Can Information Be Too Much? Information Source Selection and Beliefs[*]

Andrea Amelio [†]

October 17, 2023

Abstract

Agents undertaking economic decisions are exposed to an ever-increasing amount of information sources. This paper investigates how the number of available information sources impacts agents' ability to (i) select reliable sources and (ii) use their content effectively to update their beliefs. To answer these questions, I set up an online experiment informed by a simple automata decision-making and belief-updating model. Participants' source selection performances deteriorate as the number of available sources increases. Also, ceteris paribus, their performance in updating their beliefs using the selected sources worsens, showing a trade-off between source selection and belief updating performances. These results may help to guide policy-making decisions, providing evidence on externalities of information production.

Keywords— *Information Overload, Complexity, Belief Updating, Experiments*
JEL Codes: D91, C91, D83

[*]Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2126/1 – 390838866.

[†]University of Bonn, Bonn Graduate School of Economics (BGSE): aamelio@uni-bonn.de

1 Introduction

Addressing the question of whether access to expanding information sources is beneficial is arguably both timely and crucial, given the relentless surge of information individuals are exposed to, a phenomenon that's seemingly beyond our immediate control. Standard economic theory postulates that having access to a larger amount of information sources may only be beneficial. However, recent literature on the impact of complexity on decision-making, rooted in Simon (1955), shows how features of the decision environment shape the outcomes of decisions and belief formation (Caplin, Dean, and Martin, 2011; Enke and Zimmermann, 2017; Oprea, 2020; Enke and Graeber, 2023).

The main contribution of this paper is to show how an increase in available information sources hinders i) source selection performance and ii) the ability to make correct inferences from the available information. I perform an online experiment, divided into two main parts. First, participants are presented with a list of information sources and are required to select one. Sources are presented in a way such that it is possible to rank them according to their precision. Participants are taught to recognize more informative sources and their understanding is tested before the main part of the experiment. Second, the selected source generates a signal concerning an unobservable, binary, state, and participants provide their posterior about that state. The provided signal is generated following the data-generating process details, which are fully disclosed to participants. In other words, participants have all the information to provide a rational guess about the probability of each state.

The experiment provides two key results, directly connected to the main contributions of this paper. Participants are 18% less likely to select the best information source as the number of available sources increases from 10 to 40. Moreover, belief updating performances decrease by approximately 7%, comparing the case with the highest number of sources with the one with the lowest, exhibiting a trade-off with source selection: intriguingly, selecting a better information source decreases belief updating performances, ceteris paribus. These findings are relevant in understanding how the complexity of the information environment may impact both the selection of information sources, but also, crucially, how individuals make inferences using those sources.

The results are consistent with a model in which finite working memory is allocated between source selection and belief updating tasks. Both selecting an information source and making an inference are costly in terms of working memory. I formalize a simple model, based on the automata literature,[1] in which an increase of available sources induces the decision-maker to switch from rational choice to less burdening source selection rules (as in Salant, 2011). Additionally, when more cognitive resources are depleted in selecting an information source, belief updating rules have to be coarser and hence less precise (similarly to Leung, 2020). In the model, this happens because both source selection and belief updating make use of the same pool of finite cognitive resources. This fact induces a trade-off between how well an agent can select an information source and how well they will be able to use it. I follow the idea that choosing the optimal element from a list may be far more complex than implementing other

[1]See Ehud (1990) and Chatterjee and Sabourian (2009) for a review.

selection rules. For this reason, people may fail to select the best available option. Similarly, for what concerns belief updating, I relate to a small theoretical literature that connects finite cognitive abilities to the emergence of biases such as conservativism (Compte and Postlewaite, 2012; Wilson, 2014) and confirmation bias (Wilson, 2014; Leung, 2020).

The key novel contribution of this paper is twofold. First, I provide evidence of complexity playing a role in the domain of information source selection. Second, I relate information source selection and inference, documenting a trade-off between source selection and belief updating performances, in line with the theoretical framework.

This paper ties into several literature branches. First, this paper adds to the literature on the relationship between choice and complexity. A result of this literature is that people may fail to make the optimal choice when exposed to a large number of options (Caplin, Dean, and Martin, 2011; Caplin and Dean, 2015; Lleras et al., 2017). In failing to apply rational choice, people may recur to other selection rules, such as satisficing (Caplin, Dean, and Martin, 2011), given its lower implementation complexity (Salant, 2011). Oprea (2020) shows how procedurally complex choices, in the automata sense, are harder to implement for participants and generate a higher willingness to pay to be avoided. I bring an empirical contribution to this literature, providing evidence of the role of complexity in the domain of information source selection. Indeed, while I partially build on Caplin, Dean, and Martin (2011) experimental design, I extend their evidence to a different, relevant, domain: participants are selecting information sources through which they will have to make inferences in a following belief-updating step.

Second, this paper contributes to the theoretical literature on complexity and beliefs. This body of literature sets itself apart by arguing that complexity affects choices through beliefs, rather than impacting decision-making directly. The combination of limited cognitive capacity (finite working memory states) and complexity has been theorized to generate conservativism (Compte and Postlewaite, 2012; Wilson, 2014), confirmation bias (Wilson, 2014; Leung, 2020), and non-Bayesian inference (Chauvin, 2023). This paper contributes to this literature by postulating and experimentally showing a relationship between information source selection and belief updating. I provide evidence of a trade-off between source selection and belief updating performance, which becomes starker as the complexity of source selection increases.

Third, in close relation with the first two literature branches, this paper is related to the business and marketing literature on *choice* and *information overload*. The former literature features a large number of works with a leitmotif: an exceedingly large amount of available options may be detrimental to the choice quality or ex-post satisfaction. Having to select between an extremely large number of products (Iyengar and Lepper, 2001; Chernev and Hamilton, 2009) or comparing products with a large array of attributes (Hoch, Bradlow and Wansink, 1999; Chernev, 2003; Greifeneder, Scheibehenne, and Kleber, 2010) may decrease: (i) the likelihood of purchase (Iyengar and Lepper, 2001; Chernev, 2003), (ii) the ex-post satisfaction and confidence in the choice (Hoch, Bradlow and Wansink, 1999; Botti and Ivengar, 2004; Haynes, 2009), and (iii) the choice quality (Diehl, 2005; Dijksterhuis et al., 2006).[2] Tightly related, the information overload

[2]For an exhaustive literature review and meta-analysis of the on choice overload and information overload literature see Chernev, Bockenholt, and Goodman (2015).

literature argues that providing a decision-maker with an overabundant amount of information[3] may lead to lower quality decisions (Jacoby, Speller, and Kohn Berning, 1974; Chen, Shang, and Kao, 2009; Splinder, 2011), to a lower decision satisfaction (Jacoby, March 1984; Reutskaja and Hogarth, 2009; Messner and Wänke, 2011) and, in the beliefs domain, to confirmation bias (Götte, Han, and Leung, 2020). [4] This paper contributes to these literature branches in that it provides evidence of choice overload in the novel domain of information sources. Additionally, this paper shows how intertwined choice and information overload are, reporting evidence of a trade-off between information source quality and belief updating performance.

More broadly, this paper contributes to the literature on the Information Age and misinformation. Information fruition and production underwent numerous and complex dynamics in the last decades. It is a common and accepted position that the advent of the Internet represented one of the major changes in this field. However, consensus on the mechanisms and the direction of these changes remains elusive. On the one hand, it has been argued that the internet and social media increase the risk of ideological segregation and the creation of ”filter bubbles” (Pariser, 2011; Flaxman, Goel, and Rao, 2016) and ”echo chambers” (Sunstein, 2001). On the other hand, also the point that the use of the same means may increase exposure to diverse ideas has been advanced and studied (Benkler, 2006; Gentzkow and Shapiro, 2011; Flaxman, Goel, and Rao, 2016).[5] It is indisputable, however, that the advent of the internet dramatically increased the number of information sources available to individuals, which is the point that this paper tries to address and investigate. I inform this debate by showing how having access to a large set of information sources may hinder source selection and the ability to make inferences. My results point toward the fact that information may generate negative externalities, with potential implications for regulators.

The remainder of the paper is structured as follows. Section 2 illustrates the theoretical frameworks, aimed at conveying the key intuitions and guiding the experimental investigation. Section 3 details the experimental design and procedures. Section 4 reports results on the relationship between the number of information sources and source selection rules, and Section 5 characterizes such rules. Section 6 reports evidence of the trade-off between belief updating and source selection performance. Section 7 concludes and discusses the relevance of these results in applied settings.

2 Theoretical Framework

In this section, I provide the theoretical framework that grounds and guides the experimental investigation. First, I provide a stylized representation of the decision problem, split in the

[3]In this, the two branches of literature partially overlap, as, in some instances of the information overload literature, the concept of information corresponds to the number of attributes of a choice (e.g. Jacoby, Speller, and Kohn Berning, 1974).

[4]See Roetzel (2019) for a more exhaustive literature review on information overload in business and related domains.

[5]For example, Golin and Romarri (2022) document a positive effect of the level of internet penetration in Spanish municipalities on reported attitudes towards migrants.

information source selection step and in the belief updating step. Second, I define how information sources are ranked in this context. Third, I provide an illustration of how cognitive limitations (or finite working memory states) impact source selection and belief updating, formulating testable predictions concerning the impact of said limitations on performances and on the emergence of a trade-off between the two steps. The illustration is based on the representation of decision rules through *automata* or *finite states machines* and is carried out following Salant (2011) for source selection and Leung (2020) for belief updating. Finally, I provide an example, illustrating the key points of the model. Importantly, the purpose of this section is not to provide a general model for source selection and belief updating, but rather to formally describe the intuitions on which this work is based and to guide the empirical investigation.

2.1 General Setup

Consider a decision maker (DM henceforth) whose optimal choice depends on an unobservable state of the world $\theta \in \Theta$. The DM holds some prior about the state of the world $P \in \Delta(\Theta)$ and has access to a set of L information sources, which will be defined more rigorously later on. For now, imagine that the DM has some criteria to rank the available information sources, with better information sources being synonymous with higher chances of a correct assessment of the state of the world θ and of an optimal choice. Hence, the DM's problem is to select a good, or the best if that exists, source and then update her beliefs on the base of the information received by the source. This paper focuses on these two steps of the decision process and on their relation.
The DM is assumed to be cognitively limited, in the form of finite working memory $M \in \mathbb{N}$: both *source selection* and *belief updating* are cognitively costly to implement. In order to lay out the relationship between finite working memory and both source selection and belief updating, these processes are represented using *finite state machine* or *automata*. These stylized representations allow to formally isolate the effect of finite working memory on source selection and belief updating. In what follows, I first formalize what an information source is in this framework, defining two ways to rank a set of sources. Given these ranking criteria, I define formally source selection and belief updating using automata, linking them to the main hypotheses tested in the experiment.

2.2 Information Sources

An *information source* I is a random variable $I : \Theta \times S \to S$, where Θ is the set of possible, unobservable, states of the world and S is the set of possible signals that the information source can generate. An agent using the information source can only observe the signals generated by the said source, although the outcome is defined through the state of the world $\theta \in \Theta$ and the drawn signal $s \in S$, given the state of the world.

For simplicity, and in line with the experimental design, consider the case of a binary state $\Theta = \{A, B\}$. Moreover, assume that the signal space corresponds to the state space, that is $S = \theta$. In other words, any information source generates only two possible signals: A or B. For any state θ, define the probability of truthful reporting for source I as $p^*_I(\theta) = p(s = \theta \mid \theta)$. In this binary setting, any source I can be fully characterized as $\{p^*_I(\theta)\}_{\theta \in \{A,B\}}$, that is any source can be described through the probability of truthfully reporting the state, for both possible states.

Definition 1. *(Source Dominance) A source I_1 is said to dominate source I_2 ($I_1 \succ I_2$) if $p^*_{I_1}(\theta) > p^*_{I_2}(\theta)$ for all $\theta \in \{A, B\}$.*

Hence, a source dominates another source if the probability of reporting the state truthfully is higher for any possible state. This definition of dominance induces a partial ordering on the set of possible sources, which has some implications for data analysis as will be discussed more in-depth in Section 4. An alternative way to compare sources, which instead induces a complete ordering, is the following.

Definition 2. *(Source Ranking) A source I_1 is ranked higher than source I_2 if $\mathbf{E}_\theta[p^*_{I_1}(\theta)] > \mathbf{E}_\theta[p^*_{I_2}(\theta)]$, that is $\sum_{\theta\in\{A,B\}} P(\theta)p^*_{I_1}(\theta) > \sum_{\theta\in\{A,B\}} P(\theta)p^*_{I_2}(\theta)$. Hence, given a set of information sources I, the rank of information source $I_j \in I$ is:*

$$R(I_j, I) = |\{I_k \in I : \mathbf{E}_\theta[p^*_{I_j}(\theta)] > \mathbf{E}_\theta[p^*_{I_k}(\theta)], k \neq j\}| + 1$$

.

In other words, a source is ranked higher than another if the ex-ante probability of truthful reporting is higher for that source. Given some set of sources of cardinality L, it is possible to define a *best* source unambiguously, in a way that is consistent with both dominance and ranking.

Definition 3. *(Best Source) Given a set of L sources $\{I_1, I_2, .., I_L\}$, if there exists $I_i \in \{I_1, I_2, .., I_L\}$ such that $I_i \succ I_j$ for all $j \neq i$, then:*

1. *I_i is the best source in $\{I_1, I_2, .., I_L\}$*
2. *I_i is the highest ranking source in $\{I_1, I_2, .., I_L\}$*

Point (ii) follows from the fact that dominance also implies higher ranking, while the opposite does not hold. Note that an equivalent definition of the best source was used to instruct participants during the experiment.[6]

2.3 Automata, Source Selection, and Belief Updating

I represent both source selection and belief updating through *finite state machines* or *automata*. The aim is to provide a common theoretical framework linking the two steps of the decision problem. This framework is convenient as it can naturally feature a decision-maker with finite working memory and a related definition of complexity, common to both source selection and belief updating.

In this section, I first provide a formal illustration of automata and report some relevant results, following Salant (2011). I then show how this framework can be applied to source selection and to belief updating, linking the two steps through the finite working memory of the decision-maker. Finally, I illustrate the predictions that are subsequently investigated experimentally.

[6]For further details on the instructions, see Appendix C.

Automata and Complexity

An automaton is a tuple of several elements. First, a finite set of memory states $\mathcal{M} = \{m_1, m_2, ..., m_{\mathcal{M}}\} \cup \{Stop\}$. A memory state represents the current information that the DM holds, which impacts how she computes the additional inputs she receives. For example, in the case of belief updating, a state represents the current belief held by the DM about the state of the world. When the $\{Stop\}$ state is reached, the automaton stops processing additional inputs and switching state.[7] A transition function $g : \mathcal{M} \times X \to \mathcal{M}$ determines how the DM switches between memory states, with X being the set of inputs the DM may receive. In the case of information source selection, X is the set of available sources, while in the case of belief updating X is the set of signals that the DM may observe to update her beliefs. Additionally, it is necessary to define an initial state $m_0 \in \mathcal{M}$, from which the transitions will start. Finally, only for the case of source selection, it is necessary to specify an output function $f : \mathcal{M} \times X \to X$, to determine which element is selected from the list of information sources. $f(m, x)$ is specified as follows: if $m = \{Stop\}$ or x is the last element of the list, then x is selected.

In this framework, it is possible to define an automaton's complexity. Following, Salant (2011) and Oprea (2020), I use *state complexity*:

Definition 4. *(State Complexity) Given an automaton with memory states $\mathcal{M}$, its state complexity is $|\mathcal{M}|$.*

Later in this section, I provide some examples of automata of different complexity, for both source selection and belief updating.

Source Selection

Let the set of possible sources be an ordered list or a vector of sources $I = [I_1, I_2, .., I_L]$, that is $I = X$ in this case. This formally introduces the idea of a DM that evaluates sources sequentially, following the order indicated by the vector index. A prominent example of this kind of sequential evaluation would be a user looking for a set of keywords on a search engine on the internet, in which the results of the search would appear in a specific order. This example has a broad application range, as information search in this fashion is quite common, and may extend to fields such as collecting information about medical treatments or referenda on technical issues.

The following proposition is a reformulation of two results from Salant's (2011) paper. The core idea is to show that, in the domain of source selection rules represented through automata, rational choice and satisficing represent an upper and lower bound in terms of complexity[8].

Proposition 1. *Consider a list of information sources of length L, $I = [I_1, I_2, .., I_L]$:*

[7] This state needs to be specified only for the source selection case, as the selection has to eventually stop and produce an output. Belief updating, instead, could potentially never stop. However, for convenience, I generally include the $\{Stop\}$ state in $\mathcal{M}$.

[8] An automaton implementing rational choice always selects the best source from the list, that is the source with the highest ranking. An automaton implementing some satisficing rule, instead, selects the first source in the list which satisfies some minimum precision requirement.

1. *The state complexity*[9] *of an automaton implementing rational choice is* $L-1$.

2. *The state complexity of a rule is 1 if and only if it is a satisficing rule.*

Combining the two points from the previous proposition, it is possible to draw two observations. First, as L increases, an agent with finite memory states will eventually have to switch to a different source selection rule. Second, the only rule that has always minimal state complexity is satisficing. From these two observations, two corresponding empirical predictions follow.

Prediction 1. *The share of participants correctly implementing rational choice decreases with* L.

Prediction 2. *As L increases the share of participants implementing a satisficing rule increases.*

Belief Updating

Here, I provide a formalization that is a simplified version of the one presented by Leung (2020). An automaton representing belief updating has the same components as one representing source selection, except for a stopping state and the related output function. This comes from the fact that a belief updating procedure may be iterated potentially infinitely many times. Moreover, the interpretation of the other components is also different. Each element of the set of memory states $\mathcal{M}$ represents a different belief the DM holds about the state of the world, with the initial state $m_0 \in \mathcal{M}$ representing her prior. The set of inputs X corresponds to the set of possible signals the DM may observe. Finally, the transition function is a (potentially stochastic) mapping $g : \mathcal{M} \times X \to \Delta\mathcal{M}$, which characterizes how the decision maker combines her current belief $m \in \mathcal{M}$ and the observed signal.
Importantly, in this setup, state complexity $|\mathcal{M}|$ also represents how fine-grained the belief updating can be: the more memory states are employed to represent beliefs, the larger the variety and the potential precision of those beliefs. Considering an agent with M memory states, and considering source selection and belief updating jointly, it is clear that the higher the state complexity of the source selection rule, the lower the state complexity, and hence the precision, of the belief updating rule. As previously discussed, the state complexity of source selection is related to both the number of available sources L and to the source selection rule, with rational choice representing the upper bound in complexity for a given L. From these considerations, an empirical prediction follows.

Prediction 3. *Belief updating performance decreases in L and in source selection performance.*

Belief updating performance is defined as the absolute distance of the reported belief from the Bayesian benchmark, as specified in Section 4. As, on average, better-performing source selection rules have a higher state complexity than satisficing, with such complexity increasing in L, better performance in source selection will correspond to fewer available memory states

[9]Note that for any source selection rule, that are infinitely many automata implementing that rule. For the purpose of this work, when considering state complexity of an automaton implementing a given source selection rule, I always refer to the state complexity of the *minimal* automaton implementing that rule, that is the automaton with the lowest state complexity which implements some source selection rule.

to allocate for belief updating. I now provide a working example conveying the main intuitions, before illustrating experimental design and results.

2.4 Example

Consider a DM with a working memory of $M = 4$, with a list of three information sources $I = [1, 2, 3]$, with $3 \succ 2 \succ 1$. Once the DM picks a source from the list, the source produces a signal S about the binary state $\Theta = \{A, B\}$, and the DM updates her beliefs.

First, consider an instance in which the DM applies rational choice to the list of information sources I. Following Proposition 1, and as represented in Figure 1, this source selection rule complexity would be equal to $L - 1 = 2$.

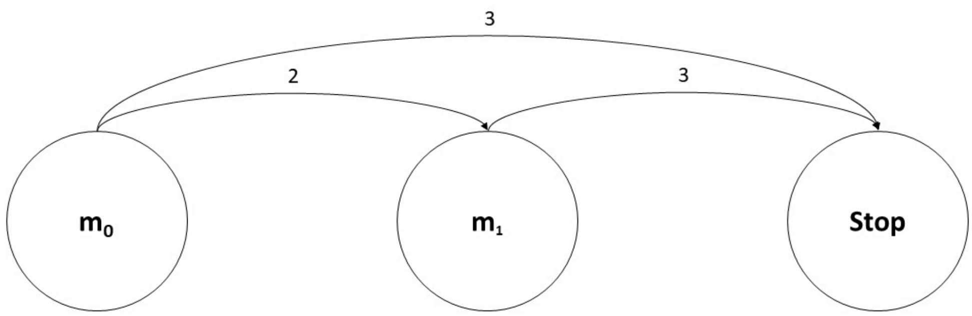


Figure 1: Automaton representing rational choice implemented for a list of three information sources.

Figure 2 represents a possible belief updating rule with the remainder working memory. Hence, the states would just be two: *"A is more likely"* and *"B is more likely"* in this case. When the DM observes $S \in S_A$, that is $P(S \mid A) > P(S \mid B)$ then she believes A to be more likely than A, and vice versa for the case of $S \in S_B$. Figure 3, shows a less coarse belief updating rule, that encompasses a third, intermediate, state: "*A and B are equally likely*". Clearly, this allows the DM to hold more fine-grained beliefs about the underlying state.

However, with a working memory of $M = 4$, the DM can not implement a finer belief updating rule without reducing the complexity of the source selection rule. Figure 4 shows an automaton for a satisficing rule with a threshold of 1: the first encountered source that is strictly better than 1 is selected. Unlike rational choice, this source selection rule would be implementable along with the belief updating rule in Figure 3, as satisficing complexity is always one.

This simple example stresses the intuition behind Prediction 3. On the one hand, keeping L constant, increasing the belief updating performance, through a finer rule, decreases source selection performance, and vice versa. On the other hand, as L increases, to keep the source selection performance constant, the DM has to opt for a coarser belief updating rule.

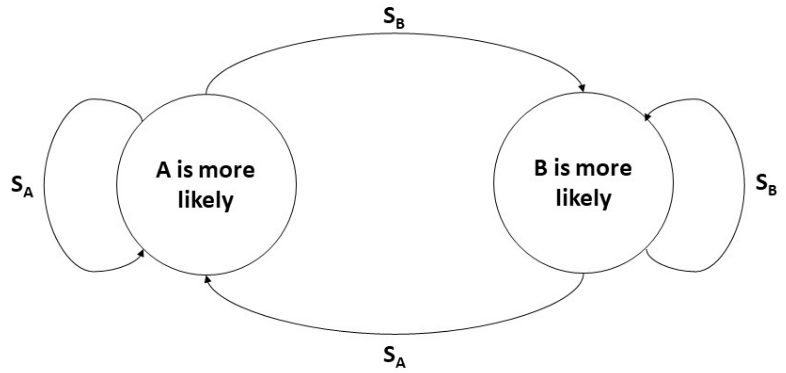


Figure 2: Automaton representing a belief updating mechanism with two memory states.

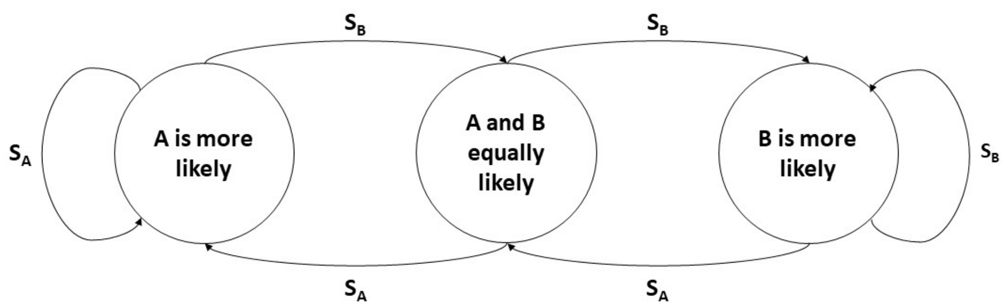


Figure 3: Automaton representing a belief updating mechanism with three memory states.

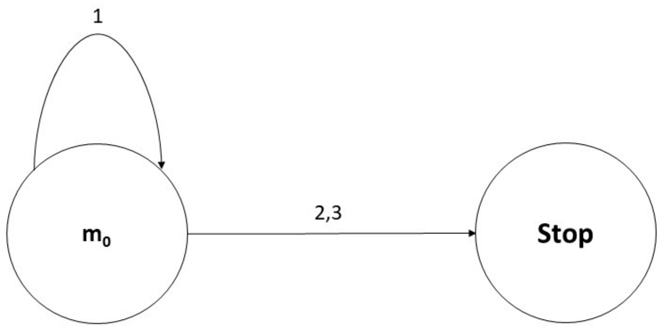


Figure 4: Automaton representing satisficing, with a threshold of 1, applied to a list of three information sources.

3 Experimental Design

An experimental framework to investigate how the number of available sources impacts source selection and related belief updating should have the following features: i) information source selection and belief updating should co-exist in the same task, ii) the decision-maker has to be able to distinguish good and bad sources (sources should be ranked), and iii) it should be possible to vary the number of available sources freely. The experimental design fulfills these requirements and consists of four stages: i) information source selection, ii) belief updating, iii) working memory task, and iv) final survey. The first two stages are repeated several times before moving to the next, to vary the task parameters and to collect multiple observations per participant (for a graphical summary of the design see Figure 5). In the next sections, I provide further details of stages (i) and (ii). The working memory task consists of a simple forward digit-span task.[10] In the final survey stage participants are asked about their age and education level.

[10]For a literature review on the use of these kinds of tasks as a measure of working memory, see Conway, Kane, Bunting, Hambrick, Wilhelm, et al. (2005).

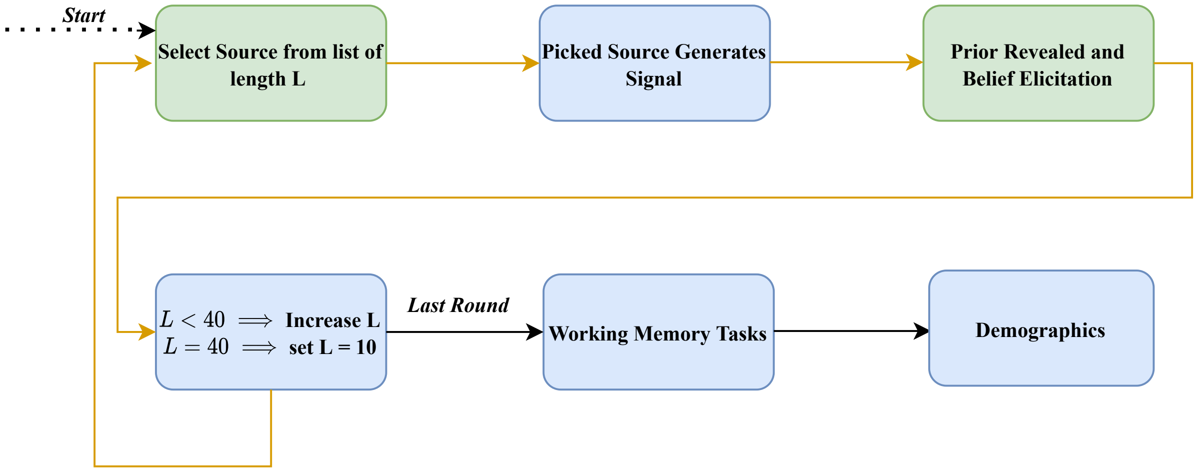


Figure 5: Summary of the experimental design. The green boxes represent financially incentivized tasks.

3.1 Information Source Selection

In each source selection task, participants observe a list of information sources of length L. Sources are represented as 2x2 tables, as shown in Figure 6.[11] The possible list lengths are $L \in \{10, 20, 40\}$, with L always being equal to 10 in the first task. After participants selected a source they undergo the associated belief updating task. Then, they face a new source selection task with a longer list, unless in the previously completed task $L = 40$, in which case the length starts back from 10. In total, each participant undergoes 9 source selection tasks, that is 3 repetitions for each possible list length.

Source Suggestion

Real State		A	B
	A	69%	31%
	B	28%	72%

Figure 6: Example of a source, as presented to participants. For any state $\theta \in \{A, B\}$, the diagonal elements represent the probability of a signal $s \in \{A, B\}$ being truthful $P(s = \theta \mid \theta)$.

Each source in a list is covered by a white block and can be uncovered by hovering over it with

[11] Participants are explained how to interpret the content of the table and their understanding is tested in a preliminary comprehension check. For further details about the instructions and the comprehension questions, see Appendix C

the cursor (see Figure 7). This setup ensures that participants can only evaluate one source at a time. Additionally, through the use of mouse tracking data, this design allows measuring which sources participants evaluated and how much time they spent evaluating them.

Figure 7: Example of sources in a list as presented to participants. In the first panel, the cursor is not hovering on any source, while in the second panel, the cursor is hovering on Source 2, uncovering the features of that source.

Sources Dominance and Best Source

Participants are explained that some sources are better than others and that each list always contains the best source. Source i is considered better than source j if and only if both diagonal elements of source i are larger than those of source j. This illustration is in line with the definition of dominance illustrated in Section 2. The order in which sources are presented in the list, and hence the position of the best source, is randomly determined in each round. Importantly, longer lists contain on average better sources, as well as the best source of longer lists always dominates the best source of shorter ones, as explained more in-depth below. The fact that the average source quality increases with list length is rooted in two considerations. First, this generates a tension, a trade-off, between the number of available sources and source quality, reproducing in a stylized way the idea that as the number of sources increases, it is also possible to find better sources. Second, this setup generates a framework in which studying the states guessed by participants is insightful. On the one hand, having better, more precise, sources should improve participants' chances of correctly guessing the unobservable state. On the other hand, the cognitive load induced by selecting a source from a longer list may hinder the gain of having access to better sources. This is in line with what I show in the results on state guesses: participants do not improve their state guesses for longer, although better, lists of sources.

Sources generating Algorithm

All the information sources used in the experiment are generated a priori, using an algorithm.

Recall that each source is characterized by the two probabilities of truthful reporting for each state, that is drawing an information source is equivalent to drawing these two probabilities. The algorithm was thought to implement three key criteria in randomly drawing the sources: i) each list of sources should contain a dominant source, ii) sources in longer lists should be on average more precise and iii) the dominant source in a longer list should always be dominant in the shorter one.

Before diving into how the algorithm works, it is convenient to define the *maximum precision* associated with a given length. The maximum precision ($M(L)$) is the highest possible probability of truthful reporting associated with a given length. Following criterium (ii), $M(10) = 70$, $M(20) = 75$ and $M(40) = 80$. The source-generating algorithm operates as follows, for each $(L, M(L))$ couple:

1. Draw 2 integers (one for each state) in the $[50, M(L)]$ interval.
2. Repeat this procedure for L times.
3. If there is not a dominant substitute the last source with a dominant one.
4. If the dominant source in the list dominates also all sources in the list with $L' < L$ then proceed to the next $L, M(L)$ couple.

Main and *Satisficing* Treatments

The condition in which participants are asked to select the best available source is the baseline, or *Main*, treatment. The experiment features an additional condition, the *Satisficing* treatment. The *Satisficing* treatment is identical to *Main*, except that participants are asked to select the *first* source in the list that meets a given precision requirement. More specifically, participants are asked to select the first source in the list with a probability of truthful reporting exceeding some threshold, for both states. There is a one-to-one mapping between the length of the information sources list and the used threshold.[12]

The goal of this treatment is to isolate the effect of a larger amount of available information sources on the computational costs of rational choice. Following the theoretical framework, the complexity costs of satisficing should not vary with list length: any impact length has on performance, in this case, should not be due to a more complex source selection rule. Hence, comparing *Main* with *Satisficing* treatments allows to distinguish the impact that increased list length has through the complexity channel, as rational choice becomes harder to implement, from other possible channels (e.g. longer lists may confuse participants).

Key Outcomes of Interest

There are three key outcomes from this task. The first is an indicator for the selected source being the *correct* source in the list. For the *Main* treatment, that source is the best or dominant source in the list. For the *Satisficing* treatment, that source is the one that fulfills the satisficing

[12]More specifically, the threshold is increasing in list length, as the average source quality is also increasing. The threshold for length 10 is 57%, for length 20 is 60% and for length 40 is 63%.

decision rule, that is the first source in the list that satisfies the specified precision requirements. The second outcome of interest is the ranking of the selected source,[13] constructed using the sum of its diagonal elements, that is the sum of the probabilities of truthful reporting from the source. The first measure can be used to study how the probability of selecting the best source varies, ceteris paribus, as the amount of available sources varies. The second measure, which is the operationalization of the ranking defined in Section 2, can be interpreted as a way to measure the quality of the selected source, to study both the extensive and the intensive margins of the relationship between the number of information sources and source selection.
The third outcome is the position of the selected source in the list. This variable is used to study participants' selection rule and how it varies with the number of available sources, in line with predictions.

3.2 Belief Updating

The belief updating tasks follow immediately after each source selection task. Participants have to provide their guesses about the probability of each state given some prior and a suggestion produced by the selected information source. Participants may observe at any time the source that they selected in the previous step, hovering over a box on the screen. The prior $P(A)$ varies in each of the 9 belief updating tasks, with the set of possible priors being $P(A) \in \{\frac{1}{10i}\}_{i=1}^{9}$, and the order being determined randomly. Figure 8 below shows an example of a belief elicitation screen.

[13]This outcome is only relevant for the *Main* treatment, as in the *Satisficing* treatment what is relevant for a successful implementation of the rule is not the goodness of the source, but to select the first source that meets the indicated precision requirements.

You picked this source:

Source Suggestion

Real State	A	B
A	62%	38%
B	35%	65%

The computer picked **state A with 50% probability** and **state B with 50% probability.**
The suggestion from the picked source is **B.**

Please state your **guess** about the **probability of each state**.

State A	0
State B	0
Total	0

Figure 8: Example of belief elicitation screen, as presented to participants. The screen is taken before any input is provided. After a guess about any of the two states is provided, the guess about the other state is automatically filled with the hundreds complement of the other guess.

Key Outcomes of Interest

The main outcome of interest for the analysis is the distance between the provided guess and the Bayesian benchmark, which in this case is simply the absolute difference between the guess from the participant and the normatively correct answer.[14] I also analyze the belief updating performance in terms of *implicit guess*: when a participant assigns more than 50% probability to a certain state, her implicit guess is that state. Hence, the state that a participant deems more likely is compared to the true, unobservable, state drawn by the computer. This measure allows to study how performances in guessing the true state, regardless of the Bayesian benchmark, vary as i) on the one hand the amount of information sources increases, while, ii) on the other hand, longer lists contain better sources on average and always contain at least one source that dominates all sources in shorter lists.

[14]This paper's focus is not to measure a specific, directional, bias (e.g. underinference, conservativeness, base rate neglect). Hence, the absolute value represents a fitting measure of the assessment quality.

4 Amount of information sources and source selection

The first result concerns the probability of selecting the best source from a list of information sources and how this probability varies as the number of available sources increases. Later, I show how this first result is robust to using relative source ranking as a measure of source selection performance.

Result 1. ***Non-rational source selection:*** *The probability of selecting the best information source decreases with the number of available sources.*

First, I report preliminary evidence on the relationship between the length of information sources lists and the share of *rational* choices implemented by participants, focusing on participants in the *Main* treatment. Figure 9 shows how the share of choices in which the best source was selected by participants decreases from approximately 70% with 10 available sources, to approximately 55% with 40 available sources.

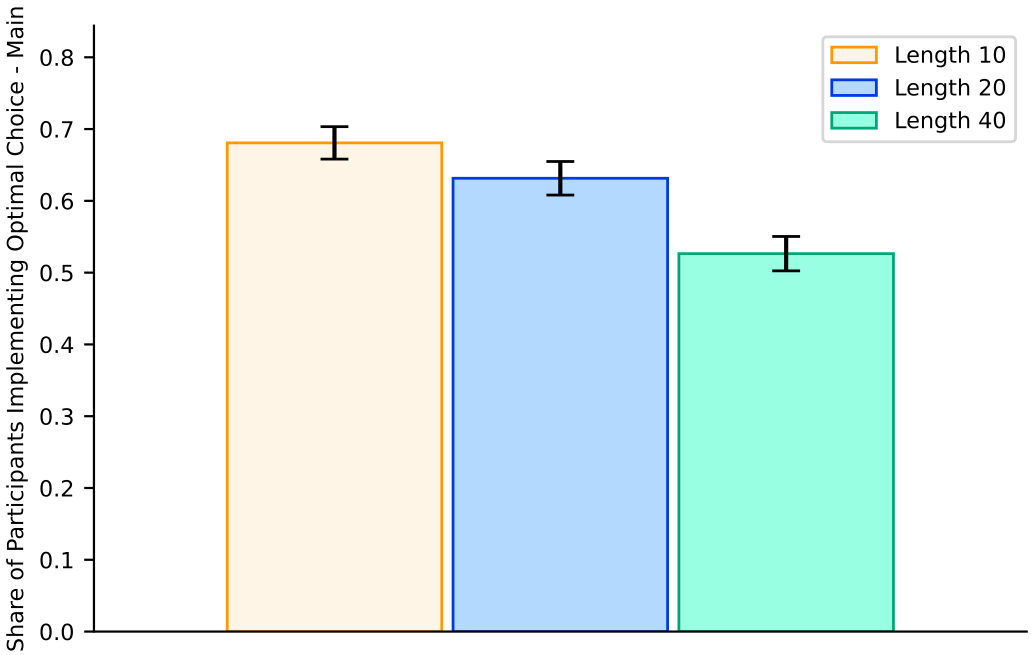


Figure 9: Share of participants selecting the best source, by list length. Error bars represent 95% *confidence intervals.*

Next, I provide more formal evidence of the pattern shown in Figure 9, estimating the following equation through OLS:

$$I(bestsource_i) = \alpha + \beta L_i + \gamma X_i + \varepsilon_i, \quad (1)$$

where $I(bestsource_i)$ is equal to one if the best source is selected in choice i. L_i is the length of the list of information sources in choice i. X_i is a set of control variables, among which the position of the best source in the list and the total amount of time spent hovering over sources in that specific round. As the features of a source were revealed only when hovering over it, the latter can be interpreted as a measure of time spent acquiring and elaborating information on the quality of the sources.

Table 1 reports the estimation results. Column (1) only includes the number of sources as the dependent variable. Columns (2) and (3) progressively include the best source position, the total time spent hovering on sources, and the performance in the working memory task, to the full specification in column (4). In all four specifications β, the main coefficient of interest is negative and significant. It is important to note how β is the estimated marginal effect of adding *one information source* to the list of available sources. Hence, according to the results, the probability of selecting the best source from the longest possible lists is 18% lower, compared to the choices in which the available sources are 10. As reported in Table B.1 in the Appendix, these results are robust to defining source selection performance using the *Source Relative Rank*. The latter is constructed by ordering sources in a list according to Source Ranking, as defined in Section 2, and dividing the resulting ordering by the number of sources in the list.[15]

	Dependent variable: Probability of Selecting Optimal Source			
	(1)	(2)	(3)	(4)
List Length	-0.006***	-0.005***	-0.006***	-0.006***
	(0.001)	(0.001)	(0.001)	(0.001)
Best Source Position		-0.002	-0.003*	-0.003
		(0.002)	(0.002)	(0.002)
Total Time on Sources			0.003***	0.003***
			(0.001)	(0.001)
Working Memory Proxy	✗	✗	✓	✓
Demographic Controls	✗	✗	✗	✓
Session FE	✗	✗	✗	✓
Priors	✗	✗	✗	✓
Observations	1,237	1,237	1,237	1,237
R^2	0.022	0.022	0.104	0.111

Notes. OLS estimates, robust standard errors are clustered at the subject level. The dependent variable is an indicator, equal to one if the selected source corresponds to the best available one. $^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$

Table 1

To ensure that results are not driven by other factors related to the length of the list (e.g. participants are confused by longer lists), I compare the *Main* and the *Satisficing* treatment. In the latter, participants are required to select the first source in the list that fulfills a given precision requirement for both states. Following the theoretical framework, the complexity of this source selection rule does not vary with list length. Hence, if Result 1 depends on the increased complexity of rational choice, and not on any other factor related to length, two related predictions follow: i) the success rate of implementing satisficing does not decrease with

[15]More formally, given a set of L sources $I = \{I_1, I_2, ..., I_L\}$, let source j ranking, according to Definition 2, be $R(I_j)$- Then I_j relative ranking is $R(I_j)/L$.

the number of available sources, and ii) Result 1 is robust when using satisficing success rate as a baseline. Figure 10 shows that the first prediction holds in the data: if anything, the probability of correctly implementing the satisficing selection rule seems to increase with the number of available sources, though not significantly. To address the second prediction I estimate through OLS the following equation:

$$I(bestsource_i) = \alpha + \beta_0 L_i \cdot I(Main)_i + \beta_1 L_i + \beta_2 I(Main)_i + \gamma X_i + \varepsilon_i, \tag{2}$$

where $I(Main)_i$ is equal to 1 if observation i belongs to the *Main* treatment. The main coefficient of interest is β_0, which is the coefficient of the interaction term. Table 2 reports the estimates for different specifications of Equation 2, with the full specification corresponding to the rightmost column. The second prediction concerning the satisficing treatment is confirmed by the results. The negative effect coefficient implies that the marginal loss in performance is larger for the *Main* condition, compared to the *Satisficing* one. This result dispels the concerns that Result 1 is driven by other factors related to the length of the sources list, as opposed to an increasing source selection complexity.

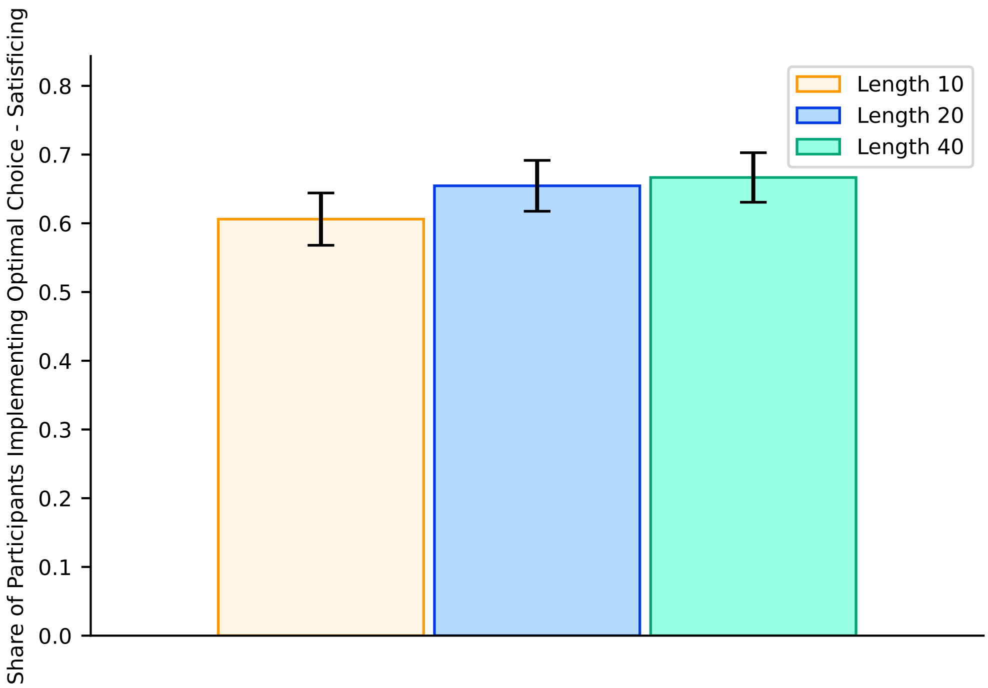


Figure 10: Share of participants correctly implementing the Satisficing rule, by list length. Error bars represent 95% confidence intervals.

	Dependent variable: Probability of Correctly Implementing the Rule			
	(1)	(2)	(3)	(4)
Effect	-0.008***	-0.008***	-0.008***	-0.008***
	(0.001)	(0.001)	(0.001)	(0.001)
1 if in Main Treatment	0.147**	0.147**	0.145**	0.149**
	(0.069)	(0.069)	(0.072)	(0.072)
List Length	0.002*	0.003**	0.004**	0.004***
	(0.001)	(0.001)	(0.001)	(0.001)
Best Source Position		-0.003	-0.003*	-0.003*
		(0.002)	(0.002)	(0.002)
Total Time on Sources			0.001***	0.001***
			(0.000)	(0.000)
Working Memory Proxy	✗	✗	✓	✓
Demographic Controls	✗	✗	✗	✓
Session FE	✗	✗	✗	✓
Priors	✗	✗	✗	✓
Observations	1,732	1,732	1,732	1,732
R^2	0.017	0.019	0.032	0.038

Notes. OLS estimates. Robust standard errors are clustered at the subject level. The dependent variable is an indicator, equal to one if the selected source correctly implements the requested rule. For the *Main* treatment this means selecting the best available source. For *Satisficing*, instead, it means to select the first source in the list with a given precision for both states. $^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$

Table 2

5 Selection Rule Switch

The second result concerns the position of the selected sources in the list, and how this position varies with the number of available sources. Moreover, I present additional evidence that fosters the interpretation of the observed pattern representing a change in the source selection rule, due to increasing computational costs for participants, as the number of available sources increases.

Result 2. ***Selection rule switch:*** *The position in the list (absolute and relative) of the selected source decreases in the number of available sources.*

Figure 11 shows the pattern of the average position of the selected source across different length conditions. It is possible to observe a decrease in the average position, although not a particularly marked one. However, note that this implies that the relative position of the selected source is markedly decreasing across length conditions, as shown in Figure A.1 in the Appendix. Hence, already from this qualitative evidence, it is possible to deduct two aspects of the average source selection rule. First, rational choice is excluded. As the position of the best source is random, the fact that the average selected sources are approximately the fourth in both "Length 20" and "Length 40" conditions, indicates that, on average, participants were not implementing rational choice. This evidence fosters the evidence provided in the previous section. Second, it seems that the average strategy is not to consider a fixed share of the available sources, nor, as will be more clear from the formal analysis and the additional evidence, to consider a fixed amount of sources in each list. Figure A.3, reporting the average share of considered sources by condition, strengthens the point that participants seem to adopt different source selection strategies, depending on the amount of available sources. Indeed, the fact that participants' share of considered sources significantly decreases as the number of sources goes up points towards a switch to a satisficing selection rule.

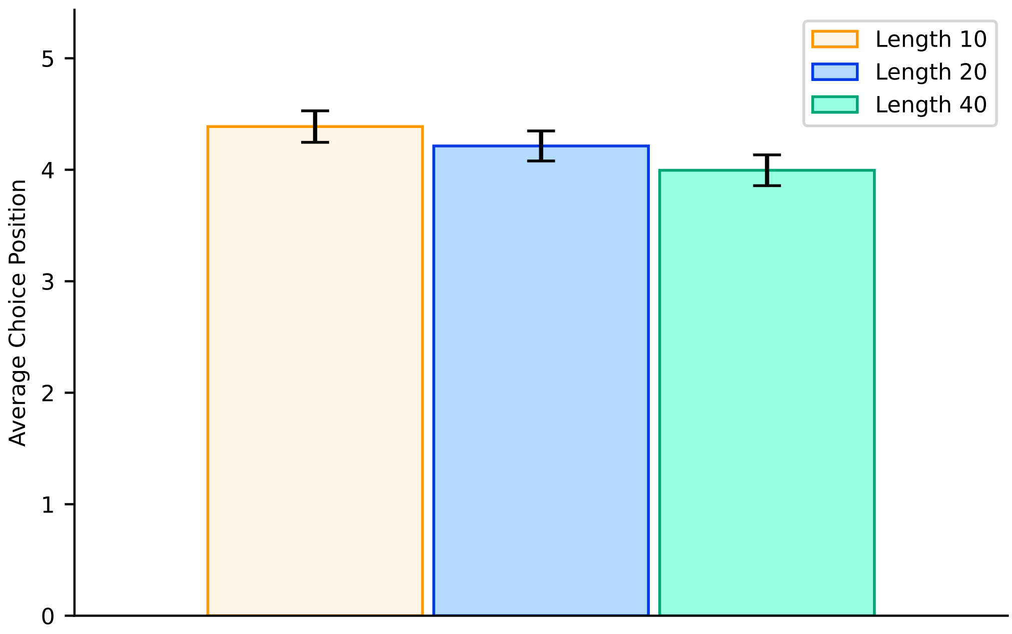


Figure 11: Average position of the selected source in the list. Error bars represent 95% *confidence intervals.*

The formal analysis is carried out similarly to the previous section. I estimate through OLS an equation identical to Equation 1, except that the dependent variable is the position of the selected sources. Table 3 reports the coefficient estimates for different specifications of the linear model, equivalent to the four specifications of Table 1. Interestingly, controlling for the position of the best source and other relevant factors, such as the time spent hovering on information sources, the estimated coefficient of list length is negative and significant. Hence, as the amount of available sources increases, the position of the selected source decreases.

	Dependent variable: Selected Source Position			
	(1)	(2)	(3)	(4)
List Length	-0.012*	-0.026***	-0.027***	-0.028***
	(0.007)	(0.009)	(0.010)	(0.010)
Best Source Position		0.027**	0.025**	0.026**
		(0.012)	(0.012)	(0.012)
Total Time on Sources	✗	✗	✓	✓
Working Memory Proxy	✗	✗	✓	✓
Demographic Controls	✗	✗	✗	✓
Session FE	✗	✗	✗	✓
Priors	✗	✗	✗	✓
Observations	1,237	1,237	1,237	1,237
R^2	0.003	0.008	0.013	0.015

Notes. OLS estimates. Robust standard errors are clustered at the subject level. The dependent variable is the position of the selected source in the list. $^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$

Table 3

6 Belief Updating vs Source Selection Trade-Off

In what follows I discuss some results concerning the interaction between the source selection and the belief updating parts of the task. More specifically, I look into how i) the number of available sources and ii) the source selection performance, formalized through relative source rank, impact the belief updating performance. The latter is defined in relation to the Bayesian benchmark. Additionally, I report evidence concerning the performance in terms of state guess: as the computer actually draws a (hidden) state for each task, it is possible to compare the state (implicitly) guessed by participants, through their probability assessment, and the actual drawn state.

Bayesian Benchmark

Result 3. ***Belief updating performance trade-offs*: *Belief updating performances are decreasing in the number of available sources and in source selection performances.***

Before delving into the analysis it is necessary to define how the key outcome and the relevant variables of interest are constructed. The belief updating performance is defined, using the Bayesian posterior as the benchmark, as follows:

$$\text{belief\_performance}_i = 100 - |P_i(A \mid s) - P(A \mid s)|, \tag{3}$$

with $P_i(A \mid s)$ being the probability attributed by the participant to state A, in choice i, having observed signal s. $P(A \mid s)$ is the Bayesian posterior of state A given signal s. Hence, the closer the guess to the normative benchmark, the higher the performance. Beyond the length condition, the other variable of interest is the source selection performance, constructed using the relative source rank. Given a list of sources and each source rank, as defined in Section 2, the relative source rank is the quantile of that source rank.[16]

Table 4 reports the result from estimating an equation identical to Equation 1, except that the dependent variable is the belief updating performance. The results are consistent with the hypothesis of a trade-off between information source selection and belief updating performances. The coefficients of *List Lenght* and of *Source Quantile Rank* shed light on two different aspects of such trade-off mechanisms. First, the estimated negative coefficient of *List Length* (amount of available sources) reveals a decrease in performances in belief updating tasks following source selection from longer lists. This is consistent with the notion that participants incur higher working memory costs when having access to a larger number of information sources, whatever their selection rule is, and that these costs are carried forward in the related belief updating task. Second, the estimated negative coefficient of *Source Quantile Rank* (source selection performance) shows that ceteris paribus, better performance in the source selection task impacts negatively the related belief updating task. Assuming that, on average, the selection of better sources implies a better-performing source selection rule, then this result supports the view of a trade-off, in terms of cognitive resources, between information source selection and belief updating. Following the theoretical framework, better-performing source selection rules are also more demanding from a working memory perspective, the extreme instance of this being rational choice.

To sum up this first result on belief updating performances, it shows how two different potential sources of working memory depletion in the source selection part of the task, the number of available sources and source selection performances, have a negative impact on performances in the following belief updating task. This is consistent with a model featuring an agent with finite working memory, which needs to be allocated between selecting a source and mapping the information generated from the source and the prior into a posterior belief.

[16] See Section 4 for additional details on how the variable is constructed.

	Dependent variable: Belief Updating Perfomance		
	(1)	(2)	(3)
List Length	-0.105**	-0.116*	-0.154**
	(0.043)	(0.065)	(0.068)
Best Source Position		-0.007	-0.012
		(0.094)	(0.094)
Source Quantile Rank		-4.747*	-5.551**
		(2.431)	(2.449)
Total Time on Sources	✗	✓	✓
Working Memory Proxy	✗	✓	✓
Demographic Controls	✗	✗	✓
Session FE	✗	✗	✓
Priors	✗	✗	✓
Observations	1,237	1,237	1,237
R^2	0.004	0.009	0.020

Notes. OLS estimates. Robust standard errors are clustered at the subject level. is the belief updating performance, constructed as 100 minus the absolute difference between the Bayesian and the reported posteriors. $^{*}p < 0.1,^{**} p < 0.05,^{***} p < 0.01$

Table 4

Unobservable State Guess

A different approach to evaluate participants' belief performance is to compare their guess about the state with the true, unobservable, state. This measure is sensible also in light of the motivating examples of this work, in which a decision-maker needs to form beliefs about an unobservable state to perform some action, the optimality of which depends on the state realization.

In the experimental setting, participants do not provide a direct guess about the state, but an assessment of the probability of each state. Hence, I consider the *implicit* guesses, that is:

$$state\_guess_i = \begin{cases} 1, \text{if } P_i(\theta \mid \theta) \geq 0.5 \\ 0, \text{if } P_i(\theta \mid \theta) < 0.5. \end{cases} \tag{4}$$

Hence, the state is considered correctly guessed if and only if the probability assigned by the participant to state θ, when θ is true, is at least 50%.

Result 4. ***State Guess Performances*****:** *The probability of a correct (implicit) state guess does not vary significantly with the number of available sources.*

Figure 12 reports a preliminary comparison of the share of correct state guesses across different list lengths. The probability of correctly guessing the state seems to vary across the different conditions, but not monotonically: the share of correct guesses is lowest when the available sources are 20. Hence, it seems that, although the quality of information sources increases significantly with length, the probability of correctly guessing the state does not follow the same pattern. Figure A.2, in the Appendix, shows how participants select on average more precise sources when more information sources are available. This can be attributed to the fact that longer lists contain more precise sources. However, the selection of more precise sources does not translate into improved guesses about the unobservable state, as shown in Figure 12.

Table 5 reports the coefficient of OLS estimation of the impact of list length on the probability of a correct state guess. For all three different specifications, the coefficient is very close to 0 and not significant. Hence, there is no evidence of the probability of correctly guessing the state being different across the different length conditions, also controlling for all other relevant factors. This holds despite the quality of the information source for longer lists being systematically higher, as illustrated in Section 3. These result, jointly with Result 3, stresses the idea that the cognitive load caused by a larger amount of available sources can compensate for the advantages brought by better source quality.

	Dependent variable: Probability of Correct State Guess		
	(1)	(2)	(3)
List Length	0.002	0.001	0.001
	(0.001)	(0.002)	(0.002)
Best Source Position		0.002	0.002
		(0.002)	(0.002)
Source Percentile Rank		0.125^{**}	0.117^{**}
		(0.054)	(0.052)
Total Time on Sources	✗	✓	✓
Working Memory Proxy	✗	✓	✓
Demographic Controls	✗	✗	✓
Session FE	✗	✗	✓
Priors	✗	✗	✓
Observations	1,237	1,237	1,237
R^2	0.002	0.010	0.023

Notes. OLS estimates, robust standard errors are clustered at the subject level. The dependent variable is an indicator, equal to one if the implicitly guessed state is correct. A state is considered implicitly guessed if the posterior attributes more than 50% probability to that state. $^{*}p < 0.1,^{**} p < 0.05,^{***} p < 0.01$

Table 5

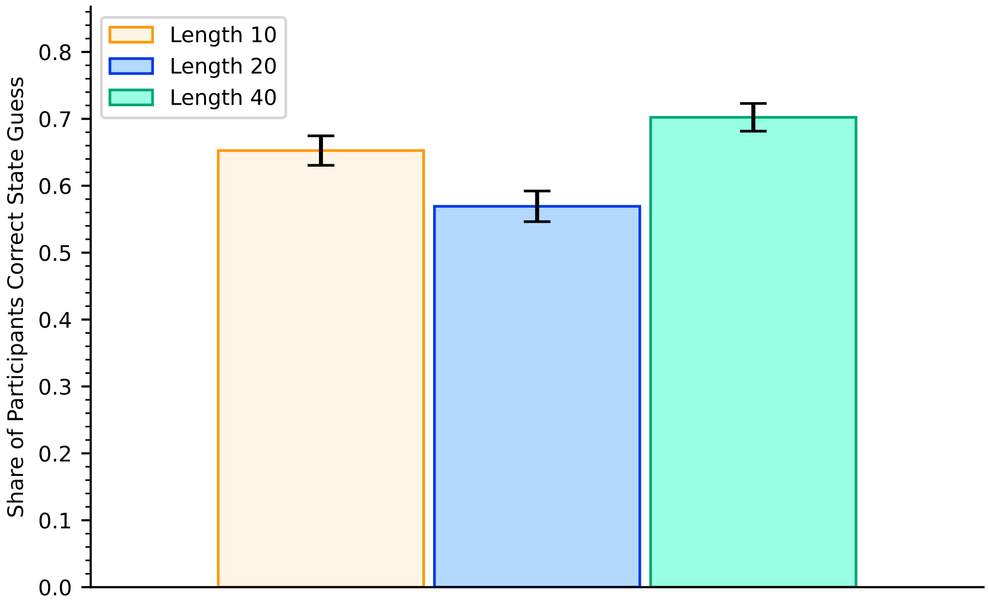


Figure 12: Share of correct state guesses across all choices. A participant is considered guessing a state if she attaches more than 50% *to it. Error bars represent* 95% *confidence intervals.*

7 Discussion and Concluding Remarks

This paper investigates the impact of the number of available information sources on people's ability to select informative sources and make inferences based on the selected sources. The investigation is carried out through an online experiment. The design of the experiment is informed and guided by a theoretical framework based on automata models of decision-making. First, the data show that the probability of selecting the best available source decreases significantly as the number of available sources increases. I propose that this is caused by an increased complexity of implementing rational choice, as the available sources increase in number. Through the *Satisficing* treatment it is possible to exclude that the high number of sources itself confuses participants, instead of the increased complexity of rational choice. Second, I report a trade-off between the source selection and the belief updating performances: ceteris paribus, participants selecting better sources perform worse in the belief updating task.

Considered jointly, the results support a model in which finite working memory is allocated between source selection and belief updating tasks. Also, consistently with the theoretical framework, the results suggest that individuals switch to different source selection rules, as the cognitive load caused by the number of available sources varies. A larger number of sources increases the complexity of the source selection environment, with negative spill-overs on belief updating. Additionally, the results on belief updating and state guess show how the costs associated with more information sources can compensate for the advantage of having access to better sources. Indeed, despite longer lists containing better sources on average, participants' performance in guessing the unobservable state is weakly worse when the number of sources increases. This result has two relevant applied implications. First, information sources seem to be akin to a good generating negative externalities, when available in an overabundant quantity. Second, mechanisms to filter and select information sources play a key role, the importance of which increases with the number of available sources. Indeed, as complexity increases, because of additional available sources, individuals may resort to other means to select information, for instance outsourcing the procedure to an algorithm. This mechanism is not explored in the stylized framework of this paper, but the results point towards the importance of regulating also these alternative source selection procedures not directly controlled by individuals.

References

Benkler, Yochai. 2006. *The Wealth of Networks: How Social Production Transforms Markets and Freedom*. Yale University Press. [4]

Botti, Simona, and Sheena S. Iyengar. 2004. "The psychological pleasure and pain of choosing: When people prefer choosing at the cost of subsequent outcome satisfaction." 87 (3): 312–26. [3]

Caplin, Andrew, and Mark Dean. 2015. "Revealed Preference, Rational Inattention, and Costly Information Acquisition." *American Economic Review* 105 (7): 2183–203. [3]

Caplin, Andrew, Mark Dean, and Daniel Martin. 2011. "Search and Satisficing." *American Economic Review* 101 (7): 2899–922. [2, 3]

Chatterjee, Kalyan, and Hamid Sabourian. 2009. "Game Theory and Strategic Complexity." In. *Encyclopedia of Complexity and Systems Science*. Edited by Robert A. Meyers. New York, NY: Springer New York, 4098–114. [2]

Chauvin, Kayle. 2023. "Euclidean Properties of Bayesian Updating." Working Paper. [3]

Chen, Yu-Chen, Rong-An Shang, and Chen-Yu Kao. 2009. "The effects of information overload on consumers' subjective state towards buying decision in the internet shopping environment." *Electronic Commerce Research and Applications*, (02): 48–58. [4]

Chernev, Alexander. 2003. "When More Is Less and Less Is More: The Role of Ideal Point Availability and Assortment in Consumer Choice." *Journal of Consumer Research* 30 (02): 170–83. [3]

Chernev, Alexander, Ulf Bockenholt, and Joseph Goodman. 2015. "Choice Overload: A Conceptual Review and Meta-Analysis." *Journal of Consumer Psychology* 25 (04): 333–58. [3]

Chernev, Alexander, and Ryan Hamilton. 2009. "Assortment Size and Option Attractiveness in Consumer Choice Among Retailers." *Journal of Marketing Research* 46 (06): 410–20. [3]

Compte, Olivier, and Andrew Postlewaite. 2012. "Belief Formation." Working Paper 12-027. [3]

Conway, Andrew, Michael Kane, Michael Bunting, Zach Hambrick, Oliver Wilhelm, and Randall Engle. 2005. "Working memory span task: A methodological review and user's guide." *Psychonomic bulletin review* 12: 769–86. [11]

Diehl, Kristin. 2005. "When Two Rights Make a Wrong: Searching Too Much in Ordered Environments." *Journal of Marketing Research* 42 (3): 313–22. [3]

Dijksterhuis, Ap, Maarten Bos, Loran Nordgren, and Rick Baaren. 2006. "On Making the Right Choice: The Deliberation-Without-Attention Effect." *Science* 311 (03): 1005–7. [3]

Ehud, Kalai. 1990. "Bounded Rationality and Strategic Complexity in Repeated Games." In *Game Theory and Applications*. Edited by Tatsuro Ichiishi, Abraham Neyman, and Yair Tauman. Economic Theory, Econometrics, and Mathematical Economics. San Diego: Academic Press, 131–57. [2]

Enke, Benjamin, and Thomas Graeber. 2023. "Cognitive Uncertainty." Working Paper. [2]

Enke, Benjamin, and Florian Zimmermann. 2017. "Correlation Neglect in Belief Formation." *Review of Economic Studies* 86 (1): 313–32. [2]

Flaxman, Seth, Sharad Goel, and Justin M. Rao. 2016. "Filter Bubbles, Echo Chambers, and Online News Consumption." *Public Opinion Quarterly* 80 (S1): 298–320. [4]

Gentzkow, Matthew, and Jesse M. Shapiro. 2011. " Ideological Segregation Online and Offline." *Quarterly Journal of Economics* 126 (4): 1799–839. [4]

Gerald, Spindler. 2011. "Behavioural Finance and Investor Protection Regulations." *Journal of Consumer Policy* 34 (09): 315–36. [4]

Golin, Marta, and Alessio Romarri. 2022. "Broadband Internet and Attitudes Towards Migrants: Evidence from Spain." IZA Discussion Papers 15804. Institute of Labor Economics (IZA). [4]

Götte, Lorenz, H. J. Han, and B. T. K. Leung. 2020. "Information Overload and Confirmation Bias." Working Paper. [4]

Greifeneder, Rainer, Benjamin Scheibehenne, and Nina Kleber. 2010. "Less may be more when choosing is difficult: Choice complexity and too much choice." *Acta Psychologica* 133 (1): 45–50. [3]

Haynes, Graeme A. 2009. "Testing the boundaries of the choice overload phenomenon: The effect of number of options and time pressure on decision difficulty and satisfaction." *Psychology & Marketing* 26 (3): 204–12. [3]

Hoch, Stephen J., Eric T. Bradlow, and Brian Wansink. 1999. "The Variety of an Assortment." *Marketing Science* 18 (4): 527–46. [3]

Iyengar, Sheena, and Mark Lepper. 2001. "When Choice is Demotivating: Can One Desire Too Much of a Good Thing?" *Journal of personality and social psychology* 79 (01): 995–1006. [3]

Jacoby, Jacob. 1984. "Perspectives on Information Overload." *Journal of Consumer Research* 10 (4): 432–35. [4]

Jacoby, Jacob, Donald E. Speller, and Carol Kohn Berning. 1974. "Brand Choice Behavior as a Function of Information Load: Replication and Extension." *Journal of Consumer Research* 1 (1): 33–42. [4]

Leung, Benson Tsz Kin. 2020. "Limited cognitive ability and selective information processing." *Games and Economic Behavior* 120: 345–69. [2, 3, 5, 8]

Lleras, Juan Sebastián, Yusufcan Masatlioglu, Daisuke Nakajima, and Erkut Y. Ozbay. 2017. "When more is less: Limited consideration." *Journal of Economic Theory* 170: 70–85. [3]

Messner, Claude, and Michaela Wänke. 2011. "Unconscious information processing reduces information overload and increases product satisfaction." *Journal of Consumer Psychology* 21 (01): 9–13. [4]

Oprea, Ryan. 2020. "What Makes a Rule Complex?" *American Economic Review* 110 (12): 3913–51. [2, 3, 7]

Pariser, Eli. 2011. *The filter bubble : what the Internet is hiding from you.* New York: Penguin Press. [4]

Reutskaja, Elena, and Robin M. Hogarth. 2009. "Satisfaction in choice as a function of the number of alternatives: When "goods satiate"." *Psychology & Marketing* 26 (3): 197–203. [4]

Roetzel, Peter Gordon. 2019. "Information overload in the information age: a review of the literature from business administration, business psychology, and related disciplines with a bibliometric approach and framework development." *Business Research* 12 (2): 479–522. [4]

Salant, Yuval. 2011. "Procedural Analysis of Choice Rules with Applications to Bounded Rationality." *American Economic Review* 101 (2): 724–48. [2, 3, 5–7]

Simon, Herbert A. 1955. "A Behavioral Model of Rational Choice." *Quarterly Journal of Economics* 69 (1): 99–118. [2]

Sunstein, Cass R. 2001. *Echo Chambers: Bush V. Gore, Impeachment, and Beyond.* Princeton University Press. [4]

Wilson, Andrea. 2014. "Bounded Memory and Biases in Information Processing." *Econometrica* 82 (6): 2257–94. [3]

Appendix A Additional Figures

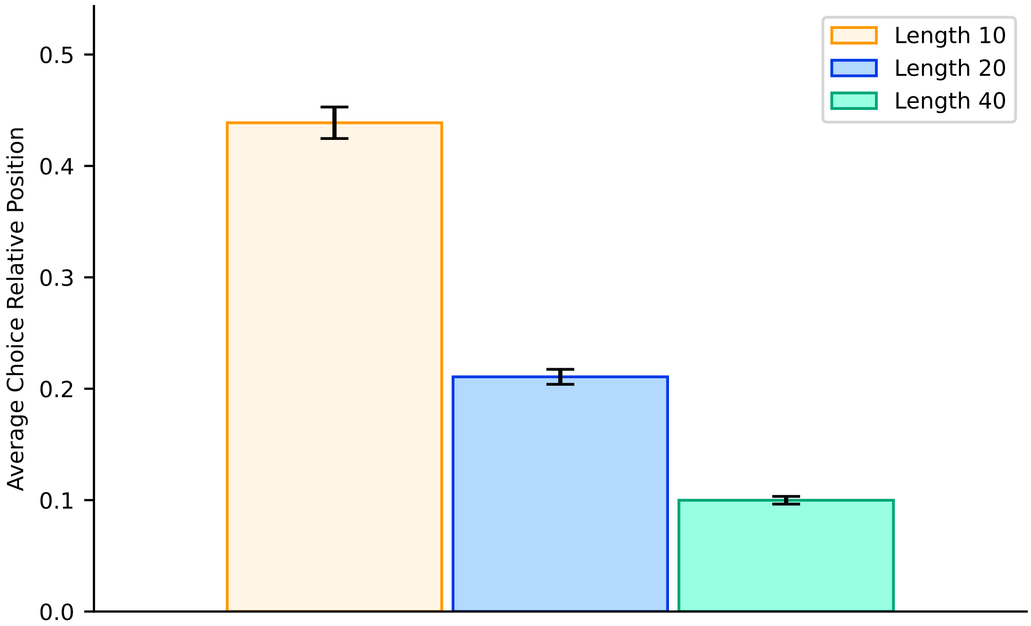


Figure A.1: Average relative position of the selected source, by amount of available sources. The relative position is constructed divided the position of the selected source by the number of available sources. Error bars represent 95% *confidence intervals.*

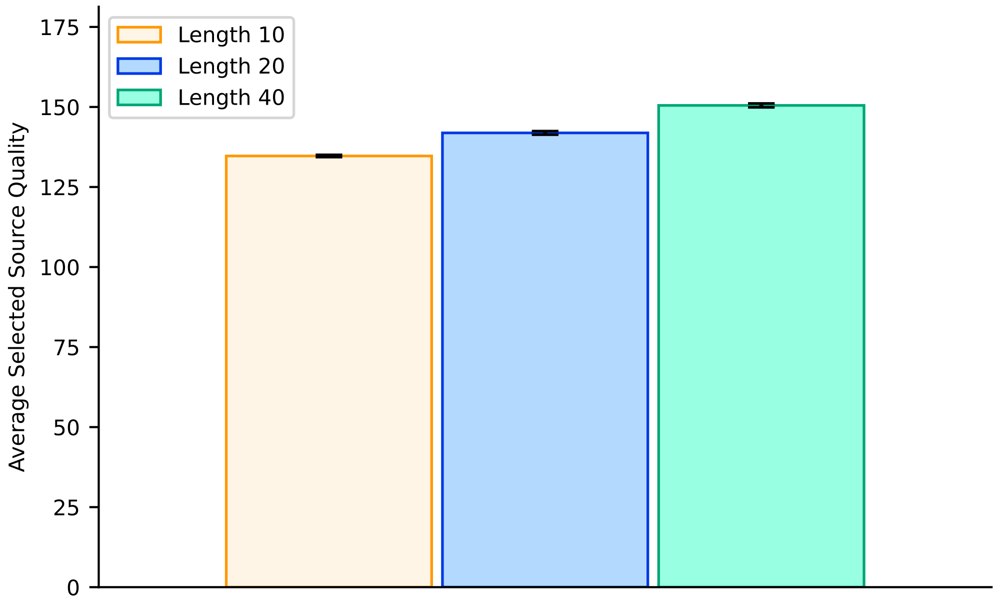


Figure A.2: Average quality of the selected source, by amount of available sources. Source quality is constructed summing the probabilities of truthful reporting for both states. Error bars represent 95% confidence intervals.

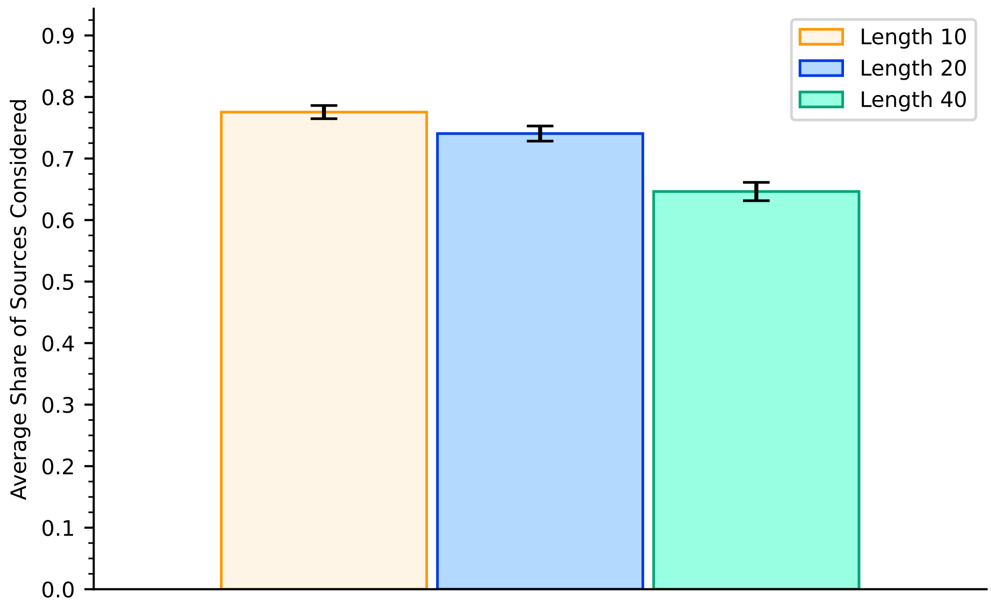


Figure A.3: Average share of sources considered by participants in Main condition. Error bars represent 95% confidence intervals.

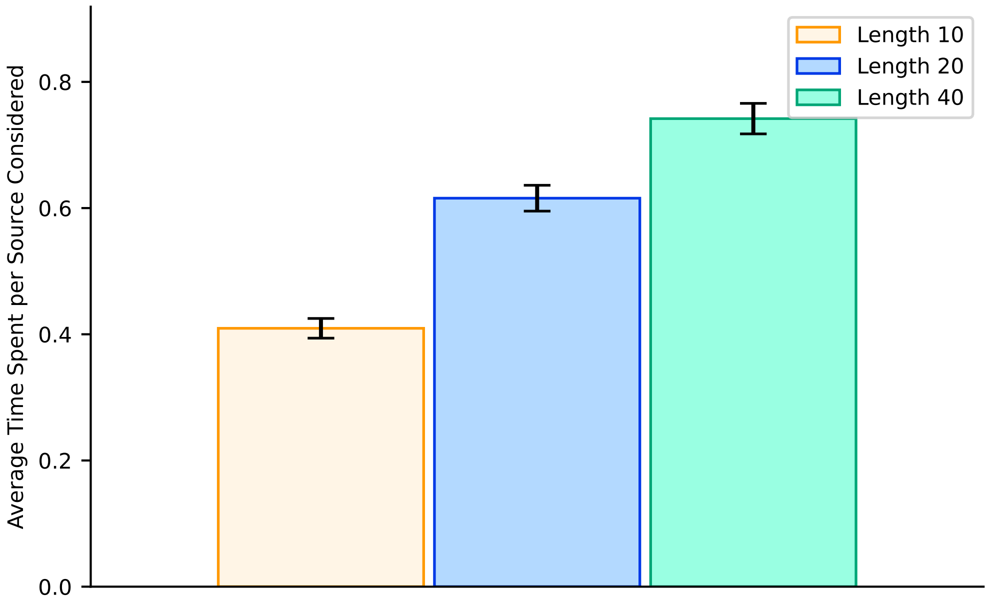


Figure A.4: Average time spent per considered source by participants. Error bars represent 95% confidence intervals.

Appendix B Additional Tables

	Dependent variable: Source Percentile Rank			
	(1)	(2)	(3)	(4)
List Length	-0.001**	-0.001	-0.001*	-0.001**
	(0.000)	(0.001)	(0.001)	(0.001)
Best Source Position		-0.000	-0.001	-0.001
		(0.001)	(0.001)	(0.001)
Total Time on Sources			0.001***	0.001***
			(0.000)	(0.000)
Working Memory Proxy	✗	✗	✓	✓
Demographic Controls	✗	✗	✗	✓
Session FE	✗	✗	✗	✓
Priors	✗	✗	✗	✓
Observations	1,237	1,237	1,237	1,237
R^2	0.002	0.002	0.062	0.075
Note:				*p<0.1; **p<0.05; ***p<0.01

Table B.1

Appendix C Experimental Material

In what follows, I provide screenshots of the instructions and control questions not provided in the main text. *Main* the *Satisficing* treatments do not differ, except for which sources participants were instructed to select. For the *Main* treatments, participants were told: "*In all the following tasks you will have to select* ***the best*** *information source contained in the list.*". For the *Satisficing* treatment, instead, participants were told: "*In all the following tasks you will have to consider the sources in order and select the first information source that satisfies the requirement.*". Also, on the decision screen, a specific precision requirement was indicated: "*Please* ***consider the sources orderly*** *(from first to last) and select* ***the first*** *information source in the list with at least* 57% *precision for both states.*".

C.1 Instructions

Instructions

Please take your time to read the instructions carefully. Your understanding of the instructions will be tested later.

*In this study, you will have to complete **9 similar tasks.***

*Each task is split into **2 parts**:*

1. Select an **information source**
2. Formulate a **probability guess**

An **information source** is a computer that will provide you with information about a **state (A or B)**. This computer is expressed using a table (example below).

Figure C.1: Experimental Instructions 1

What is a state?

A **state** is simply a **draw** made by another computer, which you do not select. You can think of it as something that you **would like to guess**, but that you **can't directly observe** (for example a card draw from a deck).

All the tasks follow the same structure:

1) The first computer (the one you don't pick) draws a state (A or B).

2) You are shown a list of tables. Each table **represents an information source.**

3) You select one of the information sources in the list.

4) After you select an information source, the source will provide a **piece of advice** on what is the **true state**. In other words, it will suggest either A or B.

5) You formulate a guess about the probability of states A and B.

Precision for State A

Source Suggestion

Real State		A	B
	A	70%	30%
	B	31%	69%

Precision for State B

Source Suggestion

Real State		A	B
	A	64%	36%
	B	36%	64%

Figure C.2: Experimental Instructions 2

Understanding Information Sources

Information sources (the computers you select) **may tell the truth** or **may lie**. The tables in the picture above are two examples of two different information sources. These tables **provide you with information about how likely is a source to tell the truth or to lie**.

The first computer may draw state A or state B, but you can't observe it. Let's analyze both cases:

1) Imagine the first computer drew state A. Then the first source would suggest you "A" (**tell the truth**) with 70% probability and "B" (**lie**) with 30%. The second source, instead, would suggest you "A" with 64% probability and "B" with 36% probability. Hence, **in case A was drawn** the **first source would tell the truth** with a larger probability.

2) Now, imagine the first computer drew state B. The first source would suggest you "A" (**lie**) with 31% probability and "B" (**tell the truth**) with 69%. The second source, instead, would suggest you "A" with 64% probability and "B" with 36% probability. Hence, **also in this case, the first source would tell the truth with a larger probability.**

For this reason, in this case, **the first source in the picture is the best**, as it is **more precise for both states**.

In other words, **a source is better then another source** if it is **more precise both for state A and state B.**

Figure C.3: Experimental Instructions 3

C.2 Comprehension Questions

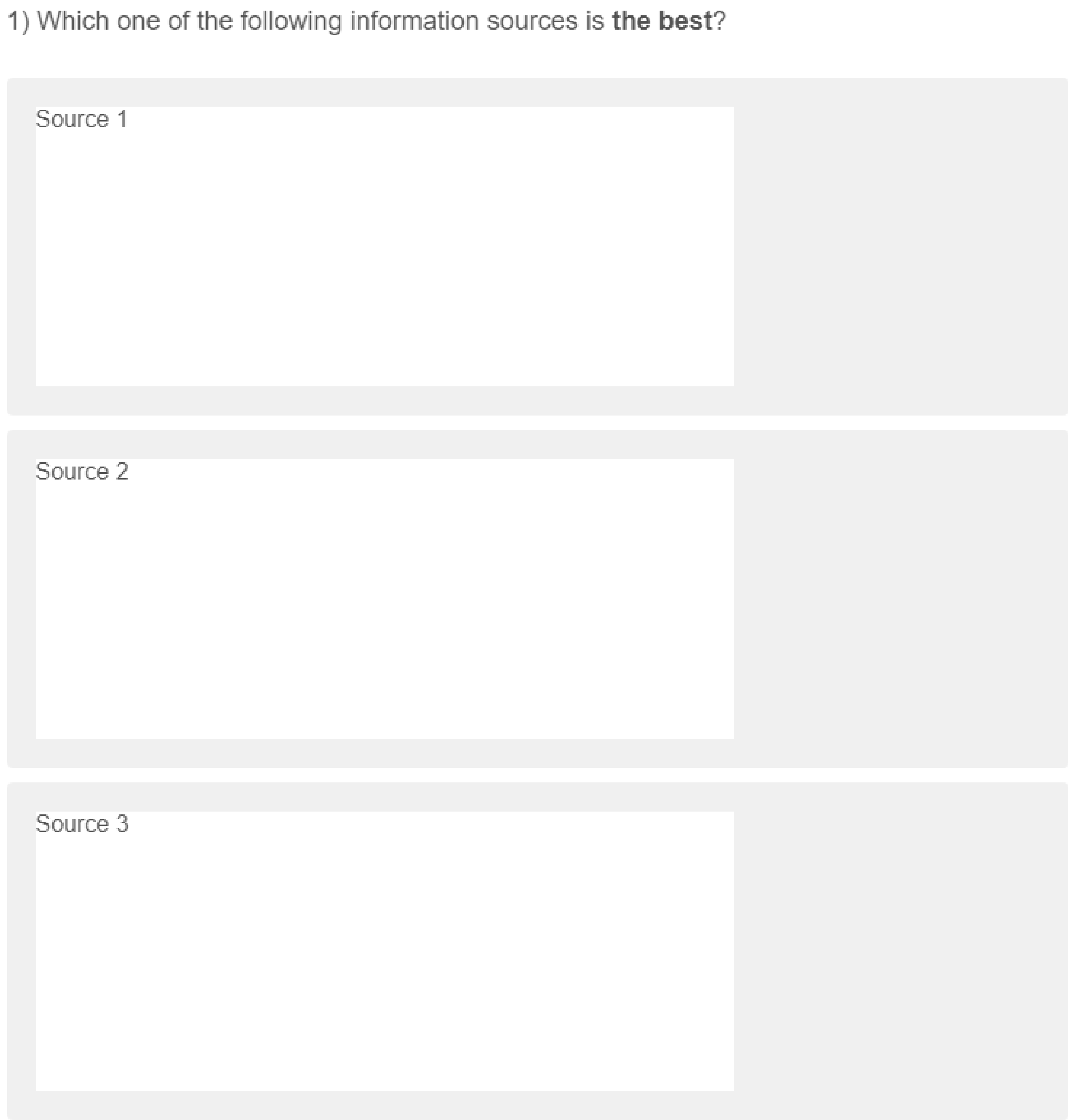


Figure C.4: Comprehension Questions 1

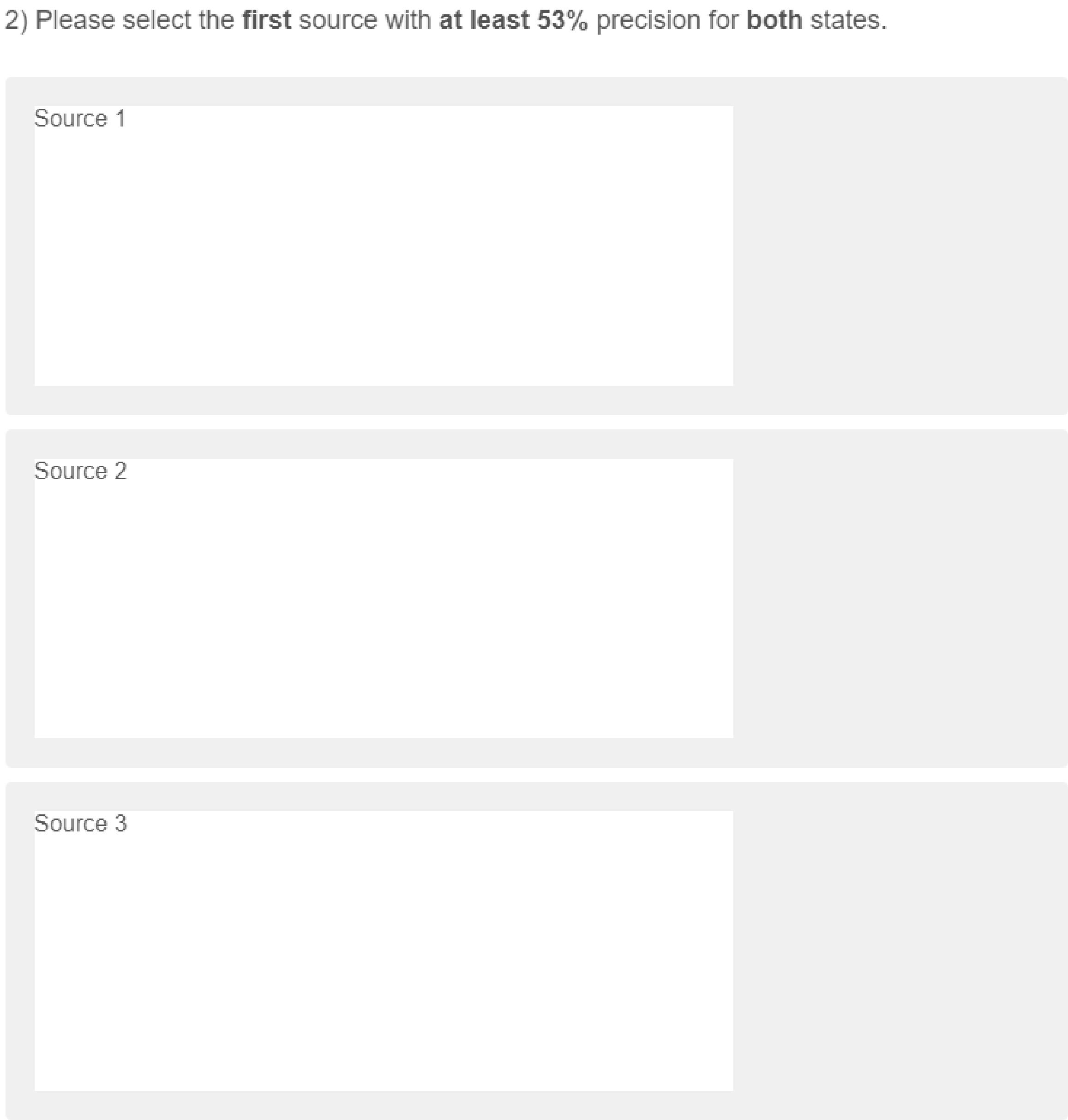
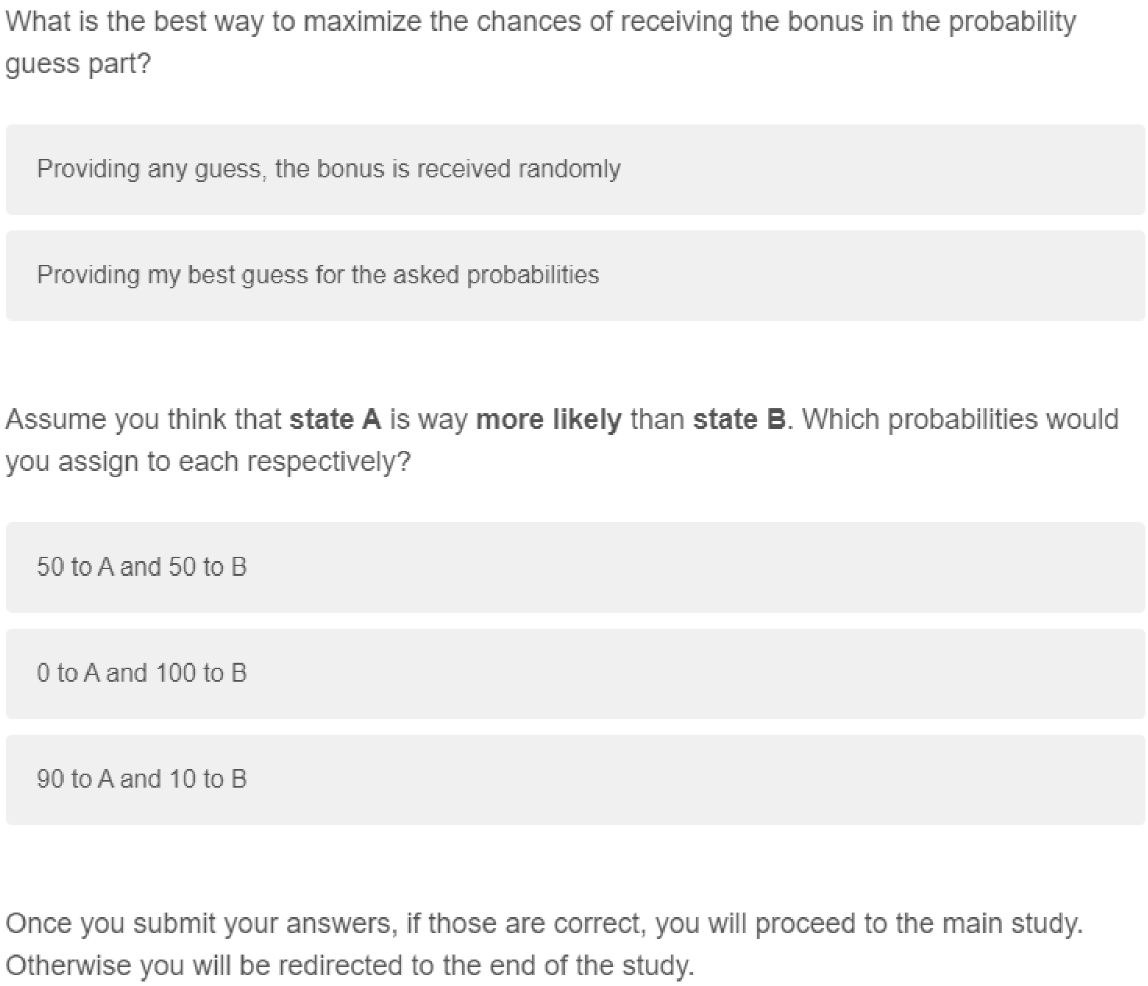


Figure C.5: Comprehension Questions 2



What is the best way to maximize the chances of receiving the bonus in the probability guess part?

- Providing any guess, the bonus is received randomly
- Providing my best guess for the asked probabilities

Assume you think that **state A** is way **more likely** than **state B**. Which probabilities would you assign to each respectively?

- 50 to A and 50 to B
- 0 to A and 100 to B
- 90 to A and 10 to B

Once you submit your answers, if those are correct, you will proceed to the main study. Otherwise you will be redirected to the end of the study.

Figure C.6: Comprehension Questions 3