Centre de Recherche en Économie et Management
Center for Research in Economics and Management

CREM UMR CNRS 6211

University of Rennes 1

University of Caen Normandie

Working Paper

Choice or information overload ?

Fabrice LE LEC

CES, Université de Paris 1, France

Marianne LUMEAU

CEPN, Université de Paris 13, LABEX ICCA, France

Benoît TARROUX

CREM, UMR CNRS 6211, Faculté des Sciences Économiques
Université de Rennes 1, France

April 2016 - WP 2016-07

UNIVERSITÉ CAEN NORMANDIE

UNIVERSITÉ DE RENNES 1

Choice or information overload?

Fabrice Le Lec, Marianne Lumeau and Benoît Tarroux *

April 2016

Abstract

This paper aims to test how the profusion of choice and information affects individuals' decisions. In particular, we investigate whether the possible choice overload effects are due to the mere presence of many alternatives or the difficulty in processing abundance of information that comes with the proliferation of options. To do so, we use the frequency with which familiar alternatives are preferred to unfamiliar ones as a behavioural measure of overload. We first propose an individual decision model, in which uncertainty about values of alternatives leads consumer to prefer familiar goods. We use this theoretical approach to devise an experiment where the level of information and the number of alternatives systematically vary. Our results show that individuals are prone to overload in the presence of larger choice sets, but that information has a small impact, if any.

Keywords Choice overload; Information overload; Bounded rationality; Familiarity; Experimental Economics

1 Introduction

Opportunities and information proliferate in modern economies. More than ever, consumers or investors face a huge number of options and have access to countless information. According to the standard microeconomics approach, the more choice and information the agents have, the better their decision: On the one hand, agents provided with more options are more likely to find one that corresponds to their preferences; On the other hand, better informed agents are able to reduce the uncertainty about the options characteristics and better identify the best alternative. However, a growing body of research in psychology, marketing and more recently in economics has suggested that the simultaneous abundance of choice and information may have negative side-effects, including decision avoidance, procrastination or lower satisfaction after a decision (Tversky

*Le Lec: University Paris 1 and Centre d'Economie de la Sorbonne CNRS; Lumeau: University Paris 13, CEPN and LABEX ICCA; Tarroux: University Rennes 1 and CREM. The authors thank Elven Priour for programming the experiment. Financial support from the MSHB and the Labex ICCA is gratefully acknowledge. The authors also thank Dorothea Kübler and the participants of the CREM seminar in Rennes 1, of the CEPN seminar in Paris 13, of ASFEE conference in Besançon (France), of Marsouin conference in Auray, of ACEI conference in Montreal, of ESA European Meeting in Prague and of OSE-PSE meeting for helpful comments.

and Shafir, 1992; Iyengar and Lepper, 2000; Scheibehenne et al., 2010a). These side-effects are often generically referred to as the *choice paradox* or *choice overload* effect.[1] Various causes of this phenomenon have been proposed, linking it to the cognitive limitations and efforts needed to process information and to make a decision or to the psychological costs of making a decision such as (anticipated) regret of bad choice, anxiety of making a choice or loss aversion (Roese, 1997; Loewenstein, 1999; Kamenica, 2008; Sarver, 2008; Ortoleva, 2013).

In line with these papers, our aim is to uncover choice paradox sources. While most situations considered in the literature are simultaneously characterized by both information and choice abundance, we attempt to isolate their respective effects. In particular, we contend that choice paradox may be explained by *information overload* or/and *pure choice overload*. The former refers to the fact that, because of limited cognitive ability, agents are not able to process appropriately the high level of information that often comes with a large number of available options. Independently from the information process, consumers can also suffer from pure choice overload,[2] which consists in having difficulty to make a decision under a profusion of options. This can be explained by the trouble individuals experience in comparing options due to unclear preferences.

To investigate this issue, we first build a model of choice and information overload, and then test its prediction in a simple consumption experiment. In both, a consumer has to choose an item in situations where the number of alternatives and information pieces about alternatives varies. The choice set contains two types of goods: *familiar* goods that she already knows and has experienced, and for which there is no uncertainty about satisfaction, and *unfamiliar* ones for which satisfaction is typically uncertain. The probability of choosing a familiar option rather than an unfamiliar one provides a simple measure of choice or information overload.

The intuition of the model is the following: A large number of opportunities or information pieces results in consumer confusion when comparing options and evaluating the best alternatives. More specifically, the beliefs on the subjective values of unfamiliar options are more uncertain, i.e., subject to a spreader noise, in the case of information or choice proliferation, than in their absence. Under the rather undisputed assumption that most individuals are risk averse, choice or information overload should imply that subjects tend to choose familiar goods more often when the choice set or the amount of information is large.

To eventually establish the relative extent of information and choice overload, we then set up an experiment where subjects could choose familiar or unfamiliar options and where the amount of information and the number of options were varied independently. The experiment proceeds

[1] Literature also makes reference to *choice fatigue*. This refers more specifically to the negative effects of the repetition of decisions on the quality of future ones.

[2] For the sake of conciseness, we refer to information and choice overload to mean information and *pure* choice overload, knowing that both could be sources of the more general phenomenon referred to in the general literature as choice overload.

as follows. Subjects are first revealed a choice set comprising both familiar and unfamiliar goods and have then to collect information about them. After the information collection stage, subjects have to rank the goods according to their preferences following an incentive compatible mechanism based on a 'real consumption' period. To study the relative importance of information and choice overload independently, the size of the choice set (*Small* or *Large*) and the information level are varied across treatments (*No Info*, *Low Info* or *High Info*).

In comparison with other studies on the choice paradox, our method focuses exclusively on the effect of choice/information abundance on the nature of the chosen alternative. In most previous studies, the effect put forth is rather choice avoidance or procrastination (Tversky and Shafir, 1992; Iyengar and Lepper, 2000; Iyengar et al., 2004; Shin and Ariely, 2004), post-decision regret or satisfaction (Reutskaya and Hogarth, 2009; Scheibehenne et al., 2010b). Yet, the evidence is relatively scarce regarding the effect of choice paradox on the final decision and the nature of the option chosen. An exception is Iyengar and Kamenica (2010) who provide evidence that the likelihood of choosing a simple-to-understand option increases with the number of available alternatives.

Our results indicate that when the amount of information increases, familiar goods are not chosen more often. This suggests that subjects are not prone to information overload (or very weakly so) and are able to process available information to reduce the uncertainty regarding the options. On the contrary, we observe a clear effect of the size of the choice set: Independently of the amount of information, subjects tend to choose familiar products more frequently with large choice sets than with small ones. This effect appears large and robust enough to suggest the presence of a strong *pure choice overload* effect. These results support the view that choice paradox is a non-negligible phenomenon and could stem from the trouble decision maker experience in subjectively assessing their own preferences.

The existence of a pure choice overload effect has several important consequences. On the theoretical side, most models in decision theory and microeconomics assume that decision processes are independent of irrelevant alternatives. Although this assumption has already been challenged empirically (Tversky and Simonson, 1993; Ariely and Wallsten, 1995), our results put forth that decision process may depend on the size of the choice set. This may impair the generalization of results obtained with small sets (typical in experiments) to real-life situations where the number of opportunities is very large.

Our study also echoes the debate on the importance of choice architecture for public policy (Sunstein and Thaler, 2003; Leonard, 2008). As empirical research tends to indicate (see Cronqvist and Thaler, 2004; Bhargava and Loewenstein, 2015; Bhargava et al., 2015), the complexity of a decision makes options difficult to compare for individuals often results in non-optimal choice or the reliance on *default* option or *status quo*. Our results also suggests that a critical aspect of a successful choice architecture may be the size of the choice set. In addition to the simplification of the information available, the information presentation format, or the existence of a default option, maintaining a small choice set may be critical to get a well-weighed, if not optimal, decision from individuals. Since our results tend to show that most influential feature of a choice situation is the size of the choice set, it might be superior to offer a small number of options than change other features.

In the realm of industrial organization, this choice overload effect may partly explain a tendency of markets to be highly concentrated on some 'superstar products' (Rosen, 1981). In almost all countries, market regulation tends to favour more competition between firms to ensure lower price but also more varieties and choice on markets. With more varieties, pure choice overload may simply lead consumers to limit their own product line to a small number of popular or habitual products, a phenomenon that may in the end and quite counter-intuitively favour big players.

The remainder of the paper is organized as follows. Section 2 develops a model of choice and information overload. Section 3 describes the experiment. Section 4 presents our main experimental findings while Section 5 discusses these findings and concludes.

2 Theoretical framework

In our model, a risk-averse consumer has to choose between familiar and unfamiliar goods. The subjective value of unfamiliar items is typically uncertain, while familiar items have a certain value. Assuming that these values are random variables, the context of choice, i.e., information or choice set size, may vary this uncertainty and may affect the probability that a consumer will choose a familiar/unfamiliar good. However, how the values change with context depends on the type of the consumer (rational or overloaded).

2.1 Setting

Let X be the set of all possible alternatives. We denote $C \subset X$ the choice set and x a generic good. We assume that the consumer is faced some level of information about the goods. For the sake of simplicity, we consider this amount as a real number denoted I.[3]

The subjective value of x is typically uncertain and represented by a random variable ω_x. This uncertainty stems from two sources. First, the characteristics of x are uncertain and the consumer can update her belief about it in light of new information. This *external* uncertainty is related to the level of information I the consumer can process. The second type of uncertainty the agent faces refers to the difficulty she may have to know her own preferences and to identify precisely

[3]We discuss the possible practical implementations of this rather vague notion of the level of information in the next section.

how x fares with respect to her preferences. This *internal* uncertainty is assumed to be related to the number of alternatives $\#C$.

Assuming the separability of these two sources of randomness, the (uncertain) subjective value of x when provided with information I and faced with the choice set C is given by:

$$\omega_x(C, I) = m_x + \tau_x^I + \mu_x^C \tag{1}$$

with $m_x = \mathbb{E}(\omega_x)$ and τ_x^I and μ_x^C two random components such as:

$$\mathbb{E}(\tau_x^I) = \mathbb{E}(\mu_x^C) = 0$$

Information or choice overload can be described as the effect on τ_x^I and μ_x^C of the amount of information and the size of the choice set respectively. If option x is perfectly known by the consumer, then $\tau_x^I = 0$. In this case, she could be still uncertain of how much she likes the item, due to 'random preferences' or uncertainty about her own preferences. This effect is captured by μ_x^C. It is also possible that $\mu_x^C = 0$, but that the consumer could be uncertain about the caracteristics of the good, as captured by τ_x^I. Note that although we assume separability, the two are not necessarily independent: it seems plausible that more external uncertainty will imply more internal uncertainty. This is formally assumed in the following:

Assumption 1 *The correlation between τ_x^I and μ_x^C is non-negative and non-decreasing in their variances $\sigma^2(\tau_x^I)$ and $\sigma^2(\mu_x^C)$.*

Let $\Delta(\mathbb{R})$ be the set of real random variables, then every option is an element of $\Delta(\mathbb{R})$ for the decision maker. To specify the behaviour of the agents, we assume that the consumer is risk averse in the sense that her subjective value is transformed into utility by a concave function, denoted u. As is conventional, the consumer maximizes her expected utility. The preferences of a consumer on the set of options are denoted with $\succ$, conditionally on the choice context, denoted $\mid C, I$. This is expressed in the following assumption:

Assumption 2 (Representation of preferences and risk aversion) *Exists $u : \Delta(\mathbb{R}) \to \mathbb{R}$ with $u' > 0$, and $u'' < 0$ such that for any two options x, y:*

$$x \succ y \mid C, I \iff \mathbb{E}\big[u(\omega_x(C, I))\big] > \mathbb{E}\big[u(\omega_y(C, I))\big] \tag{2}$$

The expected utility of an item x can be approximated in a way similar to Pratt (1964)'s "risk in the small" (see the Appendix for the derivation):[4]

$$\mathbb{E}\big[u(\omega_x(C, I))\big] \simeq u(m_x) + \frac{\sigma^2(\tau_x^I)}{2}u''(m_x) + \frac{\sigma^2(\mu_x^C)}{2}u''(m_x) + \sigma\big(\tau_x^I, \mu_x^C\big)u''(m_x) \tag{3}$$

[4]This supposes that either the random perturbation is small or that the effect of higher-order derivatives of u'' is negligible in our context.

with $\sigma^2(.)$ denoting the variance of a random variable.

2.2 Typology of consumers

We then define different types of consumers based on how $\sigma^2(\tau_x^I)$ and $\sigma^2(\mu_x^C)$ are affected respectively by I and C. A rational (i.e., standard) consumer is defined as follows:

Definition 1 (Rational consumer) *A* rational consumer *is such that* $\sigma^2(\mu_x^C) = 0$ *and* $\sigma^2(\tau_x^I)$ *decreases in* I*.*

The first part of this definition is simply that rational consumers have preferences transparent to themselvesand the size of the choice set does not affect the precision of their preferences. The second part of the definition simply says that the more information pieces an individual has, the lower variance about the subjective value of the item. Said differently, only the external uncertainty has an impact on the rational consumer.

The second type of consumer suffers from difficulty in processing information when it is provided in abundance:

Definition 2 (Information overloaded consumer) *An* information overloaded consumer *is such that* $\sigma^2(\tau_x^I)$ *increases in* I *(possibly above some threshold).*

The more information an individual gathers about options, the more confused and the less precise her representation of a given option is. Having more information is hence detrimental to the "quality" (i.e., specificity) of her beliefs. Such a phenomenon may stem from limited cognitive capacities (Simon, 1955) or limited working memory (Baddeley, 2003). Too much information may not be handled properly by the consumer, who may not be capable of properly ascribing pieces of information to alternatives and have in the end a worse representation of options than she would have had with less information.

The equivalent for choice overload is given by:

Definition 3 (Choice overloaded consumer) *A* choice overloaded consumer *is such that* $\sigma^2(\mu_x^C)$ *is increasing in* $\#C$ *(possibly above some threshold).*

Here, more choice results in difficulties (or the mere inability) to determine the best option (or just a satisfactory one), which is represented in our framework as a more widespread μ_x^C. That is, the more options the consumer is contemplating the less clear her preference is. While not modelling directly some accounts of choice overload found in the related literature, this definition relates quite closely to most of them. Indeed, the larger variance of subjective values may translate indecision or inability to choose (Danan et al., 2012; Ortoleva, 2013).[5]

[5] Regret or anticipated regret could also fit in this definition (see Loomes and Sugden (1982) or Sarver (2008) for instance): more uncertainty about the genuine subjective value of an option increases the probability of making

2.3 Behavioural predictions

It could exist some goods already known or experienced by the consumer, namely familiar goods. The external and internal uncertainty of this type of goods is low or null. If a consumer faces a choice set consisting of familiar and unfamiliar goods, we can use the choice of a familiar good as a measure of choice or information overload. Formally:

Assumption 3 (Familiar goods) *Let f be a familiar good, then ω_f is degenerated and constant with I and C, i.e., $\omega_f(C, I) = m_f$.*

Any choice set C can hence be decomposed as $C = N \cup F$, with N the set of unfamiliar goods and F the set of familiar goods. The crux of our method is that familiar goods are not affected (or negligibly so) by the context of choice (information or choice set) because the consumer has a pre-existing precise representation of this good. Using the lack of variation in the choice of familiar goods relative to unfamiliar goods provides a measure of the effect of the context of choice. Denote $\mathbb{P}(x \succ y \mid C, I)$ the probability that x is preferred to y when the context of choice is given by C for the opportunity set and I for the information available. Our aim is to elicit how the probability that an individual prefers a familiar good f to an unfamiliar one n, i.e., $\mathbb{P}[f \succ n \mid C, I]$, is related to choice or information overload. For the sake of simplicity, we assume there is no indifference between items. Moreover, no specific assumption about the distribution across individuals of m_f and m_n is needed.

For f a familiar option, and n an unfamiliar one, the probability that a rational consumer prefers f to n is constant with respect to C. Yet, the probability that a rational consumer prefers f to n decreases with respect to I: she can treat and process information so she is better informed (the variance of the lottery decreases). The following proposition summarizes this:

Proposition 1 (Rational consumer) *For a rational consumer, a familiar good f and an unfamiliar good n, it holds that: on the one hand, for $C' \subset C \in X$, $\mathbb{P}[f \succ n \mid C', I] = \mathbb{P}[f \succ n \mid C, I]$; and on the other hand, for $I' < I$, $\mathbb{P}[f \succ n \mid C, I'] > \mathbb{P}[f \succ n \mid C, I]$.*

For a same level of information, the uncertainty about the unfamiliar options is held constant, independently of the size of choice set. On the contrary, more information allows the consumer to decrease the uncertainty about $\omega_n(C, I)$ for all $n \in N$. If she is not too risk averse, she is less likely to prefer the familiar item than the unfamiliar one.

A consumer who is prone to information overload is more likely to prefer an familiar item than an unfamiliar one when the available information increases. The additional information confuses her more and renders the prospect of the unfamiliar goods more uncertain. She is not able to use the

a bad one, and consequently of experiencing regret. We use the concavity of the utility function as a measure of risk-aversion. One of the drivers for risk aversion could be regret.

additional information to decrease uncertainty about the consumption of unfamiliar items unlike the rational decision maker.

Proposition 2 (Information overloaded consumer) *For an information overloaded consumer, a familiar item f and an unfamiliar one n, it holds that: For $I' < I$, $\mathbb{P}[f \succ n \mid C, I'] < \mathbb{P}[f \succ n \mid C, I]$*

A choice overloaded consumer is expected to prefer more frequently a familiar item to an unfamiliar one when the choice set is larger because of increasing confusion about unfamiliar options without impacting that of familiar items.

Proposition 3 (Choice overloaded consumer) *For a choice-overloaded consumer, a familiar good f and an unfamiliar good n: For $C' \subset C$, $\mathbb{P}[f \succ n \mid C', I] < \mathbb{P}[f \succ n \mid C, I]$*

The beliefs of the decision maker are more vague, less precise in the presence of many options, except for familiar goods. This phenomenon increases the variance of the subjective values of the unfamiliar goods and pushes the consumer towards the less uncertain goods, i.e., the familiar ones.

To summarize, overloaded consumers can differ from the standard benchmark in two dimensions: Information overloaded consumers tend to relatively prefer familiar goods when the level of information increases, in contrast with standard consumer who tends to relatively favour unfamiliar goods, because of the reduction of the uncertainty on the quality of the unfamiliar goods. For the choice overloaded consumer, the choice set has a negative impact on the attractiveness of unfamiliar goods, while for the standard consumer, the number of options is not relevant. It is worth noting that both information and choice overload can interact (through the correlation of the corresponding noises) and that the effect of the presence of both can exceed the sum of the two effects.

3 Experimental design and procedures

3.1 Experimental design

The experiment consists of six treatments where subjects have to choose and consume an experience good in the lab, after collecting pieces of information about the alternatives. The experiment is divided into five stages, summarized in Table 1.[6] The goods proposed to subjects were content websites.[7] We chose this type of goods since the consumption of such goods can be easily implemented and controlled in the lab. Moreover, it is reasonable to assume that all subjects or almost all are familiar with at least some websites.

[6] The instructions (translated from the French) can be found in the Appendix.

[7] These websites are offering their own content, which excludes e-commerce websites and social networks.

(0)		(1)		(2)		(3)		(4)
Select a type of content	⇒	Choice set (websites)	⇒	Information (or not)	⇒	Rank items	⇒	Consumption (30 minutes)

Table 1: Course of the experiment

To study the effect of information and the choice set, we introduce two experimental variations with respect to the number of information pieces, which are to be collected by the subjects, and the size of the choice set subjects face. Table 2 summarizes the parameters of our different treatments.

During the *preliminary stage* (0), subjects have to indicate which type of content they are interested in.[8] In addition to making sure that subjects were familiar with some of the websites in the choice set, this stage also guarantees that subjects' motivation is high enough when considering all the options.

In *stage 1*, subjects are presented the choice set they will face for the whole experiment. Options are labeled through their web addresses (URL), (*e.g.*, www.nytimes.com) to provide minimal information. The number of alternatives differs across treatments: 4 options in the *Small choice set* (C^-) and 8 options in the *Large choice set* (C^+). To study the choice of familiar versus unfamiliar items, all choice sets are designed to obtain, insofar as is possible, similar proportions of familiar and unfamiliar websites at the individual level: each choice set was composed with fixed shares of "blockbuster", *i.e.* very famous websites, mildly famous ones and niche ones. We categorize websites according to the number of 'likes' collected on the local version of Facebook. The websites with a high number of 'likes' (i.e., blockbuster websites) are more likely to be familiar to the subjects, while those which collected the lower level of 'likes' (such as blogs) are very likely to be unknown.[9] In addition, the subjects' familiarity with websites was also measured *ex post* at the individual level with a questionnaire. Subjects also have to report whether they already had heard of each website before the experiment in order to control for some external information.

In *stage 2*, subjects collect pieces of information. For each website, three types of information are available: a neutral description, a popularity index (*i.e.*, the number of 'likes' on facebook.fr) and a comment from a user.[10] The total number of available information pieces depends on the size of the choice set (*i.e.*, 12 in C^- and 24 in C^+), but subjects' access to information is constrained with respect to the treatment. Three treatments are implemented. In the treatments with *Low*

[8] The available types were general news, sports, culture and pop culture, economics, games, movies and TV series, cooking.

[9] The average number of 'likes' of the most popular websites in a choice set is around 346,000 (C^+) and 313,000 (C^-), with a certain heterogeneity from 117,000 (economics) to 700,000 (news). The least popular website has around 3,000 (C^+) and 5,000 (C^-) 'likes'.

[10] The comments were drafted in the following way: 'I recommend this website/I do not recommend this website' plus an explanation. To control for the nature of the recommendations, 50% of the websites of a category (blockbusters, mildly famous ones, and niches) were provided with a positive comment and 50% with a negative comment. Moreover, the positive and negative nature of the recommendations was reversed across sessions to obtain a balanced design.

	Small Choice Set (C^-)	Large Choice Set (C^+)
No Information (I^0) (0 % of info.)	C^-/I^0 4 options 0 info. piece	C^+/I^0 8 options 0 info. piece
Low Information (I^-) (25 % of info.)	C^-/I^- 4 options 3 info. pieces	C^+/I^- 8 options 6 info. pieces
High Information (I^+) (50 % of info.)	C^-/I^+ 4 options 6 info. pieces	C^+/I^+ 8 options 12 info. pieces

Table 2: Treatment matrix

Information level (I^-), subjects have to select and consult 25% of the available information pieces, whereas in the treatments with *High Information* level (I^+), they have to consult 50% of the available information. In both, subjects are forced to open the required number of information pieces, to guarantee their exposition to a given level of information, but are free to choose which pieces of information they want to consult. In the treatments with *No Information* (I^0), the second stage is not implemented and the only information subjects have is the web address of the different items. Strictly speaking, I^0 is rather a minimal information treatment than a no information treatment: not only the website URLs can be informative on the content of a website, but subjects come to the lab with an initial information level about some of the websites (without necessarily having experienced them). As previously mentioned, subjects were asked whether they had previously heard of each website, in order to provide a tighter control of their initial level of information. From a methodological point of view, I^0 mostly serves the purpose of a control treatment. It allows to test the model developed in section 2 according to which subjects should mostly prefer familiar options. This provides a check that subjects are not motivated by other aspects of the peculiar situation, for instance, some curiosity for new websites linked to the specificity of the lab context.

The experimental variation in information is based on a *constant relative number of information pieces*: Subjects can consult the same number of information pieces *per website*. When the size of a choice set is constant, it seems indisputable that the level of information in I^+ is greater than for I^-. Yet, when the choice set varies, it is less clear what a greater level of information means. Consider for instance the treatment C^- (4 options) and the treatment C^+ (8 options): giving 6 pieces of information in C^-/I^+ and 12 in C^+/I^+ guarantees that the level of information per options remains constant, but arguably the amount of information to be processed in the second case is larger. It could hence be that any difference between C^-/I^+ and C^+/I^+ is attributable to the increase in the absolute level of information, and some apparent choice overload may in fact stem from information overload. Two features of the design allow to test (and possibly rule out) this possible confusion: first, C^+/I^- can be compared with C^-/I^+ for which the *absolute level* *of information* stays constant (equal to 6), while the choice set size varies; second, if the absolute level of information matters, we should observe differences between C^-/I^- and C^-/I^+ on the one hand and C^+/I^- and C^+/I^+ on the other hand. As a consequence, it seems that maintaining the relative level of information, rather than the absolute one, provides a better identification strategy of choice and information overload.

During *stage 3*, subjects have to rank the different alternatives according to their preferences. To induce them to reveal their true preferences, we design a real-consumption incentive mechanism: Subjects know that they will have to spend half an hour in the lab with the access to one website as the sole source of entertainment. This website is chosen randomly but in accordance to their ranking: The probability of getting their first-ranked website is greater than the probability of having the second-ranked website, and so on. Table 3 presents the probability that a website is drawn by the computer program depending on the subject's ranking. Such a procedure ensures that revealing one's true ranking is a stochastically dominant strategy, and hence the procedure is incentive-compatible.

Alternative measures could have involved having the subjects choose their preferred option only, or to ask for binary comparisons. On top of gathering less data points, the former raises a subtle statistical issue: maintaining the proportion of familiar and unfamiliar constant, the probability that the best option is a familiar one is likely to increase with the size of the choice set.[11] The second measure, that asks subjects a series of binary questions, presents two drawbacks. First, the only difference with the elicitation of the ranking is to allow for inconsistencies. Although they may be of interest in themselves, at a practical level they mostly add noise to the measure of preferences. Second, asking binary questions to subjects may simply destroy the possible effect of the size of the choice set: the situation under consideration by the subjects is not a relatively large choice set but simply two items to compare.

The *last stage* is the consumption stage. Subjects are informed of the website selected and can then consult this website, and only this one, during 30 minutes. The Internet access of the subjects' computer was locked to this particular website and subjects are not authorized to use their phone or any documents. This 'real' consumption stage was set to reinforce the incentive attached to the ranking task and to have them consider carefully the options and information available. This procedure is similar to the one used in recent works in experimental economics (Reutskaja et al., 2011; Le Lec and Tarroux, 2012).

[11] Consider two families of iid random variables (X_k) and (Y_l), all mutually independent. If for their cdf, we have $F_X < F_Y$ as is implied in our framework when considering the expected utility of options, then the probability that the maximum of all the random variables is given by a X rather than a Y increases for many usual distributions (normal, log-normal, Gamma, etc.) with k and l when k/l is held constant. When the Xs and Ys are identically distributed though this probability is constant.

Rank	Small choice set (C^-)	Large choice set (C^+)
1	0.50	0.45
2	0.30	0.25
3	0.15	0.15
4	0.05	0.07
5	-	0.05
6	-	0.02
7	-	0.01
8	-	0

Table 3: Probability of being drawn for options depending on ranks

Treatment	Choice set size (C)	Information set size (I)	Number of sessions	Number of subjects
C^+/I^+	+	+	5	66
C^+/I^-	+	-	4	60
C^+/I^0	+	0	2	47
C^-/I^+	-	+	4	60
C^-/I^-	-	-	4	60
C^-/I^0	-	0	2	48

Table 4: Characteristics of treatments

3.2 Procedures

The experiment was held in November 2013 and February 2014 at the Center for Research in Economics and Management (CREM), University Rennes 1, France. The experiment was computerized using the z-Tree program (Fischbacher, 2007) and consisted in 21 sessions, summarized in Table 4. Given the specific nature of the experiment, we used a between-subject design. In total, 341 subjects (46.6% female) were recruited among a population of undergraduate students from a variety of majors.

The payoff was a fixed show-up fee of 12 Euros. The show-up fee was unusually high to account for the fact that the main experiment does not offer subjects the possibility to earn additional money: the incentive-compatibility is ensured by the fact that subjects' choice will be implemented in the lab, where subjects spend 30 minutes with an option. In addition though, subjects could earn some money in a side experiment, which took place at the end of the consumption stage, by taking a Holt and Laury (2002) task. We add this side experiment in order to measure subjects' attitude towards risk, which theoretically plays a role in the probability of choosing a familiar option. On average, the total payoff was 15 euros. Moreover, a session lasted 90 minutes, including initial instructions, lab consumption, the side experiment and the subject payment.

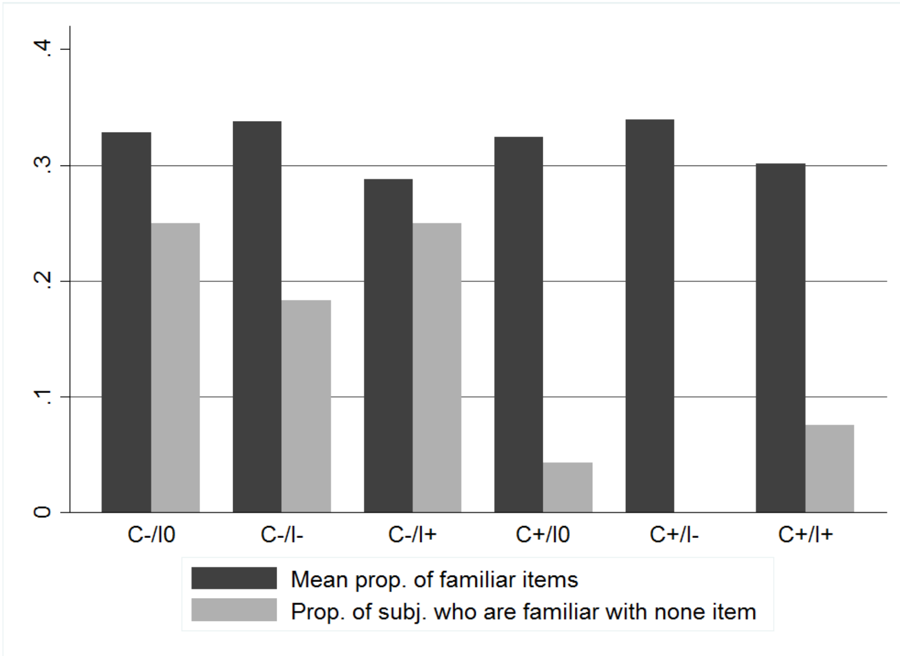


Figure 1: Difference of familiar items between treatments

4 Results

We first present a general description of the results, before studying in depth the preferences of subjects conditional to the context of choice.

4.1 Subjects' familiarity with items

Before studying the respective role of information level and the options number on the choice of familiar item, we control for the subjects' initial familiarity with the websites they face. As previously mentionned, we asked them to report for each website whether they are familiar with this website prior the experiment, that is whether they have already visited the website before they run the experiment.

On average, a subject is familiar with 31.9% of the websites, leaving 68.1% of the websites subject as unfamiliar. Figure 1 presents the mean proportion of familiar items in each treatment and shows that this proportion is rather stable across treatments. Usual pairwise tests do not suggest any significant difference between the treatments. This result indicates that our selection of the choice sets does not seem suffer from any major sampling problem in this respect. Yet, quite mechanically, the proportion of subjects who are familiar with no items is significantly larger in treatments with small choice sets than those with large choice sets.

Choice set	No info (I^0)	Low info (I^-)	High info (I^+)	All
Small choice set (C^-)	87.5	65.6	68.6	72.8
	(131 ; 36)	(169 ; 48)	(153 ; 44)	(453 ; 128)
Large choice set (C^+)	90.6	72.5	80.5	80.4
	(538 ; 45)	(741 ; 60)	(733 ; 61)	(2012 ; 166)
All	89.2	69.5	75.5	80.1
	(669 ; 81)	(910 ; 108)	(886 ; 105)	(2465 ; 294)

The sizes of the samples are given in parenthesis: the first figure correspond to the number of valid comparisons, the second one the number of subjects.

Table 5: Average frequencies with which a familiar item is preferred to an unfamiliar item (in %)

4.2 Descriptive measures of the subjects' preference for familiar items

We now study how the context of choice (choice set size and information) influences the frequency with which a familiar item is preferred to an unfamiliar one, through two statistical tools. First, we rely on a plain descriptive measure of the preference for familiar items. Second, we provide a rank-ordered logit analysis of subjects' ranking of websites.

To determine whether a given option is preferred to another one, we use the rank data. We focus only on the comparisons involving a familiar option and an unfamiliar one. For instance, for an individual reporting familiarity with one option among the four available, we obtain three valid comparisons. Subjects reporting being familiar with all the options or with no option are automatically removed from this analysis. We then compute at the individual level the frequencies with which a familiar option is preferred to an unfamiliar one, which are reported in Table 5.

These descriptive measures of the preference for a familiar option call for a few comments. On average, subjects seems to prefer a familiar item to an unfamiliar one. However, these preferences are affected by the information and choice set variations. In the absence of information, subjects overwhelmingly prefer familiar items to unfamiliar ones, in around 90% of the cases. It tends to support the basic assumptions of our theoretical framework. Table 5 also indicates that providing some information rather than none seems to strongly impact the probability of choosing a familiar option: the difference in the probability of choosing a familiar item is reduced in the treatments I^+ and I^- compared to the treament I^0.[12] In addition, it seems that the probability of choosing a familiar item is minimal for I^-, and may tend to increase from I^- and I^+ (or at least to be constant).

Regarding the effect of the choice set, the probability of choosing a familiar item increases with the size of the choice set, and that seems to occur for each information condition. This is consistent

[12]The difference between I^0 and I^- may be also used to test how useful information provided is for the subjects. The observation that they are more willing to prefer non-familiar to familiar items when information is provided indicates that this information is not totally irrelevant.

with the existence of choice overload. Moreover, it seems that the difference in the probability of choosing a familiar item linked to the choice set seems to increase with the information available: while the difference is of 3.5 % in the absence of information, it reaches 10.5 and 17.3 % for I^- and I^+.

To summarize, these raw descriptive measures suggest on the one hand the existence of choice overload, and on the other hand a more ambiguous effect of information: as for the standard decision-maker, providing some information increases the probability of choosing an unfamiliar option, but the quantity of information seems to have no effect, or a slightly detrimental effect on this probability.

4.3 Econometric analysis of subjects' preference for familiar items

To study these effects more in depth, we rely on a rank-ordered logistic model (Hausman and Ruud, 1987). Let i be an individual and ℓ the total number of individuals. Each option is referred to as j and J_i, the total number of options individual i faces. For each individual i, and each option in the choice set of i, U_{ij} is the latent score of j for i, with i prefers j to k if $U_{ij} > U_{ik}$. The latent score corresponds to the expected utility of option i in the model of section 2. U_{ij} can be decomposed as the sum of a deterministic term and a stochastic one: $U_{ij} = \upsilon_{ij} + \epsilon_{ij}$. As is usual in rank-ordered logit model, ϵ_{ij}'s are assumed to be iid with an extreme value distribution. And the deterministic term depends linearly on various variables of interest or controls (see below). The likelihood of obtaining a given rank order for individual i is given by: $L_i = \Pi_{j=1}^{J_i} \frac{e^{\upsilon_{ij}}}{\sum_{k=1}^{J_i} \delta_{ijk} e^{\upsilon_{ik}}}$ with δ_{ijk} the indicator function for the rank of j being lower than or equal to the rank of k for individual i. The log-likelihood for the sample of ℓ individuals is then given by:

$$\log L = \sum_{i=1}^{\ell} \sum_{j=1}^{J_i} \upsilon_{ij} - \sum_{i=1}^{\ell} \sum_{j=1}^{J_i} \log \big(\sum_{k=1}^{J_i} \delta_{ijk} e^{\upsilon_{ik}} \big)$$

We aim to estimate the impact of the context of choice on the effect of the familiarity of the option on the latent score U_{ij}, hence we specify the deterministic term of the score in the following way:

$$\upsilon_{ij} = \delta f_{ij} + f_{ij}[\gamma_1 t_{ij}^1 + \gamma_2 t_{ij}^2 + ... + \gamma_K t_{ij}^K] + \alpha x_{ij}$$

where f_{ij} is a dummy variable for the familiarity of option j to i, t_{ij} the variables of interest (treatments, risk attitude, gender, etc.), x_{ij} a set of option-based controls. The objective is to estimate the interaction of the variables of interest (treatments) with the familiarity of the option.

The estimations of several of these specifications are displayed in Table 6. In all the specifications, *IsFamiliar* and *IsHeardOf* are two dummy variables indicating whether the individual reported having experienced the website, or having heard of the website.[13] In the first set of specifications (1 and 2), C^+ indicates the Large Choice set treatment, *Info* is a dummy indicating whether information was provided in the treatment (I^- or I^+), I^+ simply indicating the corresponding treatment. In specification (2), individual controls have been added as independent variables. They include the result of the Holt and Laury measure, gender, and the type of content chosen by the individual. In the second set of specifications (3 and 4), independent variables are similar except for information: *RelativeInfo* is the ratio of the number of pieces of information by the number of unfamiliar websites,[14] and *TotalInfo* is the number of pieces of information available in the treatment.

Estimations displayed in Table 6 give support to our account of the simple descriptive statistics given in Table 5. First, a large choice set (C^+) is always associated with a stronger preference for familiar options (positive coefficient) and significantly so, suggesting a rather strong effect of choice overload. Second, we find confirmation for the fact that, providing some information reduces the propensity to choose a familiar item, but above some threshold, more information does not seem to have a negative impact, if any, on the preference for familiar items.[15] The second set of models (3 and 4) suggests that the relative information may matter positively (how well informed a subject is about a particular option), but that the absolute level does not seem to have an effect.

The last model (5) allows to test on a subsample whether the effect of the size of the choice set is robust to maintaining the total number of pieces of information constant. We do so by keeping the data from treatments where *the same absolute number of pieces of information is constant while the choice set varies*, that is we only keep conditions C^-/I^0, C^+/I^0, C^-/I^+ and C^+/I^-. For the first two conditions, there is no information at all so only the choice set changes while, in the last two conditions, the number of pieces of information is fixed to 6. We find that a large choice set is associated with a higher latent value for familiar items, suggesting once more that pure choice overload is at work. The estimated effect is only weakly significant, which can be partly explained by the decrease in statistical power due to the use of a smaller sample. It is noteworthy that, in particular if subjects are not information overloaded, the negative effect is still present despite the fact that a greater choice set (at least for C^-/I^+ and C^+/I^-) corresponds to subjects being much less well informed about non-familiar options.

Regarding controls, it is noteworthy that their presence or absence, or changes in the variables included, have very little effect on the estimated coefficient related to the variables of interest. We observe that some content categories lead to some preferences for less familiar websites ("Cook-

[13]When subjects were asked to report whether they have already visited the website before they run the experiment, they aslo have to indicate whether they have heard of the website. In the rank-ordered logistic model, this variable *IsHeardOf* corresponds to the set of option-based controls.

[14]As a consequence, two observations are not usable: the subjects were familiar with all the websites.

[15]We obtain similar results by working on ranks or by focusing on the preferred item of the choice set. However, these measures are less relevant regarding our data.

	Endogeneous variable: latent score for options				
	(1)	(2)	(3)	(4)	(5)
	Coef./(se)	Coef./(se)	Coef./(se)	Coef./(se)	Coef./(se)
IsHeardOf	.616***	.617***	.604***	.608***	.589***
	(.078)	(.078)	(.078)	(.078)	(.096)
IsFamiliar	2.151***	2.220***	1.405***	1.593***	2.349***
	(.212)	(.389)	(.165)	(.359)	(.238)
IsFamiliar interacted with					
C^-	(ref)	(ref)	(ref)	(ref)	
C^+	.673***	.655***	.754***	.605**	.360$^+$
	(.164)	(.171)	(.196)	(.208)	(.212)
Info	-1.572***	-1.702***			-1.490***
	(.214)	(.224)			(.216)
I^+	.576***	.587***			
	(.167)	(.176)			
RelativeInfo			-.040	-.263*	
			(.033)	(.121)	
TotalInfo			-.086	-.009	
			(.112)	(.035)	
Controls	No	Yes	No	Yes	No
n	341	341	339	339	215
LogLik.	-2133	-2107	-2151	-2124	-1320
χ^2 (df)	87 (5)	138 (14)	38 (5)	92 (14)	58 (4)
p-value	$<.001$	$<.001$	$<.001$	$<.001$	$<.001$

Significance levels. $^+$: $p < .10$; *: $p < 0.5$; **: $p < 0.01$; ***: $p < 0.001$. To compute likelihood ratio test (χ^2), we use as the null model the one with the two variables IsHeardOf and IsFamiliar.

Table 6: Rank-ordered logit estimates

ing" in particular). Women seem to significantly value more familiar choices, while the Holt and Laury risk aversion measures are associated with a positive coefficient (more preference for familiar options) but insignificantly so.

5 Concluding remarks

In this paper, we investigate how two dimensions of the context of choice, namely information and size of the choice set, affect the type of the chosen option. This allows us to test and measure choice overload, and the respective contribution of the information level and the size of the choice set to this phenomenon. To do so, we run an experiment, built on a clear theoretical distinction between choice overload and information overload. In this experiment, subjects were invited to rank items in different contexts where the number of alternatives and the information pieces vary. As all choice sets consisted of familiar and unfamiliar goods, we are able to use the preference for familiar goods as a proxy for overload.

Our main findings are the following. First, our results shed some light on the so-called choice paradox, in allowing to identify the source of choice overload. It seems that increasing information has a mixed effect on the decision maker. On the one hand, providing information leads our subjects to prefer a familiar item less frequently, as would be expected in a standard model where the availability of information reduces the uncertainty surrounding alternatives. On the other hand, the effect seems non-monotonic in the sense that an intermediate level of information seems to induce a less frequent preference for familiar items than high level of information. This suggests that, beyond a certain threshold, reading more information may entail some cognitive costs in processing and memorizing all information. Yet, this effect seems to interact with the size of the choice set and calls for a more specific work on how individuals process information. In line with recent studies on attention or memory (see for instance Reutskaja et al., 2011; Brocas et al., 2014), implementing technologies such as MouseLab or eye-tracking measures can highlight information process and how the context of choice impact it.

Finally, the clearest result of our experiment is that individuals who face a small choice set prefer familiar items less frequently than those who face a large choice set. This is the typical signature of choice overload. Subjects seem to experience more difficulties, or more confusion, in ordering options when provided a large choice set. This results is robust to controls for the absolute level of information. As a consequence, rather than in a mere informational overload, the choice paradox seems to stem from difficulty in establishing one's preferences over options. Further exploration of its potential sources like anticipated regret, loss aversion, thinking aversion or limited attention provides avenues for future research.

This result provides new evidence of the choice paradox and emphazises the pure effect of the choice set size. This is consistent with Iyengar and Kamenica (2010) who showed that individuals facing larger choice sets exhibit stronger preferences for simple and easy-to-understand alternatives. But our experiment also sugest that the main source of this effect is the size of the choice set, rather than information processing. It is noteworthy that this result is obtained in relatively simple conditions of choice: in the experiment, individuals face a maximum of only eight alternatives. Extrapolating to everyday external conditions, where economic agents face hundreds of alternatives, this would lead to a massive effect on people's decisions. This bias towards familiar products may lead to some inertia in consumption patterns: If consumers tend to stick to familiar or usual products, they may be reluctant to consider and experience new-comers or alternative options. In a long-term perspective, this may lead to stronger individual consumption routines as choice sets seem to expand endlessly in contemporary economies.

References

Ariely, D. and Wallsten, T. S. (1995). Seeking subjective dominance in multidimensional space: An explanation of the asymmetric dominance effect. *Organizational Behavior and Human Decision Processes*, 63(3):223–232.

Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience*, 4:829–839.

Bhargava, S. and Loewenstein, G. (2015). Behavioral economics and Public Policies 102: Beyond nudging. *American Economic Review: Papers and Proceedings*, 105(5):396–401.

Bhargava, S., Loewenstein, G., and Sydnor, J. (2015). Do individuals make sensible health insurance decisions? Evidence from a menu with dominated options. NBER Working Paper No. 21160.

Brocas, I., Carrillo, J. D., Wang, S. W., and Camerer, C. F. (2014). Imperfect choice or imperfect attention? Understanding strategic thinking in private information games. *The Review of Economic Studies*, 81(3).

Cronqvist, H. and Thaler, R. H. (2004). Design choice in privatozed social-security systems: Learning from the swedish experience. *American Economic Review*, 94(2):424–428.

Danan, E., Guerdjikova, A., and Zimper, A. (2012). Indecisiveness aversion and preference for commitment. *Theory and Decision*, 72(1):1–13.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2):171–178.

Hausman, J. A. and Ruud, P. A. (1987). Specifying and testing econometric models for rank-ordered data. *Journal of econometrics*, 34(1):83–104.

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

Iyengar, S. and Kamenica, E. (2010). Choice proliferation, simplicity seeking, and asset allocation. *Journal of Public Economics*, 94(7-8):530–539.

Iyengar, S. and Lepper, M. (2000). When choice is demotivating: Can one desire too much of a good thing? *Journal of Personality and Social Psychology*, 79(6):995–1006.

Iyengar, S. S., Jiang, W., and Huberman, G. (2004). How much choice is too much: Determinants of individual contribution in 401k retirement plans. In Mitchell, O. and Utkus, S., editors, *Developments in decision-making under uncertainty: Implications for retirement plan design and plan sponsors*. Oxford: Oxford University Press.

Kamenica, E. (2008). Contextual inference in markets: On the informational content of product lines. *American Economic Review*, 98(5):2127–2149.

Le Lec, F. and Tarroux, B. (2012). On attitude towards choice : Some experimental evidence of choice attitude. CREM Working Paper.

Leonard, T. C. (2008). Richard H. Thaler, Cass R. Sunstein, Nudge: Improving decisions about health, wealth, and happiness. *Constitutional Political Economy*, 19(4):356–360.

Loewenstein, G. (1999). Is more choice always better? *National Academy of Social Insurance*.

Loomes, G. and Sugden, R. (1982). Regret theory: An alternative theory of rational choice under uncertainty. *The Economic Journal*, 92(368):805–824.

Ortoleva, P. (2013). The price of flexibility: Towards a theory of thinking aversion. *Journal of Economic Theory*, 148(3):903–934.

Pratt, J. W. (1964). Risk aversion in the small and in the large. *Econometrica*, 32(1-2):122–136.

Reutskaja, E., Nagel, R., Camerer, C. F., and Rangel, A. (2011). Search dynamics in consumer choice under time pressure: An eye-tracking study. *American Economic Review*, 101(2):906–926.

Reutskaya, E. and Hogarth, R. M. (2009). Satisfaction in choice as a function of the number of alternatives: When goods satiate. *Psychology and Marketing*, 23(3):197–203.

Roese, N. J. (1997). Counterfactual thinking. *Psychological Bulletin*, 121(1):133–148.

Rosen, S. (1981). The Economics of Superstars. *American Economic Review*, 71(5):845–58.

Sarver, T. (2008). Anticipating regret: Why fewer options may be better. *Econometrica*, 76(2):263–305.

Scheibehenne, B., Greifeneder, R., and Todd, P. (2010a). Can there ever be too many options? a meta-analytic review of choice overload. *Journal of Consumer Research*, 37:409–425.

Scheibehenne, B., Greifeneder, R., and Todd, P. (2010b). Less may be more when choosing is difficult: choice complexity and too much choice. *Acta Psychologica*, 133:45–50.

Shin, J. and Ariely, D. (2004). Keeping doors open: The effect of unavailability on incentives to keep options viable. *Management Science*, 50(5):575–586.

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, 69(1):99–118.

Sunstein, C. R. and Thaler, R. H. (2003). Libertarian paternalism is not an oxymoron. *The University of Chicago Law Review*, 70(4):1159–1202.

Tversky, A. and Shafir, E. (1992). Choice under conflict: The dynamics of deferred decision. *Psychological Science*, 3(6):358–361.

Tversky, A. and Simonson, I. (1993). Context-dependent preferences. *Management science*, 39(10):1179–1189.

Appendix suggested for online publication

A Theoretical framework: Proofs

A.1 Equation (3)

For the sake of simplicity, we skip here all indices x, C and I.
For some specific values z and s for the random variables τ and μ, and assuming the third derivative of u is null or negligible, we can develop an usual Taylor series on the variable $z+s$:

$$\begin{aligned} u\big(m+z+s\big) &\simeq u(m)+u'(m)(z+s)+\frac{1}{2}u''(m)(z+s)^2 \\ &\simeq u(m)+zu'(m)+\frac{z^2}{2}u''(m)+su'(m)+\frac{s^2}{2}u''(m)+zsu''(m) \end{aligned}$$

Since $\mathbb{E}(\tau)=\mathbb{E}(\mu)=0$ by assumption, we obtain:

$$\begin{aligned} \mathbb{E}\Big[u\big(m+\tau+\mu\big)\Big] &\simeq u(m)+\frac{\mathbb{E}[\tau^2]}{2}u''(m)+\frac{\mathbb{E}[\mu^2]}{2}u''(m)+\mathbb{E}[\tau\mu]u''(m) \\ &\simeq u(m)+\frac{\sigma^2(\tau)}{2}u''(m)+\frac{\sigma^2(\mu)}{2}u''(m)+\sigma(\tau,\mu)u''(m) \end{aligned}$$

A.2 Propositions 1 to 3

We consider two random goods $f\in F$ and $n\in N$, that is a familiar good and an unfamiliar one. Their real subjective value is determined randomly with m_n and m_f. In the case of the familiar item, its expected utility is given by assumption 3 by $u(m_f)$ while that of the unfamiliar goods is given by:

$$\mathbb{E}\big[u(\omega_n(C,I))\big]=u(m_n)+\frac{\sigma^2(\tau_n^I)}{2}u''(m_n)+\frac{\sigma^2(\mu_n^C)}{2}u''(m_n)+\sigma\big(\tau_n^I,\mu_n^C\big)u''(m_n)$$

Given that $u''<0$ and $\sigma\big(\tau_n^I,\mu_n^C\big)$ is positive and non-decreasing in $\sigma^2(\tau_n^I)$ and $\sigma^2(\mu_n^C)$, we have $\mathbb{E}\big[u(\omega_n(C,I))\big]$ that is decreasing in $\sigma^2(\tau_n^I)$ and $\sigma^2(\mu_n^C)$.

Let F_I and F_J be the cdf of resp. $\mathbb{E}\big[u(\omega_n(C,I))\big]$ and $\mathbb{E}\big[u(\omega_n(C,J))\big]$. Note that these expectations are random variables due to the stochastic nature of m_n and m_f. We have $F_I(x)<F_J(x)$ for any x if $\sigma^2(\tau_n^I)<\sigma^2(\tau_n^J)$. Denote g the pdf of $u(m_f)$, $\mathbb{P}\Big[\mathbb{E}\big[u(\omega_n(C,I))\big]<u(m_f)\Big]$ is given by $\int_R F_I(x)g(x)dx$ while $\mathbb{P}\Big[\mathbb{E}\big[u(\omega_n(C,J))\big]<u(m_f)\Big]$ is given by $\int_R F_J(x)g(x)dx$. Since $F_I<F_J$, we have

$$\begin{aligned} \mathbb{P}\Big[\mathbb{E}\big[u(\omega_n(C,I))\big]>u(m_f)\Big] &> \mathbb{P}\Big[\mathbb{E}\big[u(\omega_n(C,J))\big]>u(m_f)\Big] \\ &\Leftrightarrow \\ \mathbb{P}[n\succ f\mid C,I] &> \mathbb{P}[n\succ f\mid C,J] \end{aligned}$$

Following the same reasoning, we obtain for C and D with $\sigma^2(\mu_n^D) > \sigma^2(\mu_n^C)$ that:

$$\begin{array}{rcl} \mathbb{P}\Big[\mathbb{E}\big[u(\omega_n(C,I))\big] > u(m_f)\Big] & > & \mathbb{P}\Big[\mathbb{E}\big[u(\omega_n(D,I))\big] > u(m_f)\Big] \\ & \Leftrightarrow & \\ \mathbb{P}[n \succ f \mid C, I] & > & \mathbb{P}[n \succ f \mid D, I] \end{array}$$

The first part of proposition 1 ensues directly from the fact that $\sigma^2(\tau_n^I) < \sigma^2(\tau_n^J)$ when $I > J$. Proposition 2 stems from the fact that $\sigma^2(\tau_n^I) < \sigma^2(\tau_n^J)$ when $I < J$. Similarly, part 2 of proposition 1 is implied by $\sigma^2(\mu_n^C) = \sigma^2(\mu_n^D)$ while proposition 3 comes from $\sigma^2(\mu_n^C) < \sigma^2(\mu_n^D)$ when $\#C < \#D$.

B Instructions (translated from the French)

Instructions

You are going to participate in two experiments. You will be informed about everything you need to know to participate in the study. You can earn a monetary amount during these experiments. The amount of money you can earn depends on your decisions. It is therefore important to make them carefully.

During this study, anonymity is guaranteed, i.e.,, your decisions cannot be linked with you. At the end of the experiment, you will confidentially receive the monetary amount you earned during the experiment in a check. You already receive an initial payment of €12 at the beginning of the study. All the earnings we mentioned in the instructions will add to this initial payment of €12.

Please note that communication is strictly forbidden during the entire experiment. It is also forbidden to use documents and materials (such as mobile phones) other than the ones made available to you at the beginning of the experiment. Communicating or playing around with the computer leads to exclusion from the study. We also inform you that you may only use the functions on the computer that are necessary for completing the study. In case of non-respect of one of these rules, you will be excluded from the experiment and all earnings will be lost.

If you have any questions, please raise your hand. We will then come and answer your question at your seat.

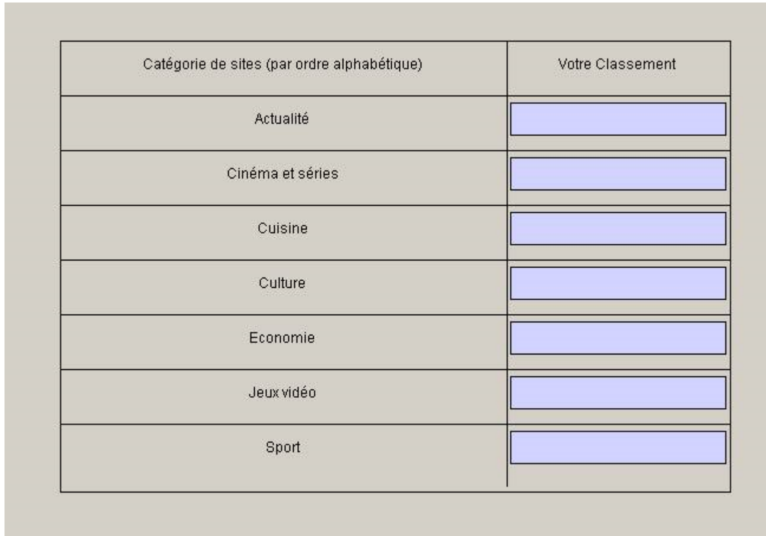
Procedure for the first experiment

Experiment 1 consists of 2 parts. In part 1, you will answer questions about various websites. In part 2, you will have access to one website for 30 minutes. The website to which you will have access will depend on the decisions you made in part 1. It is then important to answer them carefully.

1. Preliminary questions

Before these two parts, you have to rank different content categories of websites (news, cooking, sport, etc.). Your ranking will determine which kind of websites you have access to in part 2. More precisely, you will have access to one of websites belonging to the category you rank first.

The following screen displays the website categories you have to rank between 1 (the most preferred category) and 7 (the least preferred one):



You have to assign an integer between 1 and 7 to each category according to your preferences. Once your ranking is made, click on the OK button (bottom-right of the screen) to pass to part 1.

Note that you cannot give the same rank to different website categories. Thus you can click on the OK button only if each category has a different rank to the others.

2. Procedure for part 1

Part 1 of the experiment consists of 4 stages.

During the **first stage**, you will be informed about the name of the websites to which you can have access in part 2. There are [*4/8 depending on the treatment*] available websites.

Here is an example of the screen:

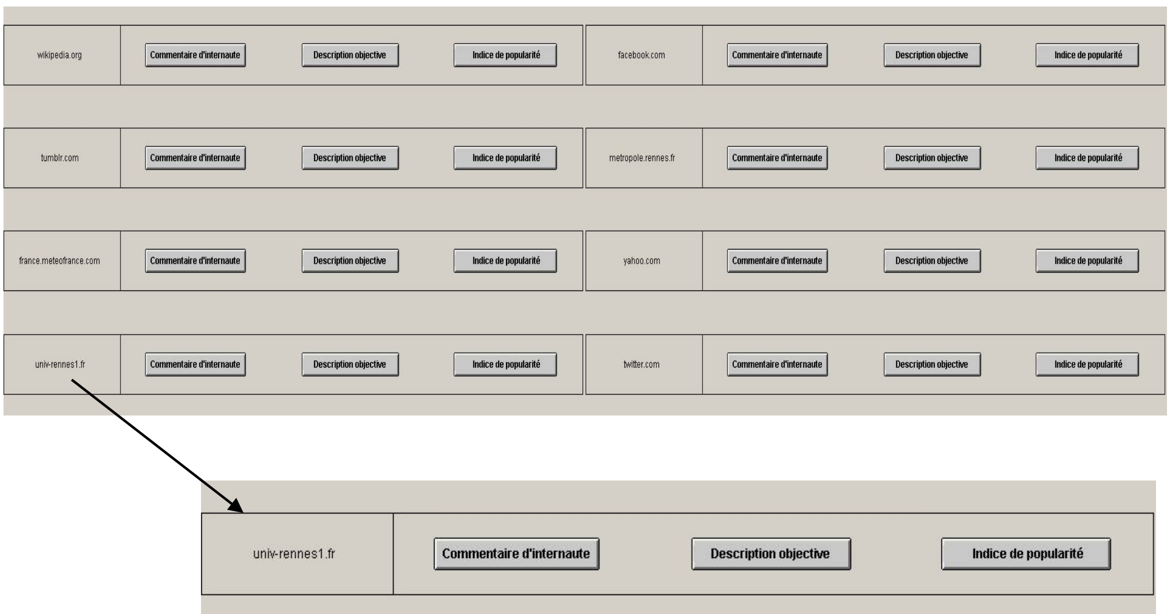
Once you take note of them, click on the OK button.

During the **second stage**, you have the opportunity to be informed of the various websites. Three types of information is available:

- An *objective description* which provides the basic characteristics of the website;
- An index of popularity which is the number of times that people declared to like the website on facebook.com, i.e., the number of Facebook likes;
- A randomly picked commentary of an internet user.

Note that you cannot read all the information pieces. More precisely, you have to read a total of [*0/3/6/12 depending on the treatment*] information pieces on any website you want.

The screen is of the following type [*treatment C+/I+*]:



Clicking on an information button opens a popup window with the information. Once you read the information, click on the OK button. The above screen appears again.

Since you have to read [*0/3/6/12 depending on the treatment*] information pieces, you have to repeat this [*0/3/6/12 depending on the treatment*] times. Note that you cannot read the same information several times. A piece of paper and a pen are at your disposal if you want to write down the information you read.

Once you have read the [*0/3/6/12 depending on the treatment*] information pieces, the third stage automatically starts. However you have to wait until all the participants of the experiment have read all the information pieces.

During the **third stage**, you have to rank the websites between 1 (the most preferred one) and [*4/8 depending on the treatment*] (the least preferred one).

The task is performed on the following screen:

Once you have made your ranking, click on the OK button to move to the next stage. Note that you cannot give the same rank to different websites. Thus you can click on the OK button only if each website has a different rank to the others.

Note that your ranking determines the second part of the experiment, where you have access to a certain website. More precisely, one of the websites is randomly picked but the probability that a website is chosen depends on its rank [*treatment C+/I+*]:

	Probability of being chosen
Website you ranked 1st	45%
Website you ranked 2nd	25%
Website you ranked 3rd	15%
Website you ranked 4th	7%
Website you ranked 5th	5%
Website you ranked 6th	2%
Website you ranked 7th	1%
Website you ranked 8th	0%

It means that the website you rank 1st has 45 chances in 100 to be chosen by the computer; the website you rank 2nd has 25 chances in 100 to be chosen; the website you rank 3rd has 15 chances in 100 to be chosen, etc. The **choice made by the computer will be implemented during the second part of the experiment**.

Ranking with respect to your preferences allows you to be more likely to have access to a website you really want to consult than one for which you have a weak preference to consult. That is to say, it is in your own interest to rank the websites carefully and honestly.

For example, let us suppose that the website you rank in the 1st position is univ-rennes1.fr. The probability that this website is selected by the computer is then 45% (see above table). If so, you have access to this website during 30 minutes in the second part of the experiment and visiting other websites is impossible.

Note that you are informed about the selection of the website at the beginning of the second part of the experiment.

After that, a **fourth stage** starts. You have to indicate whether you previously knew the website before the experiment and whether the available information was useful in order to rank the websites. Finally, the first part of the experiment is finished.

3. Procedure for part 2

Once part 1 is finished, part 2 of the experiment starts. This part lasts 30 minutes and works as follows.

You are first informed about the randomly selected website. In the previous example, you are informed that the website which is randomly drawn is: univ-rennes1.fr.

The selected choice is then implemented. For 30 minutes, you have access to a single website: the one that is randomly drawn (in the example, univ-rennes1.fr). Your browsing is limited to this website. If you try to open (intentionally or not) another webpage, a no-entry message will appear. Clicking on the red cross at the top-right of the webpage allows you to close it. You will then come back on the selected website. Please note that we cannot observe what you do on the website.

In you have any problem, do not hesitate to raise your hand. We will come and help you.

4. Evaluation of your satisfaction

Once the website is randomly selected, you have to give your level of satisfaction about the website. You will have to answer a series of questions.

****

After answering these questions, the first experiment is finished. You will be given the instructions of experiment 2.

Summary of experiment 1

Preliminary questions: ordering website content

Part 1:

- Stage 1: presentation of available websites (belonging to the content categories you prefer).
- Stage 2: information pieces about the websites.
- Stage 3: ranking the websites.
- Random draw/selection of a website according to the table below (you will be informed about the selected website at the beginning of Part 2).
- Stage 4: questions on your knowledge about the websites before the experiment and the use of the information pieces.

Part 2:

- Information about the randomly selected website
- 30 min browsing on the randomly selected website

Evaluation of your satisfaction: questionnaire

Probabilities that a website is chosen

	Probability of being chosen
Website you ranked 1st	45%
Website you ranked 2nd	25%
Website you ranked 3rd	15%
Website you ranked 4th	7%
Website you ranked 5th	5%
Website you ranked 6th	2%
Website you ranked 7th	1%
Website you ranked 8th	0%

Understanding the questions

In order to check your understanding of the instructions, please answer these statements (TRUE or FALSE).

We will give you the right answer aloud when all the participants have answered them. If, at the end of the correction, you have some questions, please raise your hand. We will then come and answer your question at your seat.

Q1. During experiment 1, I will have access to one website.

Q2. During experiment 1, I will have access to any website I would like.

Q3. The website I will have access to depends on my decisions.

Q4. During part 1, I will answer questions about 8 websites of the same content.

Q5. The content of the website available in stage 1 depends on the ranking I give in the preliminary question.

Q6. During part 1, I have to read 12 information pieces per items.

Q7. During part 2, an item I ranked is proposed.

Q8. During part 2, I can have access to all the websites I ordered and the access time depends on my ranking.

Q9. I have more chances to browse on the website I ranked first than on the one I ranked third.

Procedure for the second experiment

In this experiment, you can earn a certain amount of money, which adds to the initial endowment of €12. You have to make 10 decisions. Each decision consists in choosing between 2 options: option A and option B. For each option, you can earn some money/income with certain probabilities, as indicated on the screen that will be facing you.

The screen will indicate a list of 10 decisions. The table is of the following type:

Table: Choice of options

	Option A				Option B				Your choice	
Decision	Prob.	Earnings	Prob.	Earnings	Prob.	Earnings	Prob.	Earnings	A	B
1	10%	2	90%	1.6	10%	3.85	90%	0.1	○	○
2	20%	2	80%	1.6	20%	3.85	80%	0.1	○	○
3	30%	2	70%	1.6	30%	3.85	70%	0.1	○	○
4	40%	2	60%	1.6	40%	3.85	60%	0.1	○	○
5	50%	2	50%	1.6	50%	3.85	50%	0.1	○	○
6	60%	2	40%	1.6	60%	3.85	40%	0.1	○	○
7	70%	2	30%	1.6	70%	3.85	30%	0.1	○	○
8	80%	2	20%	1.6	80%	3.85	20%	0.1	○	○
9	90%	2	10%	1.6	90%	3.85	10%	0.1	○	○
10	100%	2	0%	1.6	100%	3.85	0%	0.1	○	○

Consider decision 4, option A gives you a chance of 40% to earn €2 and 40% to earn €1.60. Option B gives you a chance of 40% to earn €3.85 and 40% to earn €0.10. You have to make a choice between options A and B.

For each decision, you have to give your choice between A and B clicking on the option you prefer in the column on the right ("your choice").

Once you make your 10 decisions, the computer will randomly select one of them. Each decision has the same probability of being selected.

Next, the computer will randomly choose a number between 1 and 10, that determines the earnings associated with the option you have chosen.

You will be informed about your earnings and your total earnings.

Example.

Suppose that the computer selects the first line for the calculation of your earnings. For this line, option A gives you €2 if the randomly drawn number is 1 and €1.60 if it is between 2 and 10. Option B gives you €3.85 if the randomly drawn number is 1 and €0.10 if it is between 2 and 10.

If you have chosen A for this line and the computer randomly selects the number 1, then your earnings from experiment 2 are €2. The earnings are calculated in the same way for each of the decisions.

To sum up, you have to choose 10 times between an option A and an option B. When you make all the decisions, the computer randomly selects one of your decisions. Next, it randomly chooses a number, which determines your earnings.

The experiment is then finished.

****

If you have any questions, please raise your hand. We will come and answer your question at your seat.