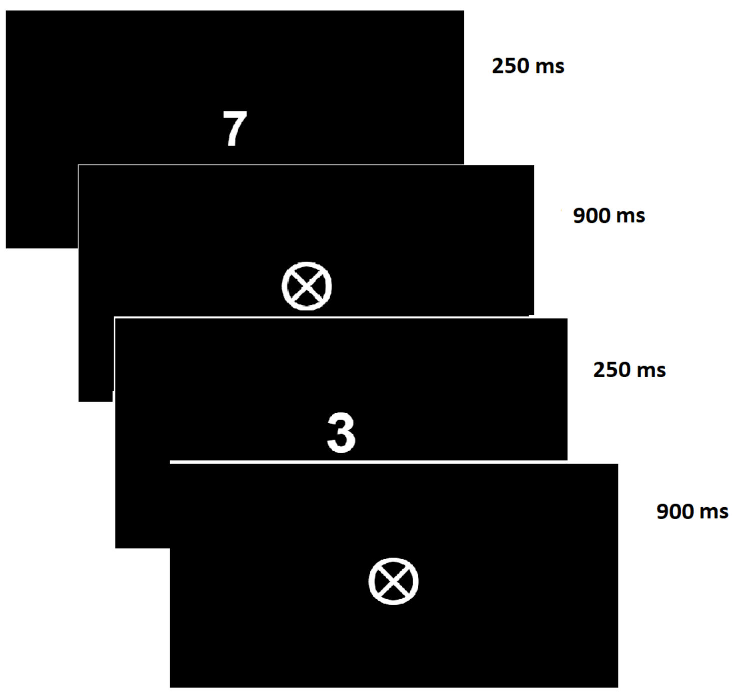


Figure S3: Main task interface after 10 clicks on each of the 9 first alternatives (in Treatment C or U)



(a) BRET



(b) SART

Figure S4: Control tasks

Instructions

Original french version

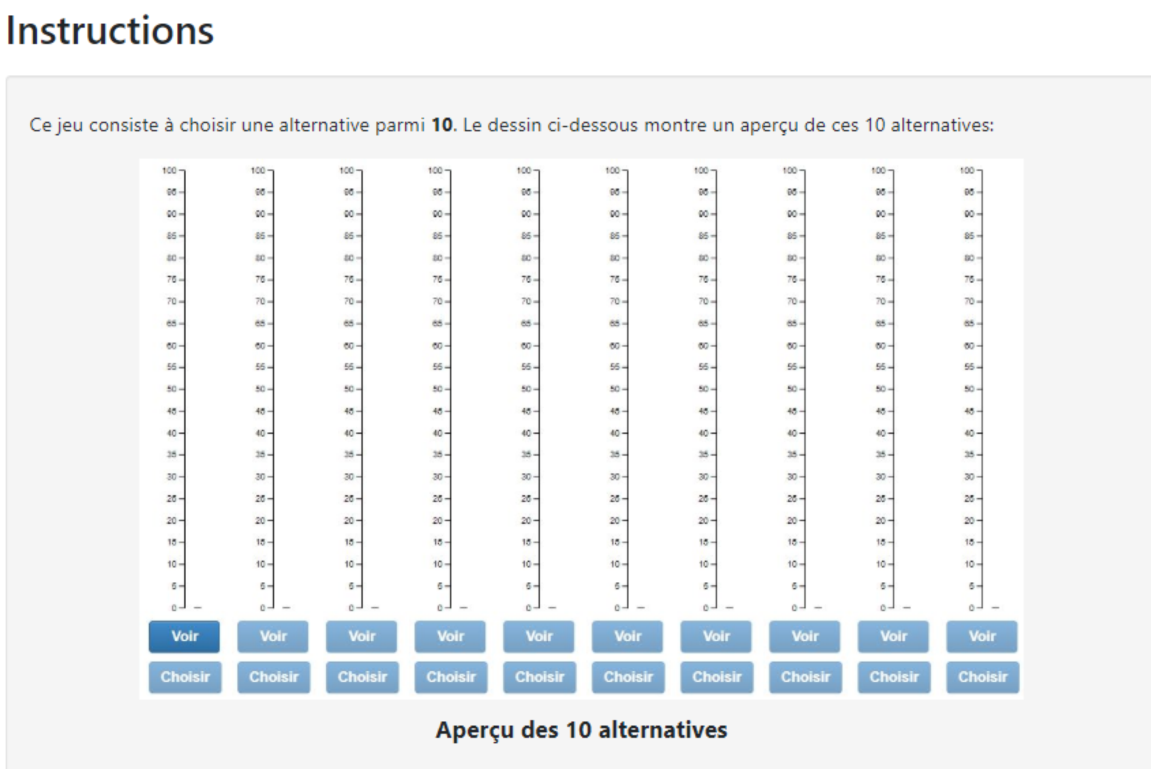


Figure S5: Instructions, before part 1, both treatment orders, original french version

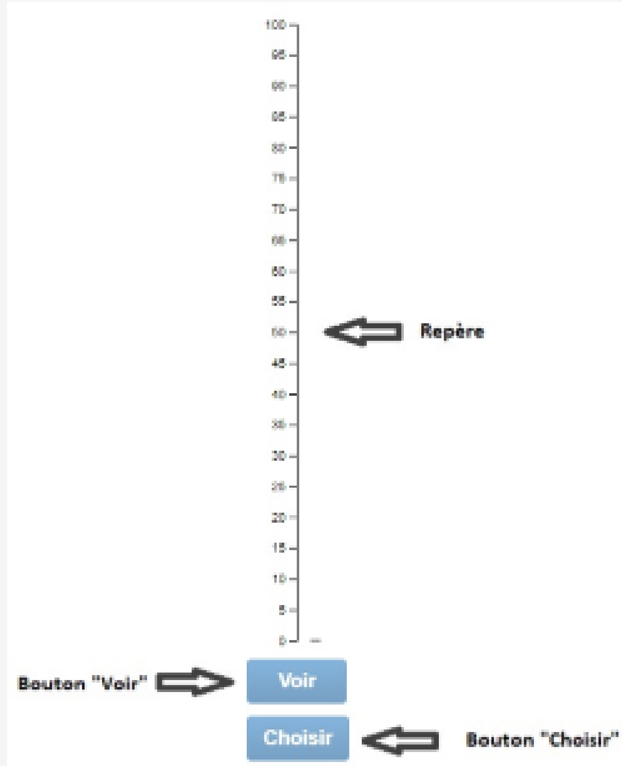
Le dessin ci-contre représente **une** alternative. Chacune des alternatives vaut entre 1 et 100. Les valeurs de ces différentes alternatives sont incertaines, c'est-à-dire que vous ne pouvez connaître qu'un intervalle dans lequel la valeur exacte d'une alternative se situe probablement. Plus cet intervalle est petit, moins vous aurez d'incertitude quant à la valeur exacte d'une alternative.

Vous allez jouer et répéter ce jeu pendant 10 périodes successives. Une seule de ces périodes, tirée au sort, sera comptabilisée pour le calcul de votre gain final.

Lors de chaque période :

Vous disposez d'un stock total de 100 clics que vous pourrez répartir entre deux utilisations :

- Découvrir la valeur d'une ou de plusieurs alternatives
- Récolter les points correspondant à l'alternative choisie



Aperçu d'une alternative

Découvrir une alternative:

La valeur d'une alternative est un nombre entier compris entre 1 et 100, chacun de ces nombres pouvant être obtenu avec la même chance.

Vous pouvez découvrir la valeur d'une alternative en cliquant sur le bouton "**Voir**" qui se trouve en dessous de l'alternative considérée. En cliquant sur ce bouton, vous obtenez la moyenne d'un certain nombre de tirages d'une distribution normale autour de la valeur exacte de l'alternative, ainsi qu'un intervalle de précision dans lequel se situe cette valeur exacte avec un degré de certitude d'environ 95%.

Cette découverte se fait progressivement. Plus vous cliquez sur le bouton "**Voir**", plus le nombre de tirages augmente et plus vous réduisez la taille de l'intervalle de précision. Au bout de 10 clics sur le bouton "**Voir**", la taille de cet intervalle de précision devient très petite. Vous êtes alors quasiment certain de la valeur exacte de l'alternative considérée.

Au fur et à mesure de vos clics, la moyenne et l'évolution du degré de précision sont illustrées graphiquement respectivement par un point et un intervalle. En positionnant la souris exactement sur le point vous pourrez lire les valeurs (arrondies) de cette moyenne ainsi que celle des bornes de l'intervalle sous la forme :

« *moyenne* (sup : *borne supérieure de l'intervalle*, inf : *borne inférieure de l'intervalle*) »

Remarque: Si la borne inférieure de l'intervalle est inférieure à 1, elle est remplacée par 1. Si la borne supérieure de l'intervalle est supérieure à 100, elle est remplacée par 100.

Dans ce jeu, si vous souhaitez découvrir une alternative, vous devez la découvrir complètement, c'est-à-dire, utiliser la totalité des 10 clics.

C'est pour cette raison que nous appellerons ce jeu le jeu de 10-clics.

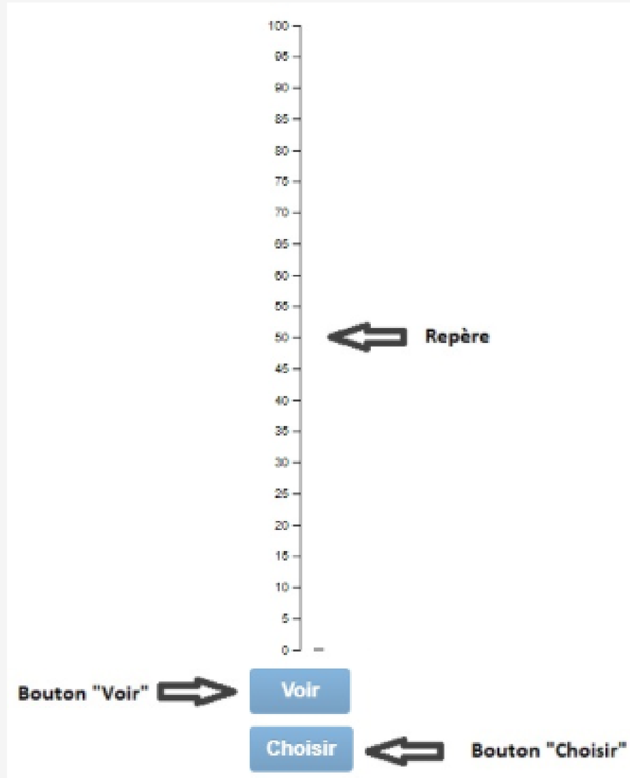
Figure S6: Instructions, before part 1 of the first played treatment, treatments order CU, original french version, cont.

Le dessin ci-contre représente **une** alternative. Chacune des alternatives vaut entre 1 et 100. Les valeurs de ces différentes alternatives sont incertaines, c'est-à-dire que vous ne pouvez connaître qu'un intervalle dans lequel la valeur exacte d'une alternative se situe probablement. Plus cet intervalle est petit, moins vous aurez d'incertitude quant à la valeur exacte d'une alternative.

Vous allez jouer et répéter ce jeu pendant 10 périodes successives. Une seule de ces périodes, tirée au sort, sera comptabilisée pour le calcul de votre gain final.

<u>Lors de chaque période :</u>
Vous disposez d'un stock total de 100 clics que vous pourrez répartir entre deux utilisations :

- Découvrir la valeur d'une ou de plusieurs alternatives
- Récolter les points correspondant à l'alternative choisie



Aperçu d'une alternative

Découvrir une alternative:

La valeur d'une alternative est un nombre entier compris entre 1 et 100, chacun de ces nombres pouvant être obtenu avec la même chance.

Vous pouvez découvrir la valeur d'une alternative en cliquant sur le bouton "**Voir**" qui se trouve en dessous de l'alternative considérée. En cliquant sur ce bouton, vous obtenez la moyenne d'un certain nombre de tirages d'une distribution normale autour de la valeur exacte de l'alternative, ainsi qu'un intervalle de précision dans lequel se situe cette valeur exacte avec un degré de certitude d'environ 95%.

Cette découverte se fait progressivement. Plus vous cliquez sur le bouton "**Voir**", plus le nombre de tirages augmente et plus vous réduisez la taille de l'intervalle de précision. Au bout de 10 clics sur le bouton "**Voir**", la taille de cet intervalle de précision devient très petite. Vous êtes alors quasiment certain de la valeur exacte de l'alternative considérée.

Au fur et à mesure de vos clics, la moyenne et l'évolution du degré de précision sont illustrées graphiquement respectivement par un point et un intervalle. En positionnant la souris exactement sur le point vous pourrez lire les valeurs (arrondies) de cette moyenne ainsi que celle des bornes de l'intervalle sous la forme :

« *moyenne* (sup : *borne supérieure de l'intervalle*, inf : *borne inférieure de l'intervalle*) »

Remarque: Si la borne inférieure de l'intervalle est inférieure à 1, elle est remplacée par 1. Si la borne supérieure de l'intervalle est supérieure à 100, elle est remplacée par 100.

Dans ce jeu, vous pouvez cliquer autant de fois que vous le souhaitez (entre 1 et 10) sur le bouton "Voir" d'une même alternative.
C'est pour cette raison que nous appellerons ce jeu le jeu de x-clics.

Attention: Si vous décidez de découvrir une alternative, vous ne pourrez plus revenir en arrière pour continuer à découvrir les alternatives précédentes.

Figure S7: Instructions, before part 1 of the first played treatment, treatments order UC, original french version, cont.

Choisir une alternative:

Le choix d'une alternative s'effectue en cliquant sur le bouton "**Choisir**".
Lorsque vous choisissez une alternative, vous récoltez un montant de points qui correspond à la valeur exacte de cette alternative multipliée par le nombre de clics restant à votre disposition, divisé ensuite par 10:

Gains de la période = clics restants * valeur exacte de l'aternative choisie / 10

Cette information figurera sur la droite de votre écran comme vous pouvez le voir sur la figure ci-contre:
Vous ne pouvez choisir qu'une seule alternative.
Si vous utilisez tout votre capital de clics à la découverte des valeurs des alternatives, vous ne pourrez plus choisir une alternative particulière et ne récolterez par conséquent aucun point.

Clics restants: 100

Valeur exacte de l'alternative choisie:

Gains si la période est sélectionnée: 0

Gains de la période

Remarque: Lors du premier clic, vous pouvez seulement uiliser le bouton "**Voir**" de la première alternative.

Commencer

Figure S8: Instructions, before part 1 of the first played treatment, both treatment orders, original french version, cont.

Instructions

Vous allez jouer et répéter ce jeu pendant 10 périodes successives. Une seule de ces périodes, tirée au sort, sera comptabilisée pour le calcul de votre gain final.

Ce jeu est identique au précédent, à une différence près

La différence est la suivante:

Vous pouvez cliquer autant de fois que vous le souhaitez (entre 1 et 10) sur le bouton "Voir" d'une même alternative.
De ce fait, nous appellerons ce jeu le jeu de **x-clics**.

Attention: Si vous décidez de découvrir une alternative, vous ne pourrez plus revenir en arrière pour continuer à découvrir les alternatives précédentes.

Commencer

Figure S9: Instructions, before part 1 of the second played treatment, treatments order CU, original french version

Instructions

Vous allez jouer et répéter ce jeu pendant 10 périodes successives. Une seule de ces périodes, tirée au sort, sera comptabilisée pour le calcul de votre gain final.

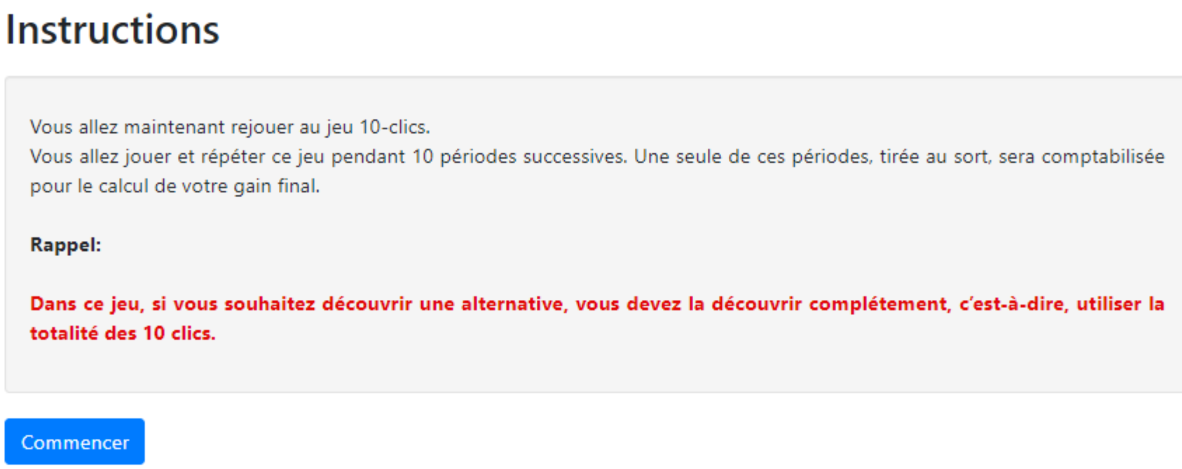
Ce jeu est identique au précédent, à une différence près

La différence est la suivante:

Dans ce jeu, si vous souhaitez découvrir une alternative, vous devez la découvrir complétement, c'est-à-dire, utiliser la totalité des 10 clics.
De ce fait, nous appellerons ce jeu le jeu de **10-clics**.

Commencer

Figure S10: Instructions, before part 1 of the second played treatment, treatments order UC, original french version



Instructions

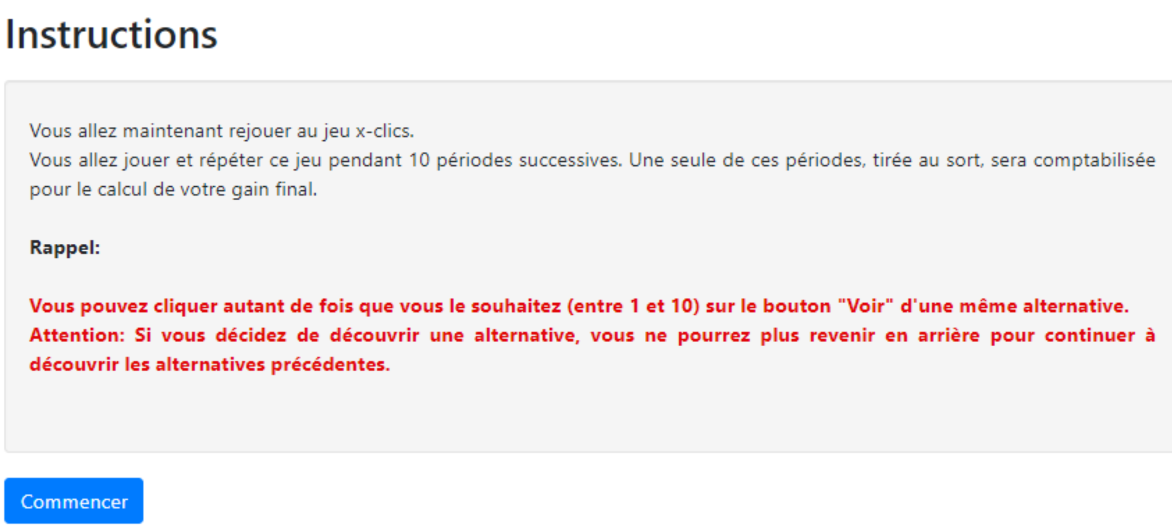
Vous allez maintenant rejouer au jeu 10-clics.
Vous allez jouer et répéter ce jeu pendant 10 périodes successives. Une seule de ces périodes, tirée au sort, sera comptabilisée pour le calcul de votre gain final.

Rappel:

Dans ce jeu, si vous souhaitez découvrir une alternative, vous devez la découvrir complétement, c'est-à-dire, utiliser la totalité des 10 clics.

Commencer

Figure S11: Instructions, before part 2 of treatment C, both treatment orders, original french version



Instructions

Vous allez maintenant rejouer au jeu x-clics.
Vous allez jouer et répéter ce jeu pendant 10 périodes successives. Une seule de ces périodes, tirée au sort, sera comptabilisée pour le calcul de votre gain final.

Rappel:

Vous pouvez cliquer autant de fois que vous le souhaitez (entre 1 et 10) sur le bouton "Voir" d'une même alternative.
Attention: Si vous décidez de découvrir une alternative, vous ne pourrez plus revenir en arrière pour continuer à découvrir les alternatives précédentes.

Commencer

Figure S12: Instructions, before part 2 of treatment U, both treatment orders, original french version

Instructions

Vous allez maintenant rejouer au jeux **10-clics** et **x-clics** pendant un total de 20 périodes successives dans un **ordre aléatoire**. Deux de ces périodes, tirées au sort, seront comptabilisées pour le calcul de votre gain final.

Rappel:

- Dans le jeu **10-clics**:
  Si vous souhaitez découvrir une alternative, vous devez la découvrir complétement, c'est-à-dire, utiliser la totalité des 10 clics.

- Dans le jeu **x-clics**:
  Vous pouvez cliquer autant de fois que vous le souhaitez (entre 1 et 10) sur le bouton "Voir" d'une même alternative.
  Attention: Si vous décidez de découvrir une alternative, vous ne pourrez plus revenir en arrière pour continuer à découvrir les alternatives précédentes.

Commencer

Figure S13: Instructions, before part 3, both treatment orders, original french version

Translated instructions:

Before part 1 of the first played treatment:

This game consists of choosing an alternative among **10**. The figure below shows a preview of these 10 alternatives.

The figure below cons represents **one** alternative. Each of the alternatives is worth between 1 and 100. The values of these different alternatives are uncertain, which means that you can only know an interval inside which the exact value of an alternative is probably situated. The smaller is this interval, the less you have uncertainty about the exact value of an alternative.

You will play and repeat this game for 10 successive rounds. One of these rounds will be randomly selected and taken into account in the calculation of your final payment.

During a round:
You have of a total stock of 100 clicks that you can allocate between two uses:

- Discovering the value of one or several alternatives
- Collecting the points corresponding to the chosen alternative

Discovering an alternative:
The value of an alternative is a number comprised between 1 and 100, each of these numbers could be obtained with an equal chance.

You can discover the value of an alternative by clicking on the button **"See"** which is situated below the concerned alternative. By clicking on this button, you obtain the mean of a given number of draws from a normal distribution around the exact value of the alternative, as well as an precision interval inside which this exact value is situated with a certainty level of around 95%.

This discovery is done progressively. The more times you click on the button **"See"**, the more the number of draws increases and the more you reduce the size of the precision interval. After 10 clicks on the button **"See"**, the size of this precision interval becomes very small. You are then almost certain of the exact value of the concerned alternative.

As you click, the mean and the evolution of the precision interval are represented graphically respectively by a point and an interval. By positioning the mouse cursor exactly on the point, you can read the (rounded) values of this means and of the bounds of the interval under the following form:
«*mean* (sup: *upper bound of the interval*, inf: *lower bound of the interval*»

Remarks: If the lower bound of the interval is lower than 1, it is replaced by 1. If the upper bound of the interval is greater than 100, it is replaced by 100.

Only in treatments order CU:
In this game, if you wish to discover an alternative, you will have to discover it completely, i.e., use the totality of the 10 clicks.

This is why we will call this game, the 10-clicks game.

Only in treatments order UC:

In this game, you can click as many times as you want (between 1 and 10) on the "See" button of a same alternative.

This is why we will call this game, the x-clicks game.

Be careful: If you decide to discover and alternative, you will not be able to go back to continue discovering the previous ones.

Choosing an alternative:

To choose an alternative, you need to click on the button **"Choose"**.

When you choose an alternative, you collect a number of points corresponding to the exact value of this alternative multiplied by the number of remaining clicks, then divided by 10:

$$\textbf{Profit of the round} = \\ \textbf{remaining clicks} \times \textbf{exact value of the chosen alternative} / 10$$

This information will be displayed on the right of your screen as you can see in the below cons figure.

You can choose only one single alternative.

If you use the totality of your clicks capital for the discovery of the values of the alternatives, you will not be able to choose an alternative and will consequently not collect any point.

Remark: For the first click, you can only use the **"See"** button of the first alternative.

Translated instructions (cont.):

Before part 1 of the second played treatment:

You will play and repeat this game for 10 successive rounds. Only one of these rounds will be randomly selected and taken into account in the calculation of your final payment.

This game is identical to the previous one, at one difference.

The difference is the following:

Only in treatments order CU:

You can click as many times as you want (between 1 and 10) on the "See" button of a same alternative.

This is why we will call this game, the x-clicks game.

Be careful: If you decide to discover and alternative, you will not be able to go back to continue discovering the previous ones.

Only in treatments order UC:

In this game, if you wish to discover an alternative, you will have to discover it completely, i.e., use the totality of the 10 clicks.

This is why we will call this game, the 10-clicks game.

Before part 2

Only in treatment C:

You will now play again to the game of 10-clicks. You will play and repeat this game for 10 successive rounds. Only one of these rounds, randomly selected, will be taken into account in the calculation of you final payment.

Reminder:

In this game, if you wish to discover an alternative, you will have to discover it completely, i.e., use the totality of the 10 clicks.

Only in treatment U:

You will now play again to the game of x-clicks. You will play and repeat this game for 10 successive rounds. Only one of these rounds, randomly selected, will be taken into account in the calculation of you final payment.

Reminder:

You can click as many times as you want (between 1 and 10) on the "See" button of a same alternative.

Be careful: If you decide to discover and alternative, you will not be able to go back to continue discovering the previous ones.

Before part 3

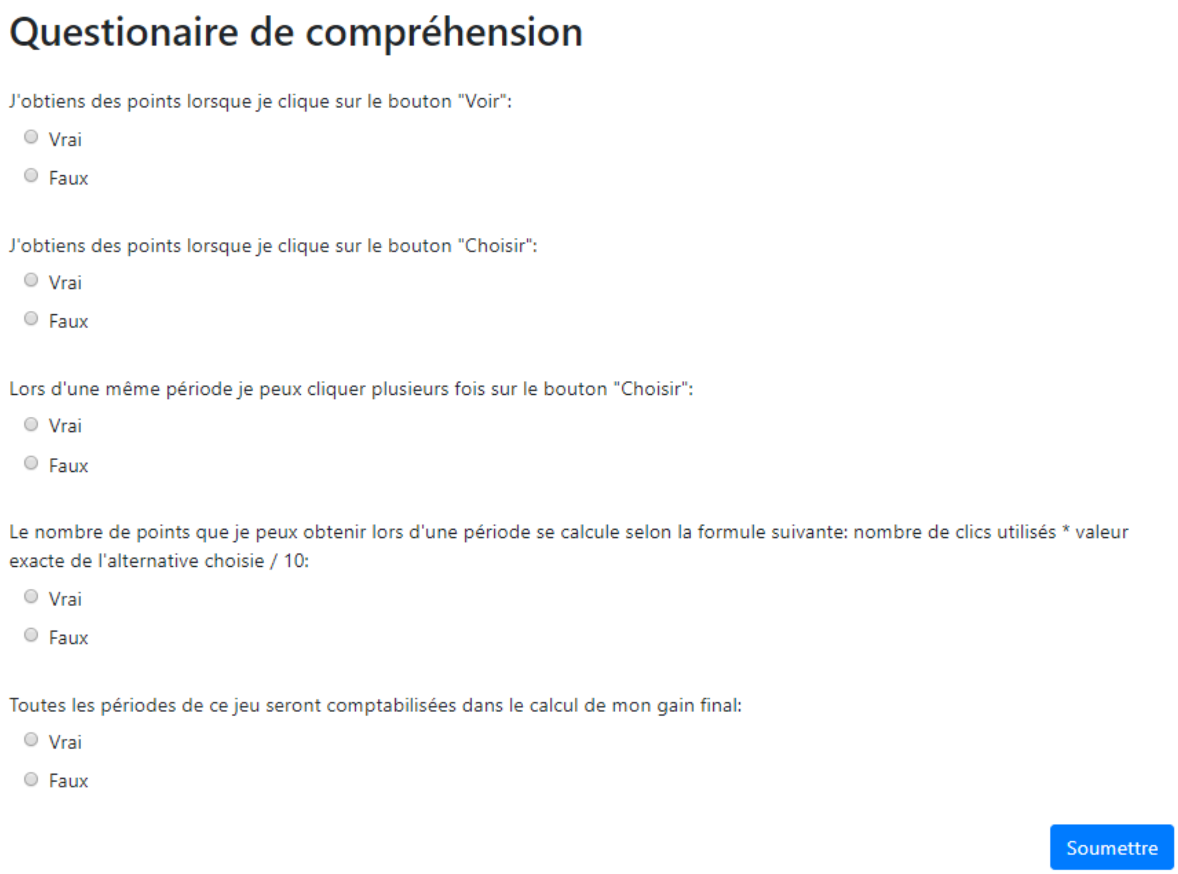
You will now play again to the games **10-clicks** and **x-clicks** for a total of 20 successive rounds in a random order. Two of these rounds, randomly selected, will be taken into account in the calculation of your final payment.

Reminder:

- In the **10-clicks** game:
  If you wish to discover an alternative, you will have to discover it completely, i.e., use the totality of the 10 clicks.

- In the **x-clicks** game:
  You can click as many times as you want (between 1 and 10) on the "See" button of a same alternative.
  Be careful: If you decide to discover and alternative, you will not be able to go back to continue discovering the previous ones.

Comprehension questionnaire:



Questionaire de compréhension

J'obtiens des points lorsque je clique sur le bouton "Voir":

- Vrai
- Faux

J'obtiens des points lorsque je clique sur le bouton "Choisir":

- Vrai
- Faux

Lors d'une même période je peux cliquer plusieurs fois sur le bouton "Choisir":

- Vrai
- Faux

Le nombre de points que je peux obtenir lors d'une période se calcule selon la formule suivante: nombre de clics utilisés * valeur exacte de l'alternative choisie / 10:

- Vrai
- Faux

Toutes les périodes de ce jeu seront comptabilisées dans le calcul de mon gain final:

- Vrai
- Faux

Soumettre

Figure S14: Comprehension questionnaire, original french version

Translated comprehension questionnaire:

I Collect points when I click on the button “See”:

- True
- False

I Collect points when I click on the button “Choose”:

- True
- False

During a same round, I can click several time on the button “Choose”:

- True
- False

The number of points that I can collect during one round is calculated according to the following formula:

$$\text{Number of used clicks} \times \text{exact value of the chosen alternative} / 10$$

- True
- False

All the rounds will be taken into account for the calculation of my final payment:

- True
- False

Simulation-based results

To compute the simulation-based results, we consider under both treatments all the possible combinations of trials allocations. In Treatment C, there are in total 9 possible combinations: allocating 10 trials to the first alternative then exploiting it, allocating 10 trials to the first alternative and 10 trials to the second alternative then exploiting the highest between these two, etc. We do not consider the combination where the player allocates 10 trials to explore all of the 10 alternatives, since the corresponding payoff to that combination is necessarily *zero*, because no trials are left for exploitation.

In Treatment U, we moreover need to consider the different combinations of the different possible number of trials spent at each alternative. In our simulations, we consider up to 4 trials by explored alternative. By doing so, we have a total of 1.398.100 possible combinations. In both treatments, for each combination we compute the corresponding payoff and the number of explored alternatives. The payoff is calculated based on the real value of the chosen alternative, even if the exploitation decision is made based on the observed value at the current trial. We then search for the optimal combination, i.e., the combination yielding the highest payoff.

In order to verify the hypothesis in our theoretical model under uncertainty, claiming to spend the minimum time (i.e. one trial or substep) on each explored alternative, we looked at the number of trials spent on each explored alternative in the simulations-based results under Treatment U. These results show that spending more than 1 trial on one alternative occurred less than 6% of the time (see Table S1). This is a support for our hypothesis.

	1 trial	**2 trials**	**3 trials**	**4 trials**
Occurrences	19911	974	234	54
	(94,04%)	(4,60%)	(1,11%)	(0,26%)

Table S1: Trials by alternative under Treatment U (uncertainty)

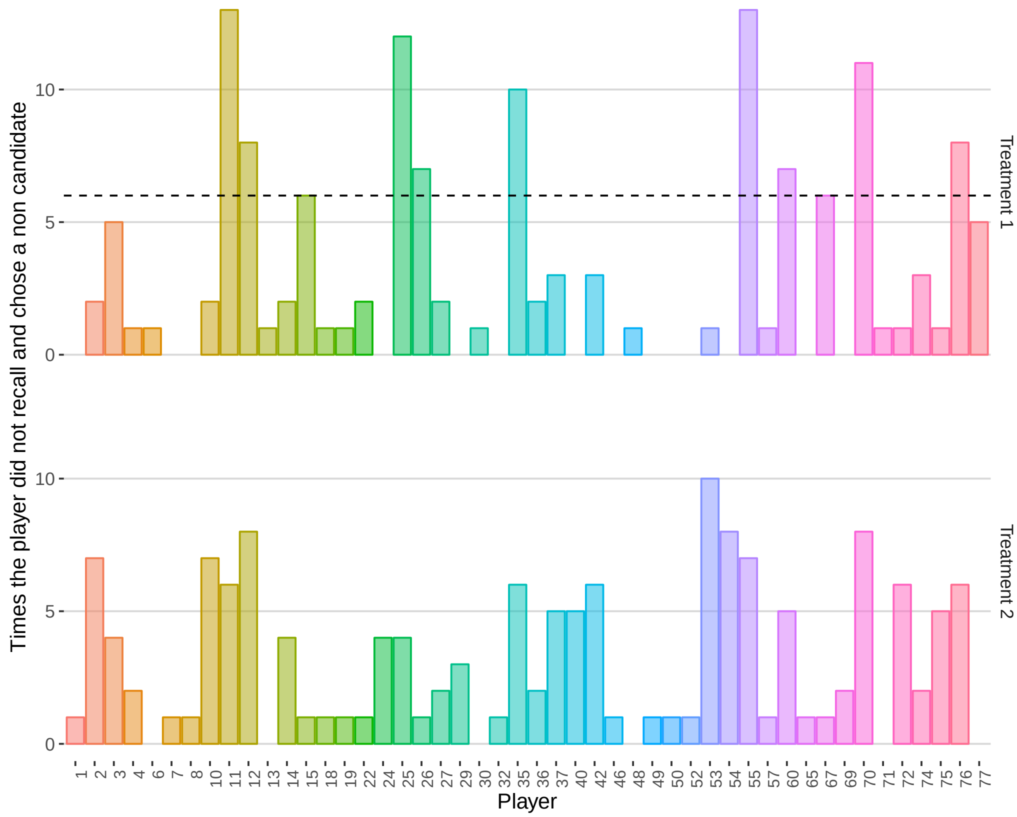
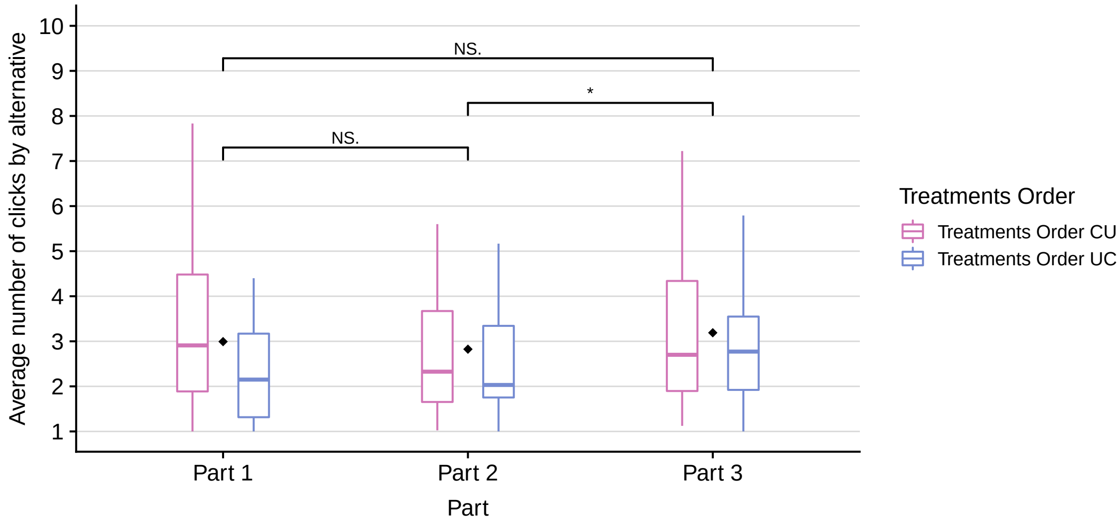


Figure S15: Individual frequencies of not recalling and choosing a non-candidate



(a)

		Part 1	Part 2	Part 3
Treatments order CU	*Mean*	3.476	3.041	3.326
	Median	2.909	2.327	2.701
Treatments order UC	*Mean*	2.480	2.595	3.045
	Median	2.150	2.031	2.772
Both	*Mean*	2.993	2.825	3.190
	Median	2.445	2.175	2.708

(b)

Figure S16: Number of clicks by alternative's evolution through the different parts of Treatment U by treatments order
(a). Means by part (black diamonds) and Boxplots by treatments order and part. Comparisons: Two-tailed paired t-test
Note: $^{NS.}p \geq 0.1$; $^{\circ}p<0.1$; $^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$
(b). Means and medians

DOCUMENTS DE TRAVAIL GREDEG PARUS EN 2019
GREDEG Working Papers Released in 2019

2019-01 Muriel Dal Pont Legrand & Harald Hagemann
Impulses and Propagation Mechanisms in Equilibrium Business Cycles Theories: From Interwar Debates to DSGE "Consensus"

2019-02 Claire Baldin & Ludovic Ragni
Note sur quelques limites de la méthodologie de Pareto et ses interprétations

2019-03 Claire Baldin & Ludovic Ragni
La conception de l'homme dans la théorie de l'Echange Composite de François Perroux : entre homo economicus et homo religiosus

2019-04 Charlie Joyez
Shared Ownership in the International Make or Buy Dilemma

2019-05 Charlie Joyez
Alignment of Multinational Firms along Global Value Chains: A Network-based Perspective

2019-06 Richard Arena & Ludovic Ragni
Nature humaine et choix rationnel : Pareto contre Walras ?

2019-07 Alain Raybaut
A First French Episode in the Renewal of Nonlinear Theory of Economic Cycles (1978-1985)

2019-08 Alain Raybaut
Bertrand Nogaro et l'économie de guerre : le Service de la main d'œuvre étrangère

2019-09 Nicolas Brisset & Dorian Jullien
Models as Speech Acts: A Restatement and a new Case Study

2019-10 Kozo Kiyota, Toshiyuki Matsuura & Lionel Nesta
On Export Premia

2019-11 Nicolas Brisset & Raphaël Fèvre
Peregrinations of an Economist: Perroux's Grand Tour of Fascist Europe

2019-12 Marco Baudino
Urbanization and Development: A Spatial Framework of Rural-to-urban Migration

2019-13 Giuseppe Attanasi, Kene Boun My, Nikolaos Georgantzís & Miguel Ginés
Strategic Ethics: Altruism without the Other-regarding Confound

2019-14 Thierry Kirat & Frédéric Marty
How Law and Economics Was Marketed in a Hostile World: L'institutionnalisation du champ aux États-Unis de l'immédiat après-guerre aux années Reagan

2019-15 Agnès Festré, Ankinée Kirakozian & Mira Toumi
La menace est plus forte que l'exécution, mais pas pour tous : sanction versus recommandation par un tiers dans une expérience de bien public

2019-16 Nicolas Brisset, Raphaël Fèvre & Tom Juille
Les années noires de la "Science de l'Homme": François Perroux, la Fondation Carrel et l'appropriation de la sociologie

2019-17 Romain Plassard
From Disequilibrium to Equilibrium Macroeconomics: Barro and Grossman's Trade-off between Rigor and Realism

2019-18 Christophe Charlier, Gilles Guerassimoff, Ankinée Kirakozian & Sandrine Selosse
Under Pressure! Nudging Electricity Consumption within Firms: Feedback from a Field Experiment

2019-19 Nicolas Brisset & Raphaël Fèvre
Prendre la parole sous l'État français: le cas de François Perroux

2019-20 Giuseppe Attanasi, Ylenia Curci, Patrick Llerena & Giulia Urso
Intrinsic vs. Extrinsic Motivators on Creative Collaboration: The Effect of Sharing Rewards

2019-21 Giuseppe Attanasi, Ylenia Curci, Patrick Llerena, Adriana Carolina Pinate & Giulia Urso
Looking at Creativity from East to West: Risk Taking and Intrinsic Motivation in Socially and Culturally Diverse Countries

2019-22 Nobuyuki Hanaki
Cognitive Ability and Observed Behavior in Laboratory Experiments: Implications for Macroeconomic Theory

2019-23 Dongshuang Hou, Aymeric Lardon, Panfei Sun & Genjiu Xu
Sharing a Polluted River under Waste Flow Control

2019-24 Sothearath Seang & Dominique Torre
Proof of Work and Proof of Stake Consensus Protocols: A Blockchain Application for Local Complementary Currencies

2019-25 Paige Clayton, Maryann Feldman & Benjamin Montmartin
Funding Emerging Ecosystems

2019-26 Frédéric Marty
Quels impacts des règles relatives à la protection des données sur la concurrence entre plateformes infonuagiques, industrielles et immobilières?

2019-27 Giuseppe Attanasi, Roberta Dessi, Frédéric Moisan & Donald Robertson
Public Goods and Future Audiences: Acting as Role Models?

2019-28 Ludovic Ragni
Note sur la conception de la science chez Cournot et Walras : critique et filiation au regard de la philosophie d'Etienne Vacherot

2019-29 Thomas Le Texier & Ludovic Ragni
Concurrence 'hybride', innovation et régulation : un modèle de duopole

2019-30 Mattia Guerini, Duc Thi Luu & Mauro Napoletano
Synchronization Patterns in the European Union

2019-31 Richard Arena
La théorie économique est-elle encore utile ?

2019-32 Richard Arena
Is still to-day the Study of the "Surplus Product" the True Object of Economics?

2019-33 Richard Arena & Katia Caldari
Léon Walras and Alfred Marshall: Microeconomic Rational Choice or Human and Social Nature?

2019-34 Maya El Hourani & Gérard Mondello
The Impact of Bank Capital and Institutional Quality on Lending: Empirical Evidence from the MENA Region

2019-35 Jean-Luc Gaffard
Le devenir du libéralisme

2019-36 Jacques Mairesse, Michele Pezzoni & Fabiana Visentin
Does Gender Matter for Promotion in Academia? Evidence from Physicists in France

2019-37 Charles Ayoubi, Michele Pezzoni & Fabiana Visentin
Does it Pay to Do Novel Science? The Selectivity Patterns in Science Funding

2019-38 Patrice Bougette, Axel Gautier & Frédéric Marty
Which Access to Which Assets for an Effective Liberalization of the Railway Sector?

2019-39 Frédéric Marty
Conventions de l'Etat et politique de concurrence européenne

2019-40 Imen Bouhlel, Michela Chessa, Agnès Festré & Eric Guerci
When to Stop? A Theoretical and Experimental Investigation of an Individual Search Task