Inattention in multi-attribute search: an experiment

Bora S. Lancee[a], Imen Bouhlel[b,*], Stephanie Rosenkranz[a]

[a] *Utrecht University School of Economics, Kriekenpitplein 21-22, Utrecht, 3508TC, The Netherlands*
[b] *Université Côte d'Azur, CNRS, GREDEG, 250 Rue Albert Einstein, Valbonne, 06560, France*

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

Preprint submitted to Journal of Economic Behavior and Organization *May 24, 2023*

Highlights

Inattention in multi-attribute search: an experiment

- We study individual multi-attribute search using a representative sample of the UK
- Search behavior is better predicted with inattention
- We replicate the relationship between attribute importance and allocated attention of earlier works
- Communicating the optimal level of search can help converge the participants' search behavior toward that level
- Attention can have a moderating effect on the change of allocated search levels

Inattention in multi-attribute search: an experiment

Abstract

Humans do not have infinite attention. Contrary to what traditional economic models would predict, only a subset of all available alternatives is considered for most decisions one makes in a lifetime. On top of this, only a limited number of attributes of these alternatives are taken into account. We study how choices between alternatives and the associated search behavior among their attributes change when an optimal search strategy is communicated. We do so by implementing a treatment manipulation targeting the amount of search and studying inattention as a moderator. Our experimental study expands the recent work on inattention. We design and implement an online search experiment with a representative UK sample, where we proxy attention levels. We study the relationship between these levels and the amount of search and investigate if we can adjust the amount of search toward a communicated optimal strategy. We find that search behavior is better described with partial attention. Additionally, the estimated levels of inattention depend on the attribute's importance—-the more important the attribute, the higher the level of attention allocated. Secondly, our intervention has a converging effect on search behavior, where so-called "*over-searchers*" diminish their search while "*under-searchers*" increase their search levels towards the communicated optimum. Lastly, we find that estimated inattention levels have a moderating role in the shift of search behavior for the group of over-searchers.

Keywords: multi-attribute search, attention, online-experiment, simulations

1. Introduction

Contrary to what traditional economic models would predict, we do not consider all available alternatives for decision-making. As a result, when choosing, for example, what coffee to buy in the supermarket, most people only think about a few alternative coffee offerings and only consider certain attributes of these alternatives when making their decision. Say, coffee's price and type of grind, rather than the numerous components related to the growing, roasting, and grinding process, that a so-called rational actor should consider. In terms of time investment and cognitive processing costs, it is impossible to incorporate all available information into all decisions one has to make. Thus, in reality, the decision-maker is either fully inattentive to certain informational pieces, e.g., not considering them at all, or partially inattentive, where the decision-maker does consider a piece of information but only partially updates prior information due to lacking attention. A recent growing body of economics studies shows that many discrepancies between observed behavior and traditional model predictions can be explained by the fact that humans have limited attention Caplin et al. (2019); Gabaix (2019); Mackowiak et al. (2021); Civelli et al. (2022).

One important facet of limited attention is that not all attributes (e.g., coffee grind) of all available alternatives (e.g., available coffees) are considered. However, the literature on decision-making has long been focused on a large variety of search problems where decision-making is based on a single attribute of the available alternatives while assuming homogeneity in all other attributes. While single-attribute search has been studied extensively Rapoport and Tversky (1966, 1970); Hey (1987); Seale and Rapoport (1997); Sonnemans (1998); Zwick et al. (2003); Bearden et al. (2006); Schunk and Winter (2009); Eriksson and Strimling (2010); Rieger et al. (2021); Descamps et al. (2022); Heinrich et al. (2022); Bouhlel et al. (2022), multi-attribute search is a more recent addition to the literature. Notable exceptions that explicitly study sequential multi-attribute optimal search strategies include Bearden et al. (2008), Gabaix et al. (2006), and Sanjuro (2017). Even though multi-attribute work like that of Gabaix et al., (2006) is well received, much of the multi-attribute evidence still concerns 3 or fewer attributes Gabaix (2019). With more than 3 attributes, it often quickly becomes intractable to calculate optimal strategies Sanjurjo (2017). However, in real life, many everyday decisions are decisions with many attributes, such as a seemingly simple decision of what coffee to buy. Within multi-attribute search, it is not only the attention allocation between the different alternatives that matters but also the (in)attention to the attributes of each alternative. Besides limited attention, there are also other reasons not to incorporate all available information, such as the protection of own image and valued beliefs Tinghög et al. (2022). Examples here off are avoiding information that might make your past judgment or decisions look questionable Mata et al. (2022), coherence with current judgment Mischkowski et al. (2021), or resolving conflict between material interest and self-perception Momsen and Ohndorf (2022). Yet, for this research we use a neutral setting where valued beliefs and self-image do not play a role, to be able to proxy inattention separate from other information ignorance stimuli.

Every decision situation comes with a choice of attention. Actors face the fundamental trade-off between processing more information to improve decisions and save mental effort Mackowiak et al. (2021). To illustrate, how people deal with inattention to information has important implications for several sub-fields in economics. For example, in macroeconomics, attention could influence expectations about the current state of the economy, and thereby the dynamics of prices and consumption (Sims, 2003; Woodford, 2012). Often, we see a delayed or dampened response to changing economic indicators, such as an increase in inflation or consumer confidence; attention models could account for this delay. For finance, inattention can be used to explain the "home bias" (Van Nieuwerburgh and Veldkamp, 2009); one candidate to explain this bias is a higher degree of attention to familiar, closer to home investment opportunities, even though the expected profits might be similar to those from less familiar opportunities. In labor economics, inattention can be used to explain some facets of searches from both firms and job applicants (Bartoš et al., 2016); research has especially focused on discrimination in the labor market and the extent to which distinct last names receive the attention of employers. Even though the notion of limited attention is nothing new in economics, recently a sharp increase in the incorporation of limited attention in economic modeling has taken place (for a history of economic thought on attention, see Festré and Garrouste, 2015). In addition to different applications, there are also multiple approaches to modeling inattention Manzini and Mariotti (2014); Lieder and Griffiths (2020); Sims (2003); Caplin et al. (2019); Gabaix (2019). To shy away from assumptions of "the *rationality of inattention*" and strict assumptions on cost functions, Gabaix's (2019) Behavioral Inattention framework seems to offer the most general and flexible framework of inattention for our particular research setting. We use this framework to predict the behavioral response to our intervention, and we model the moderating role of attention accordingly.

Our work contributes to the literature on two dimensions. On the one hand, we address the raised gap of single or few-attribute decision-making in the literature on search problems by introducing a multi-attribute setting with many alternatives. The multi-attribute setting, in turn, also allows us to measure the relationship between attribute importance and estimated allocated attention. Given the controlled setup, we can calibrate a proxy of attention levels by comparing participants' behavior to simulated search processes. We proxy attention by looking at the deviation to simulated optimal behavior, where we assume that besides partial attention our subjects are rational given the abstract setting of a neutral choice environment. Deviation from optimal behavior is also used as a proxy measurement in other well-cited papers in economics dealing with attention. e.g., Chetty et al. (2009) or Taubinsky & Rees-Jones (2018). To simplify readability, we refer to attention levels in the remainder of the paper whenever we refer to our estimated levels of attention allocated by our participants.

Another gap we address in this paper is the stability of allocated levels of search and (in)attention. Given the large role of attention in decision-making, the question arises of whether search and attention can be changed or whether it is statically allocated per person and per environment. We address this question by analyzing how final decisions and search behavior change when providing a good example of one facet of the search process. We communicate the optimal stopping behavior for our search environment and study whether the search behavior is adjusted to any greater or lesser degree towards the communicated level of simulated optimal search. The optimal level of search that we communicated is calibrated for the particular setting of our multi-attributes search task and based on an optimal range of allocated attention. We randomly exposed half of our participants to this strategy (201 participants were exposed, 199 were not). Our setting allows us to measure the change in the search order, the stopping rule, and the search depth for information. Finally, we study the moderating role of attention and compare attention levels in the treatment and control groups.

This paper presents three main findings. First, we find that search behavior within our representative sample is better predicted if we include inattention. Our calibrated levels of attention, which allow for partial updating of information, fit better with the data of our participants. We also find a similar relationship to prior meta-research Gabaix (2019), where the assigned levels of attention depend on the attribute importance. The relationship between attribute importance and attention is concave and increasing, where attention levels increase rapidly from no to little importance and slowly from high to full importance. Second, we find that communicating the optimal level of search can help converge the participants' search behavior toward that level. We find that so-called over-searchers, those with a higher allocated level of search than the rational model predicts, decrease their search levels toward the communicated optimum. In addition, we find that the depth of search within each alternative is the key feature affected by our intervention, as predicted by our rational model. On the contrary, we find that under-searchers, who have a lower amount of search than what the rational model would predict, increase their search levels toward the communicated optimum. Finally, for over-searchers, we find that attention moderates the change of allocated search levels. When search levels are lowered, the fitted attention parameters are higher, indicating that when the amount of search is lower, more attention is paid to each information piece, and the incorporation of the new information to update the prior one is more extensive.

This paper is structured as follows: Section 2 discusses the Behavioral Inattention framework of Gabaix (2019) and its applications for multi-attribute search. We will then use this framework to predict the behavioral response to our intervention, with attention as the main moderator. In Section 3, we discuss our experimental design based on the one of Gabaix et al. (2006) and the numerical simulations used to incorporate inattention in our treatment manipulations. Section 4 presents the results for our three main hypotheses, where we 1) measure attention and estimate the relationship between the attribute importance and the allocated attention, 2) estimate the relationship between the amount of search and attention, and 3) present results on the effect of our treatment intervention; the communication of an optimal search strategy. Section 5 presents the robustness checks, where we discuss an alternative measurement of attention, re-estimate our main results with different sub-samples and estimations, and discuss learning behavior over time. Finally, Section 6 presents a discussion of our main results and some concluding remarks.

2. Theoretical Framework

2.1. Gabaix's simple framework

We make use of Gabaix's (2019) simple framework for modeling attention, with some slight departures. In Appendix A.1 we discuss a simplified repetition of the most important features for our research setting. In the framework, an agent's behavior is modeled with a prior value of alternative X and adjustment toward perceived signals with Gaussian noise. We start with the assumption of a true value x, drawn from a Gaussian distribution $\mathcal{N}(x^d, \sigma^2)$, with x^d the default value, in this example, the prior mean μ and variance σ^2. The agent does not know the true value of x and instead gets a signal s, where s is:

$$s = x + \varepsilon, \tag{1}$$

where ε is drawn from an independent distribution $\mathcal{N}(0, \sigma^2)$. An agent takes action a based on the signal they receive. Thus a rational agent takes action $a(s) = \hat{x}(s)$. With action a being the behavioral response and $\hat{x}(s)$ the expected value of x given observed s. The agent starts their estimate at the prior value x^d, in this general case, the prior mean of x^d, and thereafter partially adjusts toward the signal s with a dampening factor. The average action $\hat{a}(x) := \mathbb{E}[a(s)|x]$ therefore is:

$$\hat{a}(x) = mx + (1 - m)x^d, \tag{2}$$

with $m \in [0, 1]$, where m refers to the limited reaction to the signal s. For exceptions and more details, see the chapter on Behavioral Inattention by Gabaix (2019) and Appendix A. In the case of zero noise ($\sigma_\epsilon^2 = 0$), an agent updates perfectly when receiving signal s, and therefore $m = 1$. In the opposite limit, an agent with infinite noise ($\sigma_\epsilon^2 \to \infty$) results in $m = 0$, they do not update their default x^d at all after receiving a signal, and thus their estimate remains x^d.

2.2. The multi-attribute framework

We will look at the information acquisition behavior of the agent in a multi-attribute search task. Similar to most multi-attribute settings we depart from the assumption that the complete search and decision sequence is intractable, and the model will apply a rational search action for t+1, similar to the Directed Cognition model of Gabaix et al. (2006).[1] For this purpose, we follow a similar experimental design to these authors, with some small alterations, where our setting has a number of n alternatives available, with the realistic assumption that the values of the multiple attributes for each alternative decrease in standard deviation from left to right. This feature allows the actor to distinguish the most discriminating attributes, similar to, for example, first observing the price of a product and afterward observing the exact shade of color. More details will follow in the experimental design Section 3. Action a of Equation 2 is thus three-dimensional. An agent can either 1) continue searching the attributes of her/his current alternative, 2) switch to searching the attributes of the next best alternative, or 3) stop searching and choose the current best alternative. Under full attention, the value of each alternative is defined as:

$$A_n^a = \sum_{i}^{j_n} x_i, \tag{3}$$

Where A_n refers to the updated alternative for every action a_i. Where j_n is defined as the number of search actions a_i thus far performed for alternative n. And x_i refers to the i^{th} attribute of the alternative A_n. The cost-benefit trade-off of one extra search action, as compared to stopping the search and choosing an alternative is:

$$w(\theta, \sigma) = \sigma\phi\frac{\theta}{\sigma} - |\theta|\,\Phi(-\frac{|\theta|}{\sigma}) > k, \tag{4}$$

with ϕ defined as the standard normal density function and Φ the associated cumulative distribution function. Here, θ is the estimated value gap between the alternative under consideration and its next best alternative, while σ is the standard deviation from the payoff information associated with the search alternative under consideration. For more details, see Gabaix et al. (2006).[2] The value gap θ_i of action a_i is given as:

$$\theta_i = |A_{1th} - A_{2nd}|, \tag{5}$$

where A_{1st} refers to the current highest alternative and A_{2nd} to the next best alternative as defined by Equation 3; the sum of all information acquired thus far per alternative. The features of this setup have two important implications. First, the incentive to explore the next attributes of the alternative currently under consideration decreases the larger the gap between this alternative and the next best alternative: $w(\theta, \sigma)$ is decreasing in $|\theta|$. Second, the incentive increases with the variability of the information

[1] We will refer to rational actors when we discuss the action for t+1. A complete trajectory of rational choices is not possible, and therefore each t+1 decision is only bounded rational when zooming out. Still for simplicity, we refer to the rational actor when we discuss what is rational for t+1. We make no explicit assumption on the optimization of allocated attention.

[2] In the original paper of Gabaix et al. (2006), θ is defined as x; however, to differentiate from the Equations A.6 - A.9, in Section A.2 we define the value-gap as θ for the remainder of this paper.

that the actor will obtain with the search action: $w(\theta, \sigma)$ is increasing in σ. Under full attention, we expect that 1) as long as $w(\theta, \sigma) > k$, the agent continues the search and 2) does this in the, currently, most beneficial alternative and incorporates the new information fully for the next (possible) search action.

2.3. Hypothesis development

2.3.1. The relationship between attribute importance and attention

The concise deviations of the hypotheses follow in this section. For more details about the deviation, see Appendix A. Agents are often not fully attentive. The value of one of the possible alternatives for multi-attribute search (Equation 3) would in the case of inattention give us:

$$A_n^{s,a} = \sum_i^{j_n} [m_i x_i + (1 - m_i) x_i^d], \tag{6}$$

where $A_n^{s,a}$ is the sum of all subjective values x_i^s of each attribute selected by the search actions j_n. As we can see, if $m_i = 0$, the agent does not update A_n^s at all and sticks with the default value x_i^d, so that the subjective value of the attribute $x_i^s = x_i^d$. When $m_i = 1$ there is full attention to the signal and the agent perceives the true value of $x_i^s = x_i$, thus $A_n^s = A_n$. Whenever $0 < m_i < 1$ they update with partial attention, so that $x_i^s \in (x_i^d, x_i)$.

Following the framework of Gabaix (2019), we expect attention to increase as attribute importance decreases, indicating that $m_1 > m_2$ and $m_2 > m_3$, etc. Additionally, in a meta-analysis, Gabaix found that the decrease of m_i has a concave relationship with attribute importance, where attention increases steeply for small increments of attribute importance close to zero. In contrast, the same shift of attribute importance for more important attributes will only result in a small shift in attention levels. To summarize, we expect that the most important attribute receives an attention level of $m_1 < 1$, where $m_1 > m_2$ and $m_2 > m_3$, and the decrease in attention level is concave. We expect the same to hold in our multi-attribute setting with ten attributes. This leads to our first hypothesis:

> HYPOTHESIS 1: Attributes receive partial attention, where attention increases as attribute importance increases with a concave functional form.

2.3.2. The relationship between the amount of search and attention

If Hypothesis 1 holds, we will find that $m_i < 1$.[3] The subjective value gap that determines action a_i is, therefore:

$$\theta_i^s = |A_{1st}^s - A_{2nd}^s|, \tag{7}$$

[3]For simplicity, we assume that the values of m_1 for all alternatives are identical. However, this assumption is not needed in order for our first and third hypotheses to hold. For the second hypothesis to hold under different m_i for the different alternatives, some additional assumptions on the relationship between the different m_i's are needed. If m_i of A_{1st} is larger than m_i of A_{2nd} the argument still holds. If the opposite holds, the predictions of m_i are also opposite, yet we see no theoretical reasoning for why this would be the case.

where A_1^s refers to the current highest alternative and A_2^s to the next best alternative. Given that $X \sim \mathcal{N}(\mu, \sigma^2)$, within our specific setting $\mu = 0$ (see also Gabaix et al., 2006), it holds in expectation that $\sum_i^{j_n} m_i x_i = 0 \; \forall i \neq 1$ (for all attributes $i \neq 1$ in our setting). However, for $i = 1$ the information is presented without search, just as the most important attributes are portrayed openly and predominantly in almost all real-life situations. A rational actor will thus always start their quest at the highest subjective valued alternative (based on the salient first attributes); therefore, in expectation, $x_1 > 0$ for alternatives A_1^s, A_2^s, A_3^s, and A_4^s.[4,5] Additionally, given the distribution we have $x_i^d = 0$. We can therefore simplify to a value gap in expectation of:

$$\tilde{\theta}^s = m_1(x^s_{1_{A_{1st}}} - x^s_{1_{A_{2nd}}}), \tag{8}$$

with by definition $x^s_{1_{A_{1st}}} > x^s_{1_{A_{2nd}}}$, given that a rational agent always starts her/his search process at the current highest alternative. This has the important implication that the larger the attention parameter m (more attention is allocated), the larger $\tilde{\theta}^s$. Given Equation 4, we thus find the new cost-benefit function:

$$w_s(\tilde{\theta}^s, \sigma) = \sigma\phi\frac{\tilde{\theta}^s}{\sigma} - \left|\tilde{\theta}^s\right| \Phi(-\frac{\left|\tilde{\theta}^s\right|}{\sigma}) > k, \tag{9}$$

given the properties of the benefit function $w_s(\tilde{\theta}^s, \sigma)$, the smaller $\tilde{\theta}^s$, the larger the subjective benefit function $w_s(\tilde{\theta}^s, \sigma)$. This has the important implication that under similar costs, the smaller $\tilde{\theta}^s$, the more search will take place. To summarize, we find in Equation 9 that the smaller the attention parameter m, the greater the perceived benefit $w_s(\tilde{\theta}^s, \sigma)$. Practically this means that since $m_i \left|\tilde{\theta}^s\right| < |\theta|$, and since $w_s(\tilde{\theta}^s, \sigma)$ is decreasing in $\left|\tilde{\theta}^s\right|$, the benefit of the search will be higher under lower levels of attention, and thus more search will take place under the assumption of identical costs k for $\forall m_i \in (0, 1)$. This leads to our second hypothesis[6]:

> HYPOTHESIS 2: There is a negative relationship between the amount of search and the attention parameter m.

2.3.3. The effect of communicating an optimal strategy on search behavior

The interplay between information signals x and attention m determines the amount of information taken into account by the agent and consequently the action a_i. Action

[4]The same holds for the first four subjective alternatives A^s_{1-4}.

[5]Also $x_1 = 0$ in expectation; however, given the initial transparency x_1 for all alternatives, a rational actor will select an alternative A_n with (in expectation) $x_1 > 0$ to start the search process.

[6]However if we assume that cost k of the Behavioral Inattention framework follows the costs function of $k^c = k + c_i^m$ as defined by Gabaix et al., (2019), the increase in search actions might be even higher. This indicates that the cost k^c is a combination of external costs of the search operation k plus mental processing costs c_i^m that depend on the amount of attention m allocated to information signal i. The total costs k^c are larger with full attention, given that the processing costs of $c^{m<1} < c^{m=1}$. This realistic assumption of higher processing costs for higher attention has also been predominantly found in many other works Caplin and Dean (2015); Chabris et al. (2009). Then not only the subjective benefits that *decrease* in m matter for the action a_i, but also the costs k^c that *increase* in m. Given that the direction of predicted behavioral change is the same for both cost and benefits, we do not have to make assumptions about the processing costs c_m^m related to attention, as we can only underestimate the effect.

a of Equation 2 is thus three-dimensional. An agent can: 1) continue searching the attributes of her/his current alternative, 2) switch to searching the attributes of the next best alternative, or 3) stop searching and choose the current best alternative. We use a^{stop} to refer to where the actor stops the search and makes a final decision, while with a_i^{switch} we refer to the i^{th} time the actor switches between the search of alternatives. If a^{stop} is far from the efficient level $a_e(x, m)$, a behavioral intervention informing the agent about an optimal strategy in terms of the number of search operations a_e^{stop} is expected to lead to adjustments of action a^{stop} to a'^{stop} in the direction toward the communicated reference benchmark a_e^{stop}.

A rational agent stops search whenever the benefits $w_s(\tilde{\theta}^s, \sigma)$ are lower than the costs k. This means that those agents that search too much, e.g., the over-searchers are expected to overestimate the benefits $w_s(\tilde{\theta}^s, \sigma)$ of search action, while those agents that search too little (the under-searchers) underestimate the benefits $w_s(\tilde{\theta}^s, \sigma)$.[7] We, therefore, expect that an over-searcher will (partially) adjust a^{stop} toward the proposed benchmark behavior a_e^{stop} such that the adjusted stopping behavior is somewhere between the original and efficient level; $a_e^{stop} \leq a'^{stop} < a^{stop}$, while the reverse holds true for under-searchers.

Going back to the development of Hypothesis 2, we see that under a higher m_i, an agent is expected to update θ more objectively and thus switch within the alternatives more frequently when higher attention m_i is allocated. Intuitively, this means that those who diminish their search will switch more frequently, thus searching less deeply within alternatives. At the same time, the reverse is expected for those who increase their search toward the optimal level. Therefore, we expect that communication of an optimal number of search operations for those who decrease their search $a_e^{stop} < a^{stop}$ will lead to a relatively higher ratio of a_i^{switch}, and the reverse holds for those that increase their search amount.

The subjective value gap determines the benefit function $w_s(\tilde{\theta}^s, \sigma)$. It indicates that the subjective value gap is the difference between the current highest and second-best alternative. After every search operation $a_i(x, m)$ the value gap is (subjectively) updated. If Hypothesis 2 holds and the amount of search and the attention parameter m are negatively correlated, the reverse argument is also expected to hold: given that the objective θ remains constant if a change occurs from action a^{stop} to action a'^{stop}, an optimizing agent is expected to adjust m_i as a result of a change in stopping the behavior. The above leads to our two-part third hypothesis:

> HYPOTHESIS 3.1: The communication of an optimal strategy a_e^{stop} will have a convergent effect on the amount of search toward this level, and switching behavior will be adjusted accordingly.

[7]Another possibility would be the under- (or over-) estimation of the cost k; however, given that the cost is constant over the entire task, while the benefits are unique for each search action we deem this less likely. If this mechanism is determinant for some agents, the predictions in terms of search behavior adjustments stay identical.

HYPOTHESIS 3.2: Attention parameter m is a moderator for the change in search behavior toward a_e^{stop}.

3. Experimental design

To get more insight into the (multi-attribute) search process of people, we used the information search task referred to by Gabaix et al. (2006) as the "N complex Goods" game. We chose this experimental setup since it provides an exceptionally rich, multiple alternatives (eight), multiple attributes (ten), search experiment that tracks both final decisions and the complete search order during the decision-making process. Like many real-life examples, the high dimensionality of the problems makes calculating optimal search strategies analytically and numerically intractable ((Gabaix et al., 2006, p. 1055)). The setup allows us to test the behavior of the search order, the switching behavior, and the stopping rule. We communicated the optimal strategy a_e^{stop} for half of the participants. The experimental design of Gabaix et al. (2006) and the intervention design are discussed below in detail. The experiment was programmed in Python using the oTree interface Chen et al. (2016).[8]

3.1. Experimental details

The N complex Goods game measures information search in a multi-attribute setting. In each round, subjects face eight alternatives they can choose from; the sum of all ten attributes determines the value of each alternative. Nine of the attributes are unknown to the subject, and the first attribute of each alternative is always displayed at the beginning of each round. You can find an example of one randomly generated round in Figure **??** in the **??**. Information can be gathered by clicking on one of the boxes with the mouse. After the initial opening, the information remains visible until the next round. The goal of the task is simple: maximize the total number of points over all (20) rounds, where the total number of points in one round corresponds to the value of the chosen alternative (i.e., the sum of all its ten attributes), minus the total costs devoted to opening boxes. All attributes are randomly generated, independently normally distributed, with zero mean, and linearly declining in variance, from attribute 1 (σ^2) to attribute 10 (0.1 σ^2). σ was constant throughout the game and set to 20. This gave 95% of values between -40 and 40, making it easier for participants to get used to the environment.[9] Attributes could be searched in any order, and each alternative could be chosen at any time. The cost of opening one box equals 1 point. Points were cumulative over all rounds played. After each round, a feedback screen appeared, displaying the total number of points earned thus far and the number of points earned in the last round, separated by the number of points earned by the last decision and the points spent in the opening boxes.

[8]This experiment was pre-registered before collecting any data at `https://aspredicted.org/856_WXK`. Note only 2 out of 3 treatments are presented in this paper.

[9]The intuitive explanation regarding the variance in the participants' instructions can be found in Figure **??**, the mathematical in Figure **??**, and a screenshot of a partially uncovered round can be found in Figure **??** in the **??**.

3.1.1. Experimental intervention

Half of our participants were randomly assigned to our experimental intervention (201 versus 199 not).[10] The experimental intervention consisted of the communication of a number of boxes opened following a computer-simulated optimal strategy a_e^{stop}, which led to the highest payoff on average. Participants were informed that this was the optimal number of search actions that led *on average* to the highest payoffs, as calculated by the computer over many rounds. The communicated strategy was stable throughout the game and was already present in the practice rounds.

We simulated the optimal strategy using 100,000 replications over randomly generated box values, based on a cost of opening boxes equal to 1, initial attention to the first covered attribute between 0 and 1 (using cumulative steps of 0.1),[11] and an additional decrease of attention per attribute of 0.1. We chose to communicate the number of boxes that resulted in the highest payoff. Given that previous research showed that participants often search too much in a similar setting Gabaix et al. (2006), we decided to exclude the upper bound of 14 boxes. The optimal amounts of search per attention level can be found in Table 1 and ranged from opening 6 to 11 boxes, depending on the attention levels allocated. This optimal strategy was communicated at the beginning of each round and remained stable across the entire time span of the game. More details of the simulation results of optimal search under different levels of attention can be found in Table B.1 of Appendix B.

Initial *mi*	N boxes	Payoff round
0.0	0	28
0.1	65	-3
0.2	51	11
0.3	37	24
0.4	26	35
0.5	19	39
0.6	14	41
0.7	**11**	**43**
0.8	**9**	**42**
0.9	**7**	**41**
1.0	**6**	**41**

Table 1: Simulations of search behavior

3.1.2. Experimental flow

The experiment was run online using the subject pool provided by Prolific, an academic platform dedicated to data collection. After reading the instructions for the experimental task, participants had to answer four comprehension questions. Participants were given a second try if one of the comprehension questions was answered

[10]The chosen number of participants followed a power calculation as can be found in the preregistration on AsPredicted at https://aspredicted.org/856_WXK.

[11]Note that for initial attention of 0, the optimal strategy recommends not to open any box, given that with zero attention no information will be incorporated for decision making.

erroneously on the first try. Only after correctly answering the comprehension questions could participants proceed with three practice rounds of the game. Screenshots of the experiment and a transcription of the instructions can be found in the **??**. Once participants had played all 20 rounds, they were directed to a short exit survey.

3.2. Participants

3.2.1. General characteristics

We recruited a total of 400 participants for the experiment from Prolific's subject pool, randomly divided into control and treatment groups. We used a representative sample of the UK population, with demographics matching the demographic composition of the UK in age, ethnicity, and gender, following the UK Office of National Statistics.[12] Although Prolific competes with other online platforms, research shows that Prolific respondents produced high-quality data and are far more diverse than those from other platforms Peer et al. (2017); Eyal et al. (2021). An overview of the general characteristics of the sample can be found in Table 2. Our participants were on average 45 years old. Moreover, we see a high diversity in both the level of education, with a considerable part of the sample in the highest (PhD/Masters) and lowest (High school degree) levels, and occupation (with 54% being employed). Only seven participants (1.75%) had previously participated in a similar (decision) experiment.

3.2.2. The time investment and payoffs

Participants took, on average, 26 minutes to complete the study. Following Prolific payment rules, participants were entitled to a minimum income based on the platforms' considerations of a fair hourly wage. Given our time estimates, we guaranteed payment of a £1.84 show-up fee, equivalent to £5.00 per hour. On top of the guaranteed payments, to incentivize behavior, participants could receive a bonus between 0 and up to £2.20 depending on the total number of points they made, increasing the estimated payoff to £11 per hour. The average bonus was £1.14, equivalent to a 62% increase in guaranteed income. The average final payment (show-up fee + bonus) was £2.98, which is equivalent to £6.90 per hour.

To ensure that participants fully understood the game setting, emphasis was placed on the instructions and comprehension. Before starting, participants were informed about the comprehension requirement. We can see in Table 3 that, on average, participants spend 11 minutes reading the instructions and answering the comprehension questions. Appendix B.2 reports the summary statistics of the comprehension questionnaire. The average number of correct answers was not significantly different ($p = 0.77$) between the two groups. During the practice rounds, the time per decision set was a little over 1 minute, while during the game, this decreased to a little under 30 seconds.

[12] https://webarchive.nationalarchives.gov.uk/

Variable	**Categories**	**Nº**	**Percent**
Gender	Female	201	50.2%
	Male	198	49.5%
	Other	1	0.2%
Age	18-24	42	10.5%
	25-49	191	47.8%
	50-64	113	28.2%
	Over 64	54	13.5%
Occupation	Employed	217	54.2%
	Self-employed	46	11.5%
	Student	35	8.8%
	Unemployed	36	9%
	Other	66	16.5%
Education	High-school	90	22.5%
	Vocational training	49	12.2%
	Bachelor	167	41.8%
	Masters	71	17.8%
	PhD	17	4.2%
	Other	6	1.5%
Income	0-1500 GBP	156	39%
	1501-3000 GBP	156	39%
	3001-5000 GBP	58	14.5%
	$\geq$ 5000 GBP	10	2.5%
	I don't know	20	5%
Ethnicity simplified	Asian	36	9%
	Black	27	6.75%
	Mixed	18	4.5%
	White	285	71.25%
	Other	15	3.75%
	Missing information	1	0.25%
	Consent revoked	18	4.5%
Total		**400**	

Table 2: General characteristics of sample

Statistic	Mean	Std. Dev.	Median
Total time spent on the experiment	25.96	11.78	23.72
Time spent on reading the instructions	5.14	4.69	4.13
Time spent on the comprehension questionnaire	6.00	3.85	4.97
Time spent on practice rounds	3.20	2.91	2.42
Time spent on the game	8.47	5.18	7.2
Time spent on the demographics	3.20	2.41	2.43

Table 3: Summary statistics of the time spent on the different parts on the experiment

4. Results

4.1. Measurement of attention

There are several ways to measure attention. The measurement we will follow is the deviation from optimal actions (Gabaix (2019)). When assuming fully rational decision-makers, deviations from optimal actions are considered to be due to partial attention. Other ways include eliciting beliefs (e.g., surveys) or physical measures (e.g., eye tracing and time of actions). In addition to measuring the deviations from optimal actions, we will discuss an alternative attention measurement based on decision time in Section 5. To measure the participant's actions, we use mouse tracing, programmed in oTree (Chen et al., 2016) to easily fit in the wider environment of our experiment. The tracing is similar to "Mouselab" (Payne et al., 1993; Camerer et al., 1993; Costa-Gomes et al., 2001) where we keep track of the order, time, and duration of subjects' box-opening actions.[13] One of the benefits of mouse versus eye tracing is the ability to observe final decisions and intermediate steps, revealing the attention that plays a role in cognitive processes. On the other hand, eye tracing has the advantage of capturing the gaze of where attention is allocated Brunyé et al. (2019), but it fails to reveal the intermediate steps resulting from cognitive attention processes or the recovery of errors. Mouse tracing allows for measuring the acquisition of information and what the information subjectively means to the participants.

4.1.1. Search-rules of rational search

Our setting allows us to measure the deviation from optimal actions within the space of $t+1$ actions.[14] The following search rules describe rational search behavior in our multi-attribute setting, and deviations hereof will be used to measure attention as described in DellaVigna (2009). The N complex Goods game was selected because for each $t+1$ action, the benefit function $w(\theta, \sigma)$ provides the possibility to objectively determine the best search action. Additionally, a clear distinction between the most important and least important information is made in terms of attribute importance, as described in more detail in Sections 2.2 and 3.1. Combined with the design feature that all attributes are randomly generated with a zero mean and a normal distribution, subjects can determine the rational search strategy for $t+1$. From this follows that rational, fully attentive, sequential search requires the two following search rules to hold:

> SEARCH RULE 1: Subjects will start the search process with the alternative with the highest (in expectation) value. Then, subjects will always search for (one of) the alternative(s) having the highest benefits $w(\theta, \sigma)$.
>
> SEARCH RULE 2: Subjects will always explore the most informative attribute of their current alternative.

[13]We choose to use a oTree tracing interface similar to "MouselabWEB" as we consider its data sufficient for our analysis and in addition easier to merge with the other oTree interface.

[14]The multi-attribute setting does not allow for a rational search strategy over the entire search path, but similar to Gabaix et al., (2006) we assume partially myopic optimization, where actors look at the optimal step for $t+1$.

Besides our interest in specifying the rationality of selecting information, we also wanted to test whether our subject follows a rational cost-benefit analysis for the stopping behavior a^{stop} of search. For this to hold, on top of the first two rational search rules, a rational stopping rule has to hold: the subject keeps searching until the (expected) benefit of the best search action $w(\theta, \sigma)$ (equals or) drops below the costs k. We incorporate our two search rules as well as the stopping rule into a total score that measures how often the subject: 1) opened the optimal box, i.e., respecting simultaneously search rules 1 and 2; and 2) rationally stopped and selected (one of) the alternative(s) with the highest cumulative revealed value.[15] The scores were computed per subject per round. Rounds where the subjects did not operate any search action (i.e., click) were not considered.

Table 4 shows the summary statistics of the scores obtained, pooled at the subject level. Subjects who systematically did not operate any search action (7 subjects in total out of the whole sample) are not considered. The optimal search action is obtained by assuming $m_i = 1$, for all subjects, given that this would indicate full attention and rationality. For each round, the more the subject's search actions correspond to those suggested by a particular search rule, the closer the subject's score for that search rule is to 1. A score of 0.285 means that, on average, subjects follow both search rules and the stopping rule simultaneously 28.5% of the time. To test our first hypothesis, we compare the search processes of our participants with simulated search processes.

Statistic	N	Mean	St. Dev.	Median
Total score	393	0.285	0.256	0.250
Score Search rule 1	393	0.404	0.289	0.333
Score Search rule 2	393	0.664	0.408	1.000

Table 4: Summary Statistics of the scores

Additionally, for all our three hypotheses, we look at the total sample and additionally split our sample into two types of actors: over-searchers and under-searchers. We characterize both types as searchers who, given their attention levels, search either too much (over-searchers) or too little (under-searchers). Intuitively this means that these two types over or underestimate either the benefits or the costs. To categorize our subjects, based on the score of search-rule 1, we calibrate the attention parameters at the subject level, allowing for attention to be $m \neq 1$.[16] Using these calibrated parameters, we then simulate the optimal strategy for each of the experimental series. As a second step, we then define someone as an over-searcher (under-searcher) at the round level if the total number of boxes they explored is higher (lower) than the total number of boxes that would be the rational optimal for that particular series of values.[17] It happened

[15]We do not correct the scores for the odds of executing the right action; however the more search actions already have been taken, the higher the odds of choosing the right action. Given that we expect a negative relationship between the amount of search and attention, we can thus only underestimate the effect by not correcting for the odds.

[16]We use the score of search-rule 1 as it is the one that captures most directly the processing the benefits $w(\theta, \sigma)$. For more details about the simulations, see Appendix B.3.

[17]Thus, even though the optimal number of boxes in some series under full attention might be > 6,

8.725% of the time that the exact optimal number of boxes was explored, in which case the subject was considered neither an over-searcher nor an under-searcher in that round. At the subject level, someone is an over-searcher (under-searcher) if they are an over-searcher (under-searcher) more (less) than 50% of the time. Out of the 393 subjects, 18 were not categorized, as they qualified equally for both categories over the 20 rounds. We found that 86.5% of our sample had a stable type throughout the entire task.[18]

4.2. Hypothesis 1: Attention and decision-making

Hypothesis 1 stated that attributes would receive (partial) attention, thus predicting attention levels that are below 100%. Additionally, it predicted that as attribute importance increased, attention would do so too. Finally, the functional form between the two is expected to be concave. Table 5 shows the calibrated parameters of the (in)attention estimations. As described in Section 4.1.1, attention is estimated based on the deviation from the rational search rules. For our first hypothesis, we are interested in deviations from Search-rule 1, the selection of the most beneficial (in expectation) search action, given that the benefit $w(\theta, \sigma)$ is the feature where attention plays the largest role, and subjects will perceive $w(\theta, \sigma)^s$ if attention is not full. m_1 is the initial attention level for the first attribute at the start of the search process, where $m_1 = 1$ indicates that the actor has full attention, thus rationally following (the first step of) Search rule 1, while $m_1 = 0$ indicates zero attention (or a fully inattentive actor). We find that a high level of attention is allocated to the first, most informative information attribute. Table 5 reports an estimated level of $m_1 = 0.947$. This attention level is higher than earlier work Gabaix (2019), where the highest estimates lie around $m_1 = .80$. Although our estimate of m_1 is higher compared to earlier work, we still find m_1 to be significantly different from 1, indicating an initial inattention level > 0.

The second variable of interest is m_2, which estimates the decline in attention levels as the attribute importance decreases, where $m_2 = 0$ would indicate no change in attention levels, and $m_1 = m_2$ would indicate a decline where the next attribute gets zero attention. Much in line with earlier estimates of (in)attention allocation, we find that when the attribute importance decreases, the level of attention also decreases, as seen by the positive estimate of attention allocated $m_2 = .127$. This estimate is significantly different from zero. We find that when the variance decreases by 10%, the attention level decreases by 12.7 percentage points.

The third variable we report concerns the functional form between the initial attention levels and the decline in attention as attribute importance decreases. We measure the functional form by calibrating the power, where 1 would indicate a linear decrease of m_2, 0.5 a convex relationship where attention m_2 first changes slowly as attribute importance

in another series it might be < 6, and the definition of under or over-searcher depends on the optimal strategy for that particular choice set. The optimal attention levels are fitted to determine the optimal amount of search; this is compared to observed behavior to determine the types.

[18] We compare the first five rounds versus the remaining rounds, taking a 10% "neutral" margin around the 50% level in order to have the more "border" types, e.g., those that are situated precisely or close to the limit, not to be categorized as switching types by being just above or just below the 50% level. If we implement a zero percent margin, we find that still, 63 % of our participants have a consistent type. Other magnitudes for the margin and the proportion of switchers can be found in Appendix B.4.

increases, and 2 a concave relationship where m_2 responds quickly to the first changes of attribute importance and then flattens later. In line with earlier estimates of attention Gabaix (2019), we find that attention and attribute importance does not increase linearly. We, also observe a concave relationship between the increase in attribute importance and attention levels, indicating that the distinction from zero to little importance has a larger effect on attention levels than a similar size increase in high importance levels.[19]

Profile	Variable	N	Mean	Std. Dev.	Median	Min	Max	p-value
All	m_1	393	0.947	0.049	0.96	0.79	1	<0.001 [a]
	m_2	393	0.127	0.027	0.13	0.015	0.2	<0.001 [b]
	power	393	1.13	0.156	1.125	0.625	1.636	<0.001 [a]
Oversearchers	m_1	208	0.933	0.052	0.94	0.79	1	<0.001 [a]
	m_2	208	0.118	0.027	0.12	0.015	0.17	<0.001 [b]
	power	208	1.129	0.138	1.128	0.775	1.526	<0.001 [a]
Undersearchers	m_1	167	0.966	0.039	0.975	0.8	1	<0.001 [a]
	m_2	167	0.137	0.023	0.140	0.06	0.2	<0.001 [b]
	power	167	1.136	0.179	1.125	0.625	1.636	<0.001 [a]

[a]: One Sample t-test, null hypothesis: mean is 1; [b]: One Sample t-test, null hypothesis: mean is 0

Table 5: H1 estimates of m_1, m_2 and the functional form

We find a similar relationship between m_1, m_2 and the power for both under-searchers and over-searchers if we split our sample into two types. We find that the initial attention is significantly higher for under-searchers (t = -7.000, p < 0.000), which gives some indication that there is a negative relationship between the amount of search and the level of attention; we will test this in more detail in Section 4.3. The attention decrease in the next steps is higher for under-searchers (t = -7.460, p < 0.000), compensating for the greater initial attention in the search process. Finally, for both sub-samples, we find a similar concavity between the attention allocated and the relative importance of an attribute (t = -0.422, p-value = 0.674). As can be seen in Figure 1, especially for relatively less important attributes, we observe a difference between the two types, where under-searchers often do not pay any or only marginal attention to relatively unimportant attributes.

[19]Note: Table 5 reports the change in attention as attribute importance *decreases*, while Figure 1 reports the change in attention as attribute importance *increases* to match the figures in Gabaix (2019).

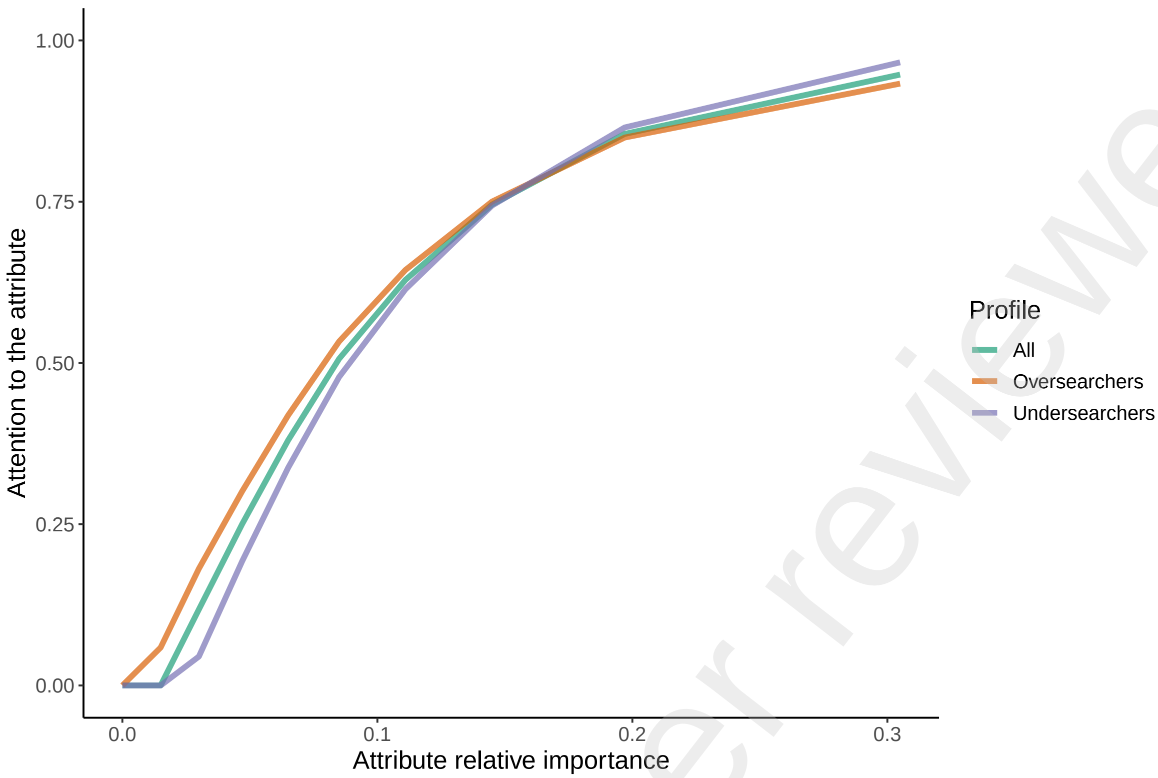

Figure 1: Relationship attention and relative attribute importance

4.2.1. Summary of Hypothesis 1

As shown in Figure 1, search behavior is better predicted when inattention is assumed to be present. Additionally, we can see that as the relative importance of the attribute increases,[20] the allocated attention also increases. Finally, we can observe a concave relationship between the attribute's importance and the attention allocated. All three findings hold for both over-searchers and under-searchers, where under-searchers allocate approximately zero attention (no consideration) for the least important attribute ($x_i > 9$) and have close to full attention for the most important attribute. On the other hand, over-searchers consider more information pieces (attention > 0); however, they more frequently allocate incomplete attention levels for the higher importance attributes by spreading their attention over the multiple pieces of information.

4.3. Hypothesis 2: The relationship between the amount of search and attention

Hypothesis 2 stated a negative relationship between the amount of search and the level of attention. As stated in Section 2.2, the smaller the attention parameter, the larger the perceived benefit of one more search action. Thus, the smaller the attention of an actor, the more search will take place. We look at the total score and the two search rules separately as proxies to estimate the relationship between allocated attention and the total amount of search.[21] First, and most importantly, we take the total score, which

[20]Given that our attributes are randomly generated, we estimated the relative importance by calculating the integral of the probability density functions for the different standard deviations.

[21]To compute the scores, we assume the same values for the attention parameters m_1 $(= 1)$, m_2 $(= 0)$, and *power* $(= 1)$ for all the participants in order to observe the pure effect of the number of

measures whether an actor followed both of the two search rules and the stopping rule consistently for each search action they make.

If we look at Table 6, showing the estimations of linear mixed-effects models at the individual level, we see that the total number of search actions (e.g., the total number of boxes opened) has a significantly negative effect on the score. We find that for each extra box opened, the score on average decreases by 1.1 percentage points. Intuitively, compared to the average score of 33.65, this is a decrease of about 3%. Furthermore, we observe that education levels have a positive effect on the total score.

When we look at the two search rules separately, we also find a negative effect of the total boxes on the score of Search-rule 1 and Search-rule 2. For each extra box opened, the score of Search-rule 1 decreases by 0.84 percentage points, which compared to the average score of 44.84, is approximately 2%. Thus, those who search more select the exact piece of information that has the highest benefit less often. Intuitively, opening five additional boxes thus results in a decrease of about 10% in the score of Search-rule 1. The effect on Search-rule 2 is of a smaller magnitude, with a 0.21 percentage point decrease for every extra box. This equals only a 0.3% change for every box in terms of percentages. On the face of it, this makes sense, as stated in H1, (in)attention predominantly works on the perception of $w_s(\tilde{\theta}^s, \sigma)$, which is the main feature of Search-rule 1. However, the fact that low levels of attention also affect Search-rule 2 can be explained by less attention to selecting the most informative attribute within the alternative, which is in line with earlier research on deviation from rational strategies and attention Gabaix (2019).

	Dependent variable:								
	Total Score (in %)			Score Search rule 1 (in %)			Score Search rule 2 (in %)		
	All	Oversearchers	Undersearchers	All	Oversearchers	Undersearchers	All	Oversearchers	Undersearchers
Total opened boxes	-1.123^{***}	-0.890^{***}	-1.940^{***}	-0.836^{***}	-0.514^{***}	-1.914^{***}	-0.206^{***}	-0.171^{*}	-0.264^{*}
	(0.061)	(0.060)	(0.160)	(0.074)	(0.075)	(0.198)	(0.060)	(0.071)	(0.133)
Round number	0.063	0.139^{*}	-0.039	0.045	0.061	0.042	0.125	0.106	0.077
	(0.048)	(0.059)	(0.087)	(0.059)	(0.066)	(0.110)	(0.084)	(0.120)	(0.123)
Male	3.148	1.265	4.871	2.517	-0.091	4.921	1.734	1.529	2.891
	(1.685)	(1.916)	(3.090)	(1.591)	(1.791)	(2.806)	(3.742)	(4.759)	(6.210)
Age	-0.056	-0.010	-0.099	-0.053	-0.049	-0.040	-0.051	0.009	-0.059
	(0.054)	(0.059)	(0.103)	(0.051)	(0.056)	(0.094)	(0.121)	(0.148)	(0.208)
Education level	1.511^{*}	1.637^{*}	1.825	1.121	1.204	1.961	1.014	0.281	2.580
	(0.707)	(0.791)	(1.361)	(0.668)	(0.740)	(1.238)	(1.571)	(1.964)	(2.734)
Monthly income	-1.101	-1.967^{*}	-0.233	-0.720	-0.569	-1.297	-2.024	-4.589^{*}	0.589
	(0.821)	(0.914)	(1.543)	(0.776)	(0.854)	(1.405)	(1.820)	(2.269)	(3.091)
Constant	33.654^{***}	30.759^{***}	37.284^{***}	44.853^{***}	39.784^{***}	50.624^{***}	67.642^{***}	73.134^{***}	57.244^{***}
	(3.644)	(4.043)	(7.017)	(3.469)	(3.822)	(6.448)	(8.048)	(10.000)	(13.975)
Observations	7,205	4,110	2,755	7,205	4,110	2,755	7,205	4,110	2,755
Subjects	393	208	167	393	208	167	393	208	167
Log Likelihood	$-31,841.780$	$-17,483.920$	$-12,657.610$	$-33,537.260$	$-18,498.260$	$-13,267.740$	$-31,759.620$	$-18,122.290$	$-12,178.800$
Akaike Inf. Crit.	63,705.560	34,989.840	25,337.220	67,096.520	37,018.510	26,557.490	63,541.240	36,266.580	24,379.600
Bayesian Inf. Crit.	63,781.260	35,059.380	25,402.350	67,172.230	37,088.040	26,622.620	63,616.950	36,336.110	24,444.740

Note: *p<0.05; **p<0.01; ***p<0.001

Table 6: H2: estimates of the relationship between the number of search actions and attention levels

We confirm the negative relationship between the amount of search and the attention

boxes irrespective of the participants' estimated attention.

measures for both over- and under-searchers. We control for the effect of the treatment manipulation as an additional explanatory variable, and the results, presented in Table B.6 in Appendix B.5 are essentially the same. In order to look deeper into the difference between over- and under-searchers we look at the interaction effects between the type of searcher and the total opened boxes. As can be seen in Table B.7 in Appendix B.6, for the score and Search-rule 1 we find a significantly larger effect for under-searchers compared to over-searchers. For Search-rule 2, we do not find a significant difference. Intuitively, this makes sense: given that under-searchers search less, they search mostly relative high-importance attributes and leave the lower attributes unexamined. The concave relationship discussed in Section 4.2 shows that by also including attributes of lower importance, attention levels are expected to decrease quickly.

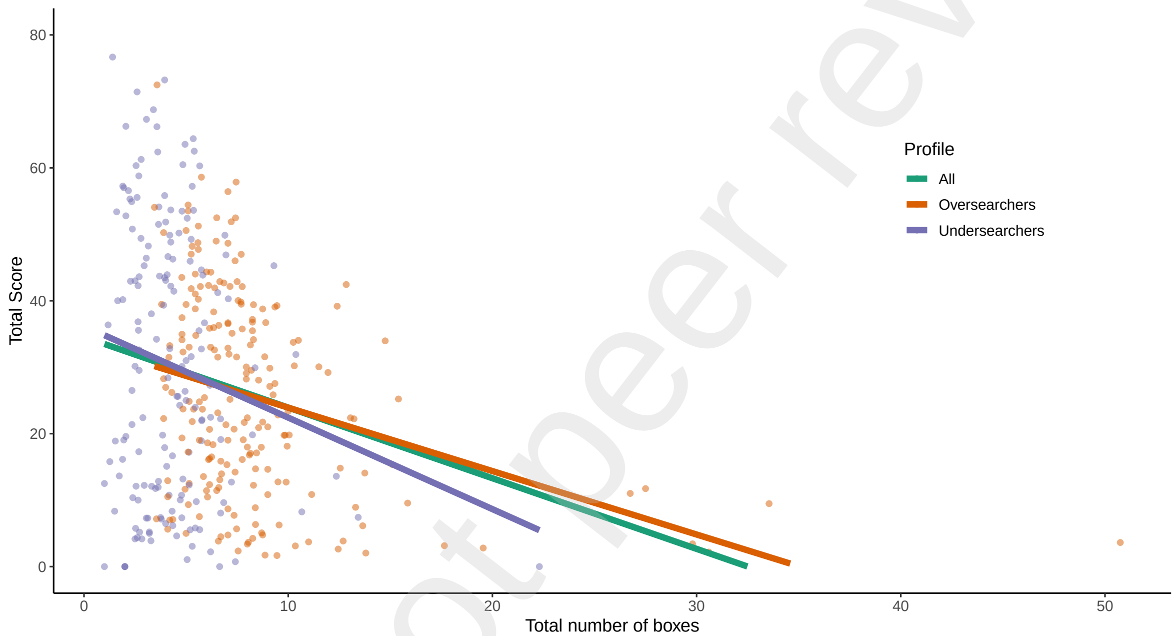

Figure 2: Relationship amount of search and total score

4.3.1. Summary of Hypothesis 2

As we can see in Figure 2, we observe a negative relationship between the number of search actions and the scores as described in the rational search-rule description in Section 4.1.1. In line with our findings in Table 6, we find a stronger reaction of under-searchers to a change in the number of search actions compared to over-searchers. Higher deviations from optimal actions indicate lower attention levels, and thus we can conclude that we find a negative relationship between the amount of search and the attention allocated to each piece of information, as predicted by our theoretical model in Section A.2.

4.4. Hypothesis 3: The effect of the communicated optimal strategy

4.4.1. The effect on the total amount of search and depth of search

For our Hypothesis 3.1, we focus on the *total opened boxes*, measuring the total number of attributes opened before reaching a decision. We hypothesized that this variable would be directly influenced by our treatment, i.e., the communication of an efficient level of search a_e^{stop}. The second variable of interest is the number of *attributes per alternative*, which states the average number of attributes explored per considered alternative, thus measuring the depth of the exploration. We expect the depth to be indirectly influenced by the communication of a_e^{stop} by the adjustments of m following the change of a^{stop} to a'^{stop}. Both measurements are discussed in Table 7 and 8. Additionally, for Hypothesis 3.2 we also estimate the effect of the communicated efficient level of search a_e^{stop} on the attention parameter m, that we hypothesized to play a role in moderating the depth of search after adjusting the amount of search. This result is discussed in Table 9.

Hypothesis 3.1 stated that the communicated optimum would have a convergent effect toward the efficient level of search. Intuitively, this would mean that the over-searchers decrease their amount of total search, while under-searchers increase their total amount of search. Additionally, the hypothesis states that the switching behavior, by moderating effects of m, follows the adjustment of the total amount of search. Our model predicts that for those who decrease (increase) their search toward a_e^{stop}, m would increase (decrease). Thereby the depth of search will decrease (increase) due to more (less) precise estimates, and a lower (higher) emphasis is put on the initial difference between the best and the second-best alternative. Table 7 shows the summary statistics of both variables and the payoff for the group without (T1) and with (T2) a simple heuristic with the optimal strategy communication. For both groups, we can see that the average number of boxes is close to 6, which falls within the range of the optimal amount of search actions. However, if we look at the median we see that T1 is further apart from this optimum than T2. Notable is that the standard deviation is smaller in T2 compared to T1 for all three variables, showing less variance between participants. Finally, there is an indication that the payoff might be higher ($p = 0.076$). In short, Table 7 shows a first indication of a convergent effect.

	T1		**T2**		**p-value**[1]
Variable (Round level)	***Mean (Std)***	***Median***	***Mean (Std)***	***Median***	
Total opened	6.006 (7.316)	4	5.946 (3.964)	6	<0.001
Attributes per alternative (depth of search)	2.093 (1.694)	1.8	2.177 (1.448)	2	0.1058
Payoff	26.877 (44.545)	27	29.953 (41.889)	31	0.0766

[1]: Mann-Whitney U test

Table 7: Summary statistics of main variables for Treatments

We see in Table 8, which shows the estimations of linear mixed-effects models at the individual level, that the average number of attributes explored, as depicted in the constant, is only a little under the optimal range of 6 - 11 boxes that were calibrated using the simulations.[22] The communication of the optimal strategy has, therefore in line with expectations, not a significant effect on average search behavior for the total sample.

[22]The full simulation can be found in Appendix B.1 Table B.1, including different levels of attention and the associated optimal levels search.

	Dependent variable:					
	Total opened			Attributes per alternative (depth of search)		
	All	Oversearchers	Undersearchers	All	Oversearchers	Undersearchers
Communication of the optimal strategy	0.185	−2.028**	0.967***	0.081	−0.392*	0.194
	(0.457)	(0.703)	(0.265)	(0.116)	(0.158)	(0.130)
Round number	−0.052***	−0.048*	−0.056**	−0.009*	0.001	−0.018**
	(0.014)	(0.019)	(0.021)	(0.004)	(0.005)	(0.007)
Male	0.447	0.424	0.050	0.098	−0.059	0.182
	(0.458)	(0.691)	(0.261)	(0.116)	(0.155)	(0.128)
Age	0.020	0.020	0.002	0.001	−0.002	0.002
	(0.015)	(0.021)	(0.009)	(0.004)	(0.005)	(0.004)
Education level	−0.011	−0.023	−0.170	−0.057	−0.083	−0.058
	(0.193)	(0.285)	(0.115)	(0.049)	(0.064)	(0.057)
Monthly income	0.123	0.321	−0.099	−0.119*	−0.150*	−0.108
	(0.221)	(0.331)	(0.130)	(0.056)	(0.074)	(0.064)
Constant	5.098***	8.203***	4.171***	2.469***	3.550***	1.850***
	(1.030)	(1.532)	(0.648)	(0.262)	(0.343)	(0.304)
Observations	8,000	4,160	3,340	8,000	4,160	3,340
Subjects	400	208	167	400	208	167
Log Likelihood	−22,092.370	−12,040.660	−8,270.026	−12,263.140	−6,717.280	−4,711.707
Akaike Inf. Crit.	44,206.740	24,103.320	16,562.050	24,548.280	13,456.560	9,445.414
Bayesian Inf. Crit.	44,283.600	24,172.990	16,629.300	24,625.140	13,526.230	9,512.665

Note: *p<0.05; **p<0.01; ***p<0.001

Table 8: H3.1 Amount of search and communication of the optimal strategy

If we split our sample into those that over-search and those that search less than theoretically optimal, we find a significant convergent effect for both groups. Over-searchers (over) search less when presented with the optimal strategy heuristic, while under-searchers search more, both significantly adjusting their behavior toward the communicated benchmark. We find that over-searchers decrease their search by approximately 25% (decrease of 2.03 boxes for an average of 8.20 boxes), while under-searchers increase their search by approximately 23%. Thus, we find a very similar convergent effect for both groups.

Similar to the total amount of search, we do not find any significant effect on the average depth for the entire sample following the communicated optimal strategy. However, when we separate our sample into the two types of over- and under-searchers, we find a significant decrease for over-searchers by 0.392 boxes, resulting in a little more than 11% decrease in depth per alternative. Similar to Hypothesis 2, we look at the interaction effect between the independent variable, the communication of optimal strategy heuristic, and the search type. We do not find any significant change in the depth of search for solely the under-searchers, as can be found in Appendix B.6 Table B.8. The depth is significantly different between over-searchers and under-searchers (t = -2.744). The round number has a negative effect on both the total search actions and the depth of the search.

4.4.2. Summary of Hypothesis 3.1

As shown in Figure 3, the communicated optimal strategy has no effect on average for the entire sample. When looking at the median, we see a small increase, bringing the number of search actions closer to the efficient level a_e^{stop} as can also be seen in Table 7. As expected, for over-searchers, we observe a strong decrease in the number of performed search actions for those in the treatment group and a small decrease in the depth of search. For under-searchers, we do see an increasing, convergent effect on the total number of search actions but do not observe a change in the depth of search. In conclusion, we find partial support for our third hypothesis; decision-makers of both types adjust their exploration behavior closer to the communicated optimal strategy. Over-searchers also diminish in the search depth, thus coming closer to the optimal simulations as reported in Table B.1 in Appendix B.1. For under-searchers, we find a convergent effect toward a_e^{stop} for the amount of search, but we do not observe the same effect on switching behavior.

4.4.3. The moderating effect of m

We also estimate our treatment variable's effect on the attention parameter m. Hypothesis 3.2 is developed under the assumption that if the amount of search is adjusted toward the optimum level a_e^{stop}, the attention parameter m would moderate this effect. Following the theoretical reasoning of Section 2.2, we would thus expect that over-searchers who decrease their total search would increase their attention levels, while under-searchers who increase their total amount of search might decrease their attention levels. In Table 9, presenting the estimations of linear mixed-effects models at the individual level, we see that the effect of the communication of the optimal strategy on the scores of the overall sample is positive but not significantly different from those participants that did not face the optimal strategy.

When we split up our sample into the two types again, we see a significant increase

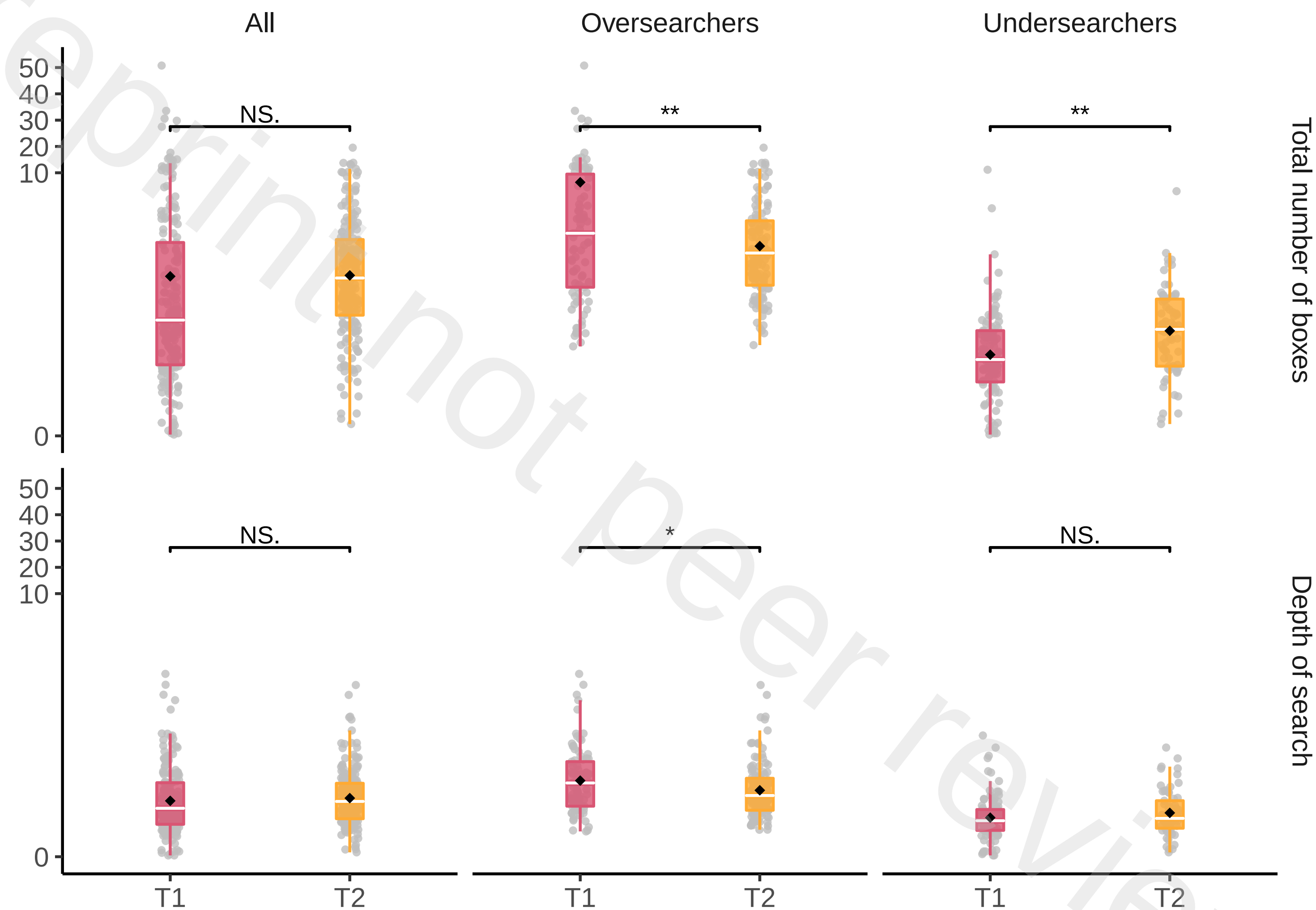

Figure 3: Comparison search amount between the two treatments for the different profiles
Note: $^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$; Two-tailed unpaired t-test comparisons
Red refers to T1 (control), orange to T2 (communicated optimal strategy)

	Dependent variable:								
	Total Score (in %)			Score Search rule 1 (in %)			Score Search rule 2 (in %)		
	All	Oversearchers	Undersearchers	All	Oversearchers	Undersearchers	All	Oversearchers	Undersearchers
Communication of the optimal strategy	2.072 (1.754)	5.950** (2.018)	−2.316 (3.159)	−0.194 (1.702)	6.455*** (1.857)	−4.838 (3.000)	5.122 (3.730)	9.339 (4.812)	−3.577 (6.291)
Round number	0.107* (0.051)	0.179** (0.062)	0.023 (0.093)	0.076 (0.061)	0.084 (0.068)	0.102 (0.118)	0.134 (0.084)	0.114 (0.120)	0.087 (0.123)
Male	2.921 (1.759)	0.983 (1.985)	5.514 (3.129)	2.470 (1.707)	−0.158 (1.826)	5.650 (2.976)	1.454 (3.739)	1.529 (4.732)	3.024 (6.212)
Age	−0.074 (0.057)	−0.038 (0.062)	−0.079 (0.104)	−0.065 (0.055)	−0.066 (0.057)	−0.025 (0.099)	−0.055 (0.120)	−0.004 (0.147)	−0.057 (0.208)
Education level	1.645* (0.738)	1.801* (0.819)	2.245 (1.384)	1.192 (0.717)	1.314 (0.754)	2.249 (1.318)	1.115 (1.570)	0.353 (1.953)	2.513 (2.747)
Monthly income	−1.460 (0.856)	−2.222* (0.948)	−0.619 (1.562)	−1.029 (0.832)	−0.636 (0.873)	−1.652 (1.489)	−1.983 (1.819)	−4.227 (2.262)	0.574 (3.091)
Constant	26.098*** (3.910)	20.913*** (4.365)	27.824*** (7.224)	40.170*** (3.809)	32.118*** (4.043)	42.676*** (6.904)	63.620*** (8.286)	65.835*** (10.394)	57.487*** (14.270)
Observations	7,205	4,110	2,755	7,205	4,110	2,755	7,205	4,110	2,755
Subjects	393	208	167	393	208	167	393	208	167
Log Likelihood	−32,008.460	−17,585.280	−12,728.900	−33,598.140	−18,515.430	−13,311.470	−31,764.550	−18,123.440	−12,180.610
Akaike Inf. Crit.	64,038.920	35,192.550	25,479.790	67,218.280	37,052.860	26,644.930	63,551.090	36,268.880	24,383.220
Bayesian Inf. Crit.	64,114.620	35,262.090	25,544.930	67,293.990	37,122.390	26,710.060	63,626.800	36,338.410	24,448.350

Note: *p<0.05; **p<0.01; ***p<0.001

Table 9: H3.2 Scores and communication of the optimal strategy

in the total score of about six percentage points for over-searchers. This is an increase of more than 28% compared to the baseline. We observe the same type of effect for the score of Search-rule 1, with an increase of about 20% compared to the baseline for over-searchers. This is in line with our assumptions regarding the moderating effect of attention; when over-searchers decrease their amount of search, compared to over-searchers that are not presented with the optimal strategy heuristic, we see an increase in the levels of attention. However, when looking at the under-searchers, we see a small but insignificant decrease in the score, yet we do find a significant difference between over- and under-searchers (interaction effects reported in Appendix B.6, Table B.8). We observe a similar pattern as in Table 8 that estimates the effect of the treatment on the depth of search, where we observe a change for over-searchers but not for under-searchers. Another thing to note in Table 9 is the positive effect of the round number for the over-searchers for the score. We do not see any significant effect for the score of Search rule 2, which is unsurprising given that the attention parameter m only feeds into Search rule 1. The changes in depth of search, hypothesized to take place due to the moderating role of m, seem to align with our model predictions.

4.4.4. Summary of Hypothesis 3.2

As we can see in Figure 4 we did not observe any difference in our attention measurement m on average between those to whom the optimal strategy was communicated and those to whom it was not. We see that the increase is mostly driven by over-searchers, where T2 results in an increase of about 31%, while the difference for under-searchers is not significant. In conclusion, we find partial support for the second part of Hypothesis 3.2, where the moderating role of m, following the communicated optimal strategy, is mostly present for over-searchers.

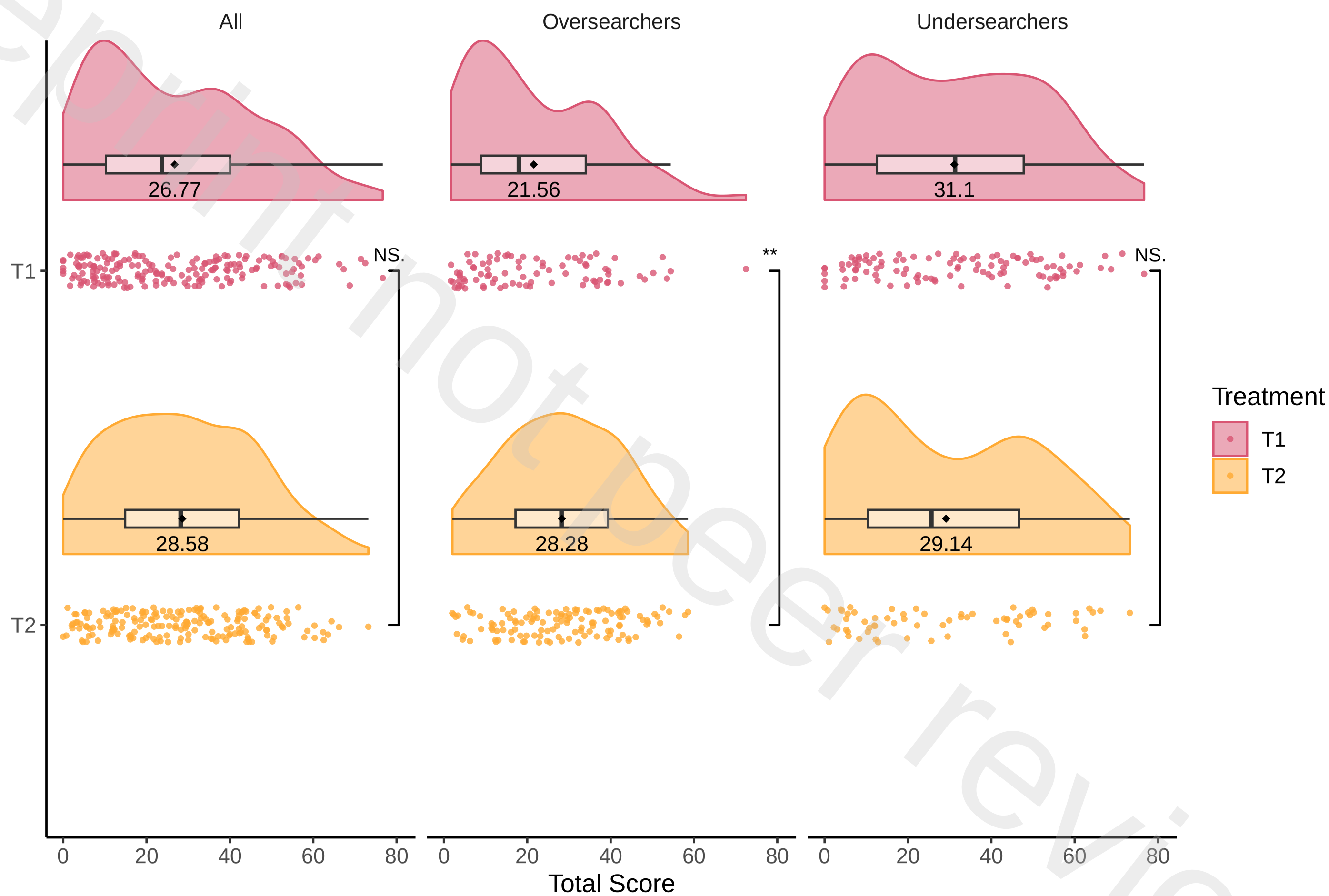

Figure 4: Comparison of total score between the two treatments for the different profiles
*Note: *p<0.05; **p<0.01; ***p<0.001; Two-tailed unpaired t-test comparisons*
Red refers to T1 (control), orange to T2 (communicated optimal strategy)

5. Robustness checks

5.1. Alternative measurements of attention

The deviation from rational behavior is only one way to measure attention. As discussed in Section 4.1, another way to estimate (in)attention is to physically measure the time between actions taken. We re-estimate the relationship between the total score and the average time between boxes in seconds using linear mixed-effects models at the individual level. We find in Table 10, in line with expectations, that for each extra second taken between two search actions, the score increases by 0.102 percentage points. Intuitively, this means that when our subjects contemplate their next decision longer, fewer deviations from rationality are present. In terms of magnitude, this means that the score increases by 0.4% for each extra second. Thus, by allocating one extra minute for a search action, the score would increase by an estimated 25%, which given the estimations of Table 4 would indicate doubling the score.

	Dependent variable:					
	Total Score (in %)			Total Score (in %)		
	All	Oversearchers	Undersearchers	All	Oversearchers	Undersearchers
Total opened boxes	-1.123^{***}	-0.890^{***}	-1.940^{***}			
	(0.061)	(0.060)	(0.160)			
Average time between boxes				0.102^{*}	0.097^{*}	0.113
				(0.050)	(0.048)	(0.190)
Round number	0.063	0.139^{*}	-0.039	0.145^{**}	0.184^{**}	0.095
	(0.048)	(0.059)	(0.087)	(0.047)	(0.058)	(0.086)
Male	3.148	1.265	4.871	2.689	1.504	4.093
	(1.685)	(1.916)	(3.090)	(1.695)	(2.000)	(3.001)
Age	-0.056	-0.010	-0.099	-0.074	-0.036	-0.074
	(0.054)	(0.059)	(0.103)	(0.054)	(0.062)	(0.100)
Education level	1.511^{*}	1.637^{*}	1.825	1.772^{*}	1.583	2.955^{*}
	(0.707)	(0.791)	(1.361)	(0.716)	(0.826)	(1.348)
Monthly income	-1.101	-1.967^{*}	-0.233	-1.652^{*}	-2.585^{**}	-0.777
	(0.821)	(0.914)	(1.543)	(0.823)	(0.953)	(1.492)
Constant	33.654^{***}	30.759^{***}	37.284^{***}	25.248^{***}	24.514^{***}	22.358^{***}
	(3.644)	(4.043)	(7.017)	(3.645)	(4.197)	(6.781)
Observations	7,205	4,110	2,755	6,673	4,032	2,318
Subjects	393	208	167	391	208	165
Log Likelihood	$-31{,}841.780$	$-17{,}483.920$	$-12{,}657.610$	$-28{,}750.560$	$-17{,}004.360$	$-10{,}287.430$
Akaike Inf. Crit.	63,705.560	34,989.840	25,337.220	57,523.130	34,030.730	20,596.860
Bayesian Inf. Crit.	63,781.260	35,059.380	25,402.350	57,597.990	34,100.050	20,660.090

Note: *p<0.05; **p<0.01; ***p<0.001

Table 10: Scores per type

5.2. Robustness check with different sub-sample and control variables

As an additional robustness check, we re-estimate our two main tables: 1) the effect of the communicated optimal strategy on the total opened boxes and the depth of search; and 2) the effect of the communicated optimal strategy heuristic on the total score, with a sub-sample of our participants, and separately with additional control variables. We include control variables that are related to capabilities and cognitive processing costs, such as the use of a calculator, self-reported mental capacity, performance in the comprehension questions, and prior participation in decision experiments. For our alternative sub-sample, we only include participants that made one or no mistakes in the comprehension questions prior to the game; this excludes 34% of our sample.

All of the estimations discussed in this paragraph can be found in Appendix B.7. For the entire sample, we find no difference (for any type of actor) in the number of search actions or the depth of search when we include additional control variables. For over-searchers, we find a positive effect of the performance in the comprehension questions, indicating that those over-searchers with more errors "over-search" even more. For this reason, we chose to re-run our estimation by excluding those that made more than one mistake in the comprehension questions. For this sub-sample, we find a slightly larger positive effect of the communicated strategy for the total number of boxes compared to the entire sample, with the treatment group ending up precisely in the communicated range of 6 to 11. The effect for over-searchers is no longer significant, with 8 boxes aptly in the communicated range. Finally, for the under-searchers, the increase is larger compared to the entire sample, thus also ending up in the communicated range. For the depth of search, we find that contrary to our full sample, the increase in depth by the under-searchers is now also significant on the 5% level. We conclude that the sub-sample does not differ significantly in most estimations but closer approximates the optimal rational behavior.

We make the same two specifications for the moderating role of attention m; they can be found in Appendix B.7, Tables B.11 and B.12. We find no significant differences for any score estimation neither for the specification of alternative control variables nor for the different sub-sample. In line with our expectations, we find a positive effect of mental calculation capacity and a negative effect of the number of errors in comprehension questions.

5.3. Learning over time

Figure 5 shows the distributions of the distance to the optimal number of boxes (top) and distance to the optimal depth of search (bottom), averaged per person over the first 10 and the last 10 rounds. We see that over-searchers are more often to the left of the zero line, thus searching more than the optimal number of search actions, while for under-searchers the distribution for both measures is often more to the right of the zero line, indicating that they search less than optimally. We compare the shaded distribution of the first 10 rounds with the colored distribution of the last 10 rounds (using two-tailed paired t-test). We find that no differences can be observed between the first 10 and the last 10 rounds for those with the communication of the optimal strategy (T2). For those without the strategy, we see some learning behavior for the total number of boxes (for all, and over-searchers; p-value < 0.01 for both). It seems that those that have the heuristic to steer behavior stay relatively steady in their search strategy, whereas those without move a little closer to the optimum over time.

Additionally, we estimate the triple interaction effect of treatment*type*last rounds, the estimation can be found in Appendix B.8. For the distance to the total number of boxes, we see that under-searchers do not improve the optimal distance significantly over time, both for those with the communication of the optimal strategy (T2) and without

(T1). We also see that over-searchers that are in T1 do worse than the under-searchers and over-searchers in T2, as can also be seen by the big tale of the distribution on the left. However, comparing the first 10 and last 10 rounds, over-searchers without optimal strategy communication do improve. Over-searchers with the optimal strategy communicated are closer to the optimum in the first 10 rounds and stay close to this distribution in the final 10 rounds. All in all, Figure 5 and the interaction effects reported in Appendix B.8 suggest that our treatment intervention is most beneficial for over-searchers, even though over-searchers without the communication also improve over time. The communicated optimal strategy helps steer behavior early onward and gets actors closer to the optimum than those without.

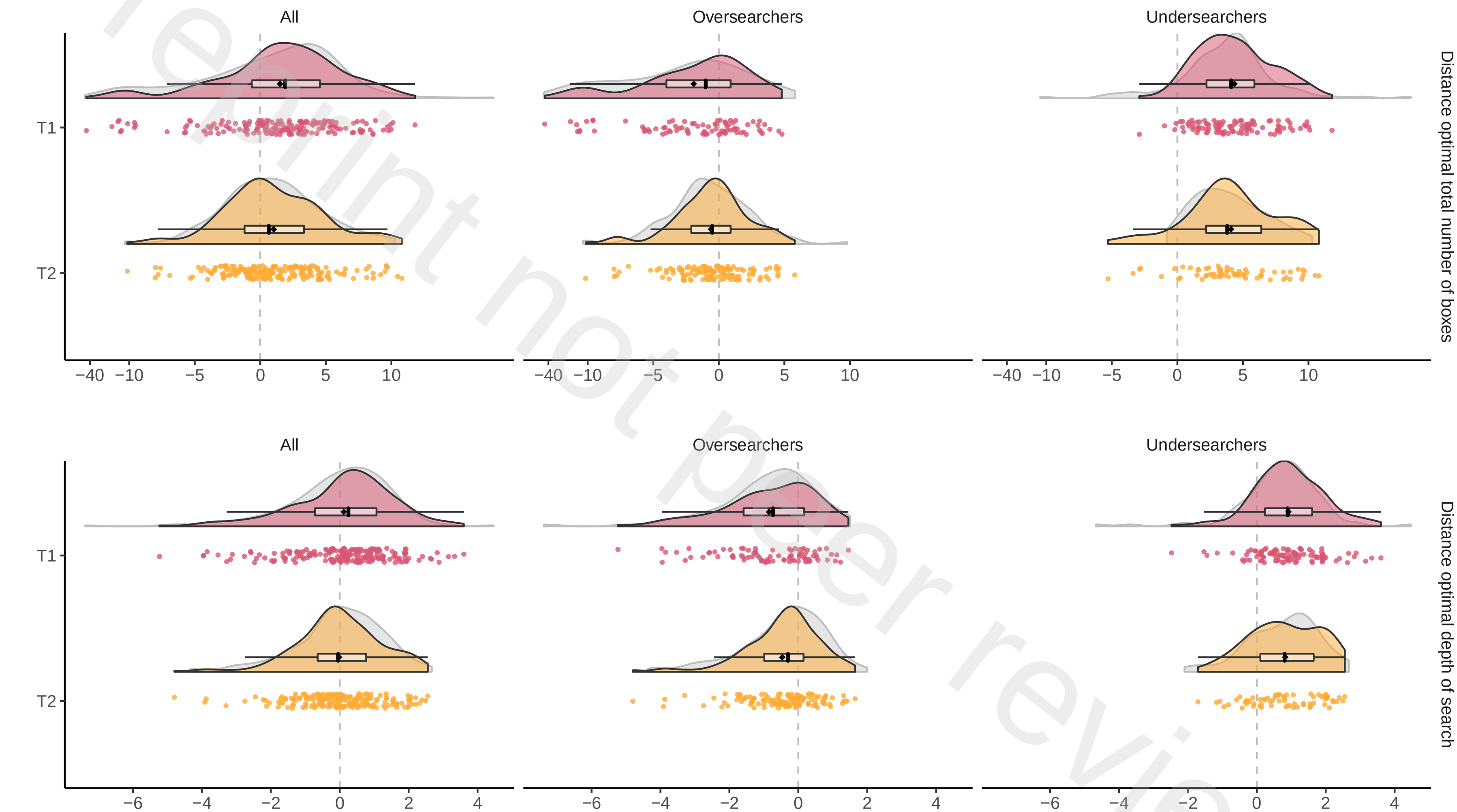

Figure 5: Evolution of the distance to the optimal strategy

Note: the shaded distributions represent the first ten rounds. The colored distributions and the boxplots represent the last ten rounds. The black bold crossbar represent the medians, and the black dots the means of the distributions of the last ten rounds.

6. Discussion

Multi-attribute search problems are essential to most day-to-day decision-making. However, contrary to most traditional models, actors deal with most of these decisions without full attention to all available information. We contribute to the understanding of multi-attribute decision-making by studying the final decisions, and the intermediate search process, and by estimating the moderating role of attention for both outcomes. Moreover, we investigate how the communication of an optimal strategy on one dimension of the search process, influences the different dimensions of the rational decision-making process, as well as the moderating attention. To the best of our knowledge, our study is the first to combine the multi-attribute element in an abstract setting where we attempt to improve decision-making with a simple heuristic of communicating the optimal amount of search.

Hypothesis 1 investigated the role of (in)attention in the decision-making process. We find that the search strategies of our actors are better described with partial attention, thus rejecting the notion of full attention as suggested in traditional economic models. Additionally, as attribute importance decreases, so does the level of allocated attention, where a concave relationship exists between the amount of attention allocated and the attribute importance. Both findings are consistent with previous work Gabaix (2019), where different studies are merged to estimate the level of attention m and its relationship with the importance of the attribute. The similarity with earlier work indicates that our multi-attribute setting is a suitable decision framework to test our other two hypotheses. Furthermore, our research adds another estimate of the attention parameter m to the literature for the so far scarce collection of estimates.

Hypotheses 2 and 3 offer insights into the interplay of the amount of search and the estimates of attention m allocated. To test Hypothesis 2, we analyzed the amount of search and the levels of m and found that those who search more allocate lower levels of attention per attribute, as also predicted by our model. To test Hypothesis 3, we investigated this interplay a step further and found that the amount of search can be altered by introducing a simple heuristic for decision-making; stating the optimal range of search actions. Second, we found that over-searchers rationally adjust their search depth accordingly. To analyze the moderating role of attention, Hypothesis 3 researched the effect of the heuristic on allocated attention m. Here, we found that over-searchers increase their attention following the communicated strategy. It indicates a causal mechanism since those who adjust their amount of search due to an external stimulus adapt their levels of attention accordingly. From a practical point of view, if full attention is of importance, less information should be stimulated to be considered. From a modeling perspective, this is also interesting, indicating that attention per decision might be finite and, when larger choice sets are available, attention is not necessarily increased accordingly, but part of the choice set might fall outside the possibilities of consideration Caplin et al. (2019).

One interesting finding is that under-searchers, contrary to over-searchers, do not adjust their depth and attention levels following the communicated strategy. One possible explanation for this behavior is that many under-searchers already have a starting position that is close to the communicated range, see Figure 6. This has the implication that for this group there is a smaller potential to get better. As we can see in the figure, the distance to the optimal range is often larger for over-searchers compared to under-searchers. Connected to this, is the average amount of search that under-searchers obtain after observing the heuristic. Even after increasing their amount of search toward a_e^{stop}, we find that under-searchers examine, on average, five attributes. Our simulations predict partial attention only if the number of attributes rises above six. If the decision-makers are fully rational, all amounts below six would indicate the same level of full attention, and thus the stability of deep search and allocated attention levels m are not surprising. Our robustness estimations support this logic, as in our sub-sample estimations under-searchers also significantly adjust their depth when performing slightly over six search actions. Lastly, to dive deeper into the differences between over- and under-searchers we look at a number of (self-reported) survey variables of our participants, as shown in Table B.3 and B.4 in Appendix B.4. We find no difference in any of our tested demographic variables or for the self-reported strategies between the two types. However, we do find a significant difference between the two types in terms of self-reported motivation. Over-searchers are mostly motivated by finding a good (21%) or the best choice (46%), while under-searchers score higher on cost-related motivations, such as not spending too many points (46%).

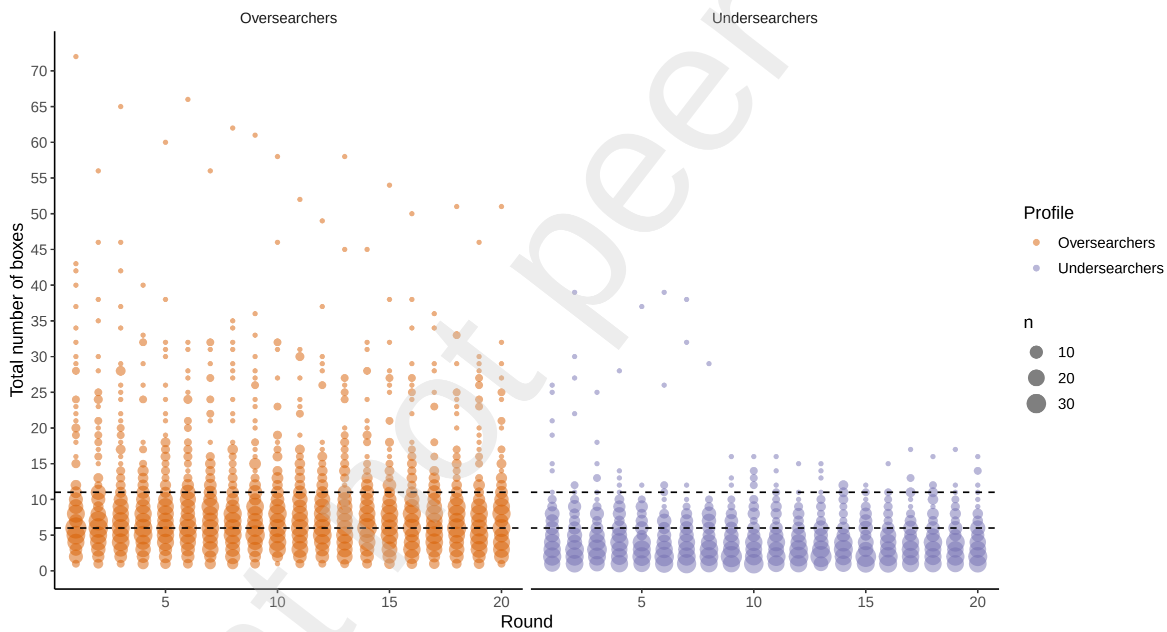

Figure 6: Evolution of average total number of opened boxes per profile

In contrast to previous findings, we found a lower level of total search and calibrated a higher level of (partial) attention. Within multi-attribute (intractable) environments, the average action $\overline{a}_i$ is often too high. This is attributed to the fact that search consists of both exploring the alternatives as well as exploring the attributes of each alternative. While the former refers to searching for a satisfactory alternative, the latter aims at reducing the ambiguity regarding the level of satisfaction of the examined alternative Sanjurjo (2017). For a decision, both "parts" need to be satisfied. While some efficient level a_e^{stop} can be derived for a given particular environment, e.g., given costs k and benefits $w_s(\tilde{\theta}^s, \sigma)$, recent studies have shown that often $a_i^{stop} > a_e^{stop}$, i.e., people tend to consider too many attributes, especially with respect to deep within alternative search, given the cost k_i (see for example Bearden and Connolly, 2008; Gabaix et al., 2006;

Sanjurjo, 2017). We observe that this over-search behavior is significantly lower and closer to the optimum than earlier work on the observed search behavior. For example, participants on average over-search in earlier research, while we have a relatively evenly split sample with 56% of our participants qualifying as over-searchers Gabaix et al. (2006); Sanjurjo (2017). An interesting direction for subsequent research is to disentangle further what determines whether an actor acts as an over- or under-searcher. Similarly, it would be interesting to see if these two types make consistently different decisions in a range of different contexts.

In line with our model predictions, this lower amount of total search also results in higher initial estimates of m, starting at 0.95 for the highest level of attribute importance and only decreasing by 0.13 for each next attribute. Previous research found estimates ranging from $m = 0.76$ for an attribute importance of 0.58 Gabaix (2019). In contrast, we find estimates of $m = 0.95$ for an attribute importance of 0.30. Several reasons can explain this difference. First, in our multi-attribute setting, the first attribute also determines which alternative will be examined first; thus, even though the attribute importance is not as high, the importance of the first examination is more deterministic for the final decisions compared to a single-attribute case where the alternative already has been chosen. Second, and more importantly, our particular setting might prompt higher levels of attention compared to the settings earlier mentioned; most research has been done on the attention estimates of so-called side-attributes (think for example about shipping costs or sales tax). Even though the payoffs are lower in our setting compared to say choosing the right car, as researched by Lacetera et al. (2012), our setting might prompt higher levels of attention. In our setting *all* attributes –the most and least important– are included in the measure for allocated attention, rather than just the most important attributes, as is often done in multi-attribute research. Additionally, search costs are made explicit. Finally, prior experiments have mostly been done with student samples in the lab. While in the laboratory, participants often have to stay until everyone is finished and thus are stimulated to explore more alternatives. Our setting allows, just as in daily life, to allocate as much time as participants prefer and perhaps examine fewer attributes with higher attention to make quicker decisions. This more realistic setting might be even more relevant for subjects with higher opportunity costs, i.e., the share of non-students in our representative UK sample. The higher attention levels have important implications for future research, as it indicates that different settings can still be used to study most mechanisms, yet with different starting levels of attention. Still, realistic levels of (in)attention should be substantially adjusted toward the exact question of interest.

One challenge of inattention research is the inability to measure attention directly. As discussed in Section 4.1, we choose mouse tracing to observe the intermediate steps and to see how our participants interpret the information by observing their next move. In the robustness checks, we investigate the time between search actions as an alternative proxy for our original attention measurements and find a positive effect between the amount of time allocated between moves and the level of attention, indicating that time could serve as another measurement of (in)attention. However, with the new advancements in eye tracing, intermediate steps can also, in the near future, be observed with increased precision. With increased precision, they can serve not only single-attribute research but also multi-attribute search and will have the potential to make big advancements in attention research. A similar challenge involves the incorporation of cognitive costs.

In this research, we did not incorporate the role of cognitive costs, as adding positive cognitive costs could only result in underestimating our hypothesized effects for our particular setting. However, to get a more precise insight into the role of (in)attention, advancements in measurements and proxies of cognitive costs of decision-making should also be incorporated in decision-tracing research.

Our research suggests that the attention given to certain information signals depends on the amount of information and the relative importance of that information compared to the other signals present. In order to develop better policy recommendations, more research should be done on the relationship between this trade-off and additionally on information volumes and information processing for instructive aspirations. Likewise, our research gives a first indication that attention and the amount of information considered for decision-making can be easily steered with a simple heuristic. Future research should dive deeper into the effects of decision context too, for this research we used a neutral experimental setting, however in real life information also influences utility by different channels. For example, Tinghög et al. (2022) discuss the identity-protective cognition, coined Homo Ignorans, that avoids, neglects, or distorts information that can threaten their own identity. Different heuristics and contexts should be tested for instructive aspirations and guidance in (commercial) decision strategies, as this could have consequential effects on how we communicate and help others make the best decisions depending on their needs.

7. Conclusion

The trade-off of how much information to incorporate in order to make a decision versus the cost that considering this extra information adds is essential for the decision-making process. Attention plays an essential role in this trade-off, even though most traditional economic models do not incorporate inattention Gabaix (2019). Therefore, our finding that attention is often incomplete in multi-attribute settings has important implications for advances in decision research. Specifically, our research suggests that not only the search process can be improved by communicating a simple heuristic, but decision-makers also adjust their attention levels as predicted. Future research should incorporate (new) attention measures and adjust economic modeling to include the moderating role of attention in more detail, as well as in different decision contexts where advancements are made to improve decision-making.

Declarations of interest

None.

A. Appendix A

A.1. Theoretical framework – extended version

We make use of Gabaix's (2019) simple framework for modeling attention, with some slight departures. Sections A.1 and A.2 are a simplified repetition of Gabaix's framework; for a more detailed description of the equations and steps between, please refer to Gabaix (2019). In the framework, an agent's behavior is modeled with a prior value of alternative X and adjustment toward perceived signals with Gaussian noise. We start with the assumption of a true value x, drawn from a Gaussian distribution $\mathcal{N}(x^d, \sigma^2)$, with x^d the default value, in this example, the prior mean μ and variance σ^2. The agent does not know the true value of x and instead gets a signal s, where s is:

$$s = x + \varepsilon, \tag{A.1}$$

where ε is drawn from an independent distribution $\mathcal{N}(0, \sigma^2)$. An agent takes action a based on the signal they receive. The agent's objective function is: $u(a, u) = -\frac{1}{2}(a - x)^2$. A rational agent therefore solves $max_a \, \mathbb{E}[-\frac{1}{2}(a - x)^2|s]$. The agent estimates the value of x given the (noisy) signal s they receive. When they optimize, the first-order search rule is:

$$0 = \mathbb{E}[-(a - x)|s] = \mathbb{E}[x|s] - s, \tag{A.2}$$

thus a rational agent takes action $a(s) = \hat{x}(s)$. With action a being the behavioral response and $\hat{x}(s)$ the expected value of x given observed s, $\hat{x}(s)$ is defined as:

$$\hat{x}(s) = \mathbb{E}[x|s] = \lambda + (1 - \lambda)x^d, \tag{A.3}$$

with λ defined as:

$$\lambda = \frac{\sigma_x^2}{\sigma_x^2 + \sigma_\epsilon^2} \in [0, 1], \tag{A.4}$$

Equation A.3 shows that the agent should start their estimate at the prior value x^d, in this general case, the prior mean of x^d, and thereafter partially adjust toward the signal s with a dampening factor. The average action $\hat{a}(x) := \mathbb{E}[a(s)|x]$ therefore is:

$$\hat{a}(x) = mx + (1 - m)x^d, \tag{A.5}$$

with $m = \lambda$, $m \in [0, 1]$. We will use the notation m to refer to the limited reaction to the signal s. In this chapter, we use the Bayesian ratio of variances λ, but this does not always need to be the case in a more general context. For exceptions and more details, see the chapter on Behavioral Inattention by Gabaix (2019). In the case of zero noise, an agent updates perfectly when receiving signal $s, \sigma_\epsilon^2 = 0$ and therefore $m = 1$. In the opposite limit, an agent with infinite noise ($\sigma_\epsilon^2 \to \infty$) results in $m = 0$, they do not update their default x^d at all after receiving a signal, and thus their estimate remains x^d.

A.2. Deterministic attention and action: Multi-attribute search

We continue with the assumption that the agent starts from the default value x^d and, based on their expectations, adjusts it toward the true x, where partial adjustments result in a new perceived value of $\hat{x}(s)$, which lies between the default value x^d and the true value x. This is a realistic assumption, as signals are generally imprecise. In similar reasoning, cognitive capacity, especially under the circumstances of information overload, also prohibits m from being 1 for all possible signals. Continuing to build on the simple framework of Behavioral Inattention gabaix2019behavioral, we follow the structure where the agent should maximize:

$$\max_a u(a, x), \tag{A.6}$$

where a, as before, is an action and x is a vector of "attributes" (think for example of prices, different characteristics of the good, social impact such as (fair) trade-marks, etc., as different attributes of a good). The rational agent thus will choose action $a^r(x) = \arg\max_a u(a, x)$. For the behavioral agent, this is replaced by the "attention-augmented decision utility,"

$$\max_a u(a, x, m), \tag{A.7}$$

with m being the degree of attention, resulting in the agent's subjective perspective of the world. They, therefore, take action:

$$a(x, m) = \arg\max_a u(a, x, m), \tag{A.8}$$

given our multi-attribute setting and similarly to other works chetty2009salience, gabaix2014sparsity, dellavigna2009psychology, the utility function takes the form of:

$$u(a, x, m) = u(a, m_1 x_1 + (1 - m_1)x_1^d, \ldots, m_i x_i + (1 - m_i)x_i^d), \tag{A.9}$$

where $m_i \in [0, 1]$ is the attention to variable x_i, and x_i^d is the default value for variable x_i (the expected value of the agent without a signal).

Agents are often not fully attentive. The value of one of the possible alternatives for multi-attribute search Equation A.5 and A.9 would in the case of inattention give us:

$$A_n^{s,a} = \sum_i^{j_n} [m_i x_i + (1 - m_i)x_i^d], \tag{A.10}$$

where $A_n^{s,a}$ is the sum of all subjective values x_i^s of each attribute selected by the search actions j_n. As we can see, if $m_i = 0$, the agent does not update A_n^s at all and sticks with the default value x_i^d, so that the subjective value of the attribute $x_i^s = x_i^d$. When $m_i = 1$ there is full attention to the signal and the agent perceives the true value of $x_i^s = x_i$, thus $A_n^s = A_n$. Whenever $0 < m_i < 1$ they update with partial attention, so that $x_i^s \in (x_i^d, x_i)$.

B. Appendix B

B.1. Simulation for intervention

Initial inattention	Additional inattention	**Number of explored boxes**	Payoff of the round	Number of explored rows
0.0	0.1	**0.0**	28.1343	0.0
0.1	0.1	**64.6813**	-2.924	8.0
0.2	0.1	**50.7477**	10.9077	7.9257
0.3	0.1	**36.8134**	24.3255	6.9749
0.4	0.1	**26.075**	34.9089	5.5791
0.5	0.1	**19.0268**	39.1766	4.5411
0.6	0.1	**14.3104**	41.2327	3.8645
0.7	0.1	**10.7786**	42.991	3.3364
0.8	0.1	**8.7915**	41.8903	3.0468
0.9	0.1	**7.0167**	41.3979	2.7728
1.0	0.1	**5.9462**	40.7623	2.5958

Table B.1: Simulations of search behavior with cost equal to 1 and the standard deviation equal to 20

B.2. Comprehension summary statistics

Statistic	Mean	St. Dev.	Median
Frequency of correct answers in question 1	0.785	0.411	1
Frequency of correct answers in question 2	0.790	0.408	1
Frequency of correct answers in question 3	0.815	0.389	1
Frequency of correct answers in question 4	0.422	0.495	0
Total number of errors in the questionnaire	1.188	0.916	1

Table B.2: Summary statistics of the comprehension questionnaire

B.3. Individual inattention parameters fitting

We individually fit the inattention using the following steps:

1. We use the participants' choices and compute for each participant and at each round the corresponding score of search-rule 1. We compute these corresponding scores based on different values of the inattention parameters m_1, m_2 and *power*. We consider several combinations of the inattention parameters. For computational power limitations, we restrict our parameters space to: $m_1 \in \{0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0\}$, $m_2 \in \{-0.1, 0.0, 0.1, 0.2\}$ and $power \in \{0.5, 1.0, 2.0\}$. We thus consider in total 84 possible combinations of parameters.[23]

[23]Note that an additional value of m_1 adds 12 combinations, an additional value of m_2 adds 21 combinations and an additional value of *power* adds 28 combinations.

2. We then fit the inattention parameters to the participant choices for each participant and at each round as follows: We consider, in the following order, m_1 that maximizes the score of search-rule 1 (we take the highest m_1 if multiple obtained values), then m_2 that maximizes the score of search-rule 1 given the fitted m_1 (we select one of the most frequently obtained values of m_2 if multiple obtained values), then *power* that maximizes the score of search-rule 1 given the fitted m_1 and m_2 (we select one of the most frequently obtained values of *power* if multiple obtained values).
3. We finally pool at the participant level by taking the average of the fitted values at each round to obtain the individual inattention parameters.

Note that the fitted inattention parameters that we obtain for m_1 and *power* (see Table 5) lie far from the boundaries that we set for our parameters space: a minimum of 0.79 for m_1, and a *power* parameter contained between 0.625 and 1.636. The minimum obtained value for m_2 (0.06) also lies far from the chosen boundary. Its maximum obtained value (0.2) is however at the boundary, but this value occurred only around 0.76% of the time.

B.4. Additional tables

Variable	**Categories**	**Oversearchers**		**Undersearchers**	
		Nº	**Percent**	**Nº**	**Percent**
Gender	Female	100	48.1%	85	50.9%
	Male	108	51.9%	81	48.5%
	Other	0	0%	1	0.6%
Age	18-24	25	12%	14	8.4%
	25-49	95	45.7%	84	50.3%
	50-64	54	26%	53	31.7%
	Over 64	34	16.3%	16	9.6%
Occupation	Employed	107	51.4%	97	58.1%
	Self-employed	26	12.5%	19	11.4%
	Student	20	9.6%	12	7.2%
	Unemployed	17	8.2%	16	9.6%
	Other	38	18.3%	23	13.8%
Education	High-school	45	21.6%	39	23.4%
	Vocational training	30	14.4%	17	10.2%
	Bachelor	81	38.9%	79	47.3%
	Masters	40	19.2%	23	13.8%
	PhD	8	3.8%	7	4.2%
	Other	4	1.9%	2	1.2%
Income	0-1500 GBP	83	39.9%	67	40.1%
	1501-3000 GBP	81	38.9%	63	37.7%
	3001-5000 GBP	28	13.5%	26	15.6%
	$\geq$ 5000 GBP	5	2.4%	4	2.4%
	I don't know	11	5.3%	7	4.2%
Total		**208**		**167**	

Table B.3: General characteristics of sample per type

Variable	**Categories**	**All**		**Oversearchers**		**Undersearchers**	
		Nº	**Percent**	**Nº**	**Percent**	**Nº**	**Percent**
Motivation	Being certain that I chose a good choice	62	15.5%	43	20.7%	17	10.2%
	Finding the best choice	170	42.5%	95	45.7%	62	37.1%
	Not spending points in opening boxes	136	34%	50	24%	77	46.1%
	Not spending too much time	28	7%	18	8.7%	9	5.4%
	Other	4	1%	2	1%	2	1.2%
Strategy	Always the same amount of boxes	4	1%	3	1.4%	1	0.6%
	Always explore the same number of alternatives	9	2.2%	5	2.4%	4	2.4%
	Always start at the top	16	4%	10	4.8%	5	3%
	When I see 2 negative in a row I switch	56	14%	34	16.3%	20	12%
	Keep opening until I have reached a certain value	20	5%	15	7.2%	5	3%
	Opening at least 1 attribute per	13	3.2%	8	3.8%	5	3%
	Choose the option with the highest open value in the beginning	239	59.8%	113	54.3%	110	65.9%
	Other	43	10.8%	20	9.6%	17	10.2%
Total		**400**		**208**		**167**	

Table B.4: Strategy and Motivation per type

Margin	Crossovers
5%	17.25%
10%	13.5%
15%	5.5%
20%	3.25%

Table B.5: Frequency of crossovers between extreme profiles

B.5. Mixed model with additional controls

	Dependent variable:								
	Total Score (in %)			Score Search rule 1 (in %)			Score Search rule 2 (in %)		
	All	Oversearchers	Undersearchers	All	Oversearchers	Undersearchers	All	Oversearchers	Undersearchers
Total opened boxes	-1.122^{***}	-0.882^{***}	-1.938^{***}	-0.837^{***}	-0.493^{***}	-1.903^{***}	-0.206^{***}	-0.168^{*}	-0.263^{*}
	(0.061)	(0.061)	(0.160)	(0.074)	(0.075)	(0.198)	(0.060)	(0.071)	(0.133)
Communication of the optimal strategy	1.591	3.516	−1.098	−0.527	5.165^{**}	−3.563	5.083	8.989	−3.415
	(1.680)	(1.942)	(3.122)	(1.588)	(1.799)	(2.816)	(3.727)	(4.808)	(6.289)
Round number	0.063	0.139^{*}	−0.039	0.045	0.062	0.044	0.125	0.106	0.078
	(0.048)	(0.059)	(0.087)	(0.059)	(0.066)	(0.110)	(0.084)	(0.120)	(0.123)
Male	3.077	1.245	4.907	2.533	−0.011	5.012	1.473	1.574	3.009
	(1.685)	(1.902)	(3.090)	(1.592)	(1.759)	(2.791)	(3.736)	(4.724)	(6.209)
Age	−0.056	−0.014	−0.099	−0.053	−0.054	−0.042	−0.052	0.0002	−0.060
	(0.054)	(0.059)	(0.103)	(0.051)	(0.055)	(0.093)	(0.120)	(0.147)	(0.208)
Education level	1.543^{*}	1.689^{*}	1.777	1.112	1.260	1.812	1.099	0.337	2.433
	(0.707)	(0.786)	(1.367)	(0.668)	(0.726)	(1.237)	(1.569)	(1.950)	(2.745)
Monthly income	−1.069	-1.859^{*}	−0.230	−0.729	−0.423	−1.288	−1.903	−4.143	0.624
	(0.821)	(0.909)	(1.543)	(0.776)	(0.841)	(1.397)	(1.817)	(2.258)	(3.090)
Constant	32.730^{***}	28.419^{***}	37.833^{***}	45.155^{***}	36.300^{***}	52.359^{***}	64.783^{***}	67.164^{***}	58.862^{***}
	(3.760)	(4.213)	(7.170)	(3.580)	(3.943)	(6.547)	(8.286)	(10.396)	(14.277)
Observations	7,205	4,110	2,755	7,205	4,110	2,755	7,205	4,110	2,755
Subjects	393	208	167	393	208	167	393	208	167
Log Likelihood	−31,841.350	−17,482.420	−12,657.550	−33,537.210	−18,494.240	−13,266.950	−31,758.710	−18,120.670	−12,178.660
Akaike Inf. Crit.	63,706.700	34,988.840	25,339.100	67,098.410	37,012.480	26,557.910	63,541.430	36,265.330	24,381.310
Bayesian Inf. Crit.	63,789.290	35,064.690	25,410.150	67,181.000	37,088.330	26,628.960	63,624.020	36,341.180	24,452.360

Note: *p<0.05; **p<0.01; ***p<0.001

Table B.6: H2: estimates of the relationship between the number of search actions and attention levels while controlling for the treatment

B.6. Mixed models with interaction effects for type of searcher

	Dependent variable:		
	Total Score	Score Search Rule 1	Score Search Rule 2
Total opened boxes	-1.927^{***}	-1.917^{***}	-0.261
	(0.135)	(0.168)	(0.134)
Oversearcher	-6.553^{***}	-14.117^{***}	2.424
	(1.918)	(1.881)	(3.914)
Total opened boxes * Oversearcher	1.028^{***}	1.403^{***}	0.092
	(0.152)	(0.189)	(0.151)
Round number	0.067	0.053	0.099
	(0.050)	(0.059)	(0.087)
Male	2.971	1.960	2.224
	(1.723)	(1.575)	(3.816)
Age	-0.049	-0.046	-0.031
	(0.055)	(0.050)	(0.122)
Education level	1.724^{*}	1.439^{*}	1.370
	(0.732)	(0.670)	(1.621)
Monthly income	-1.152	-0.732	-2.333
	(0.837)	(0.766)	(1.851)
Constant	37.063^{***}	52.513^{***}	64.959^{***}
	(3.863)	(3.594)	(8.443)
Observations	6,865	6,865	6,865
Subjects	375	375	375
Log Likelihood	$-30{,}335.790$	$-31{,}931.690$	$-30{,}306.320$
Akaike Inf. Crit.	60,697.580	63,889.370	60,638.630
Bayesian Inf. Crit.	60,786.420	63,978.220	60,727.480

Note: *p<0.05; **p<0.01; ***p<0.001

Table B.7: Interaction effects for H2

	Dependent variable:		
	Total opened	Attributes per alternative (depth of search)	Total Score (in %)
Communication of the optimal strategy	1.032	0.192	−2.331
	(0.617)	(0.158)	(2.717)
Oversearcher	6.432***	1.458***	−10.363***
	(0.571)	(0.146)	(2.498)
Communication of the optimal strategy * Oversearcher	−3.175***	−0.579**	8.550*
	(0.826)	(0.211)	(3.609)
Round number	−0.051***	−0.008	0.116*
	(0.014)	(0.004)	(0.053)
Male	0.237	0.050	3.057
	(0.403)	(0.103)	(1.760)
Age	0.014	−0.001	−0.058
	(0.013)	(0.003)	(0.056)
Education level	−0.082	−0.076	1.964**
	(0.172)	(0.044)	(0.749)
Monthly income	0.138	−0.126*	−1.439
	(0.196)	(0.050)	(0.856)
Constant	2.752**	2.002***	29.581***
	(0.940)	(0.241)	(4.065)
Observations	7,500	7,500	6,865
Subjects	375	375	375
Log Likelihood	−20,657.080	−11,523.050	−30,509.060
Akaike Inf. Crit.	41,340.160	23,072.110	61,044.110
Bayesian Inf. Crit.	41,430.160	23,162.100	61,132.960

Note: *p<0.05; **p<0.01; ***p<0.001

Table B.8: Interaction effects for H3

B.7. Mixed model sub-sample and alternative controls

	Dependent variable:					
	Total opened			Attributes per alternative (depth of search)		
	All	Oversearchers	Undersearchers	All	Oversearchers	Undersearchers
Communication of the optimal strategy	0.173	−2.022**	0.964***	0.082	−0.408**	0.216
	(0.456)	(0.685)	(0.267)	(0.116)	(0.155)	(0.131)
Round number	−0.052***	−0.048*	−0.056**	−0.009*	0.001	−0.018**
	(0.014)	(0.019)	(0.021)	(0.004)	(0.005)	(0.007)
Male	0.575	0.872	0.033	0.128	0.035	0.185
	(0.461)	(0.691)	(0.262)	(0.117)	(0.156)	(0.128)
Age	0.020	0.023	0.002	0.001	−0.002	0.002
	(0.015)	(0.021)	(0.009)	(0.004)	(0.005)	(0.004)
Education level	0.024	−0.055	−0.188	−0.047	−0.097	−0.055
	(0.194)	(0.281)	(0.116)	(0.050)	(0.063)	(0.057)
Monthly income	0.105	0.180	−0.097	−0.122*	−0.172*	−0.118
	(0.221)	(0.322)	(0.132)	(0.056)	(0.073)	(0.065)
Used a calculator	−1.778	−0.469	−1.014	0.762	1.211	0.333
	(2.282)	(3.449)	(1.185)	(0.581)	(0.780)	(0.581)
Mental calculation capacity	−0.532	−0.864	0.166	−0.112	−0.090	−0.027
	(0.352)	(0.534)	(0.197)	(0.090)	(0.121)	(0.097)
Number of errors at the comprehension questionnaire	0.392	1.319***	−0.054	0.042	0.249**	−0.049
	(0.252)	(0.390)	(0.139)	(0.064)	(0.088)	(0.068)
Participated before to a search game	−0.189	−2.437	0.032	0.130	−0.490	0.316
	(1.405)	(1.923)	(0.851)	(0.358)	(0.435)	(0.417)
Constant	5.728***	8.683***	3.929***	2.637***	3.498***	1.982***
	(1.365)	(1.928)	(0.852)	(0.347)	(0.436)	(0.407)
Observations	8,000	4,160	3,340	8,000	4,160	3,340
Subjects	400	208	167	400	208	167
Log Likelihood	−22,089.470	−12,033.400	−8,269.118	−12,261.140	−6,711.718	−4,710.989
Akaike Inf. Crit.	44,208.950	24,096.790	16,568.240	24,552.280	13,453.440	9,451.978
Bayesian Inf. Crit.	44,313.750	24,191.790	16,659.940	24,657.090	13,548.440	9,543.684

Note: *p<0.05; **p<0.01; ***p<0.001

Table B.9: Additional control variables for T2 – total boxes and depth of search

	Dependent variable:					
	Total opened			Attributes per alternative (depth of search)		
	All	Oversearchers	Undersearchers	All	Oversearchers	Undersearchers
Communication of the optimal strategy	0.734	−0.720	1.112**	0.113	−0.369*	0.326*
	(0.405)	(0.524)	(0.341)	(0.131)	(0.167)	(0.164)
Round number	−0.040*	−0.040	−0.034	−0.005	0.001	−0.013
	(0.017)	(0.022)	(0.028)	(0.005)	(0.006)	(0.008)
Male	0.534	0.344	−0.211	0.063	−0.129	0.019
	(0.407)	(0.522)	(0.329)	(0.131)	(0.167)	(0.158)
Age	0.012	0.006	−0.020	0.0003	−0.001	−0.007
	(0.013)	(0.016)	(0.011)	(0.004)	(0.005)	(0.006)
Education level	0.030	0.030	−0.133	−0.049	−0.081	−0.031
	(0.166)	(0.208)	(0.141)	(0.054)	(0.066)	(0.068)
Monthly income	−0.423*	−0.109	−0.271	−0.239***	−0.190*	−0.224**
	(0.199)	(0.279)	(0.150)	(0.064)	(0.089)	(0.072)
Constant	5.636***	7.947***	5.248***	2.610***	3.422***	2.381***
	(0.939)	(1.160)	(0.878)	(0.299)	(0.365)	(0.401)
Observations	5,280	2,840	2,060	5,280	2,840	2,060
Subjects (errors <= 1)	264	142	103	264	142	136
Log Likelihood	−14,306.380	−8,006.227	−5,016.178	−7,545.446	−4,212.867	−2,787.582
Akaike Inf. Crit.	28,634.760	16,034.450	10,054.360	15,112.890	8,447.734	5,597.165
Bayesian Inf. Crit.	28,707.050	16,099.920	10,116.290	15,185.180	8,513.201	5,659.100

Note: *p<0.05; **p<0.01; ***p<0.001

Table B.10: Sub-sample estimation for T2 – total boxes and depth of search

	Dependent variable:								
	Total Score (in %)			Score Search rule 1 (in %)			Score Search rule 2 (in %)		
	All	Oversearchers	Undersearchers	All	Oversearchers	Undersearchers	All	Oversearchers	Undersearchers
Communication of the optimal strategy	1.929	5.561**	−1.750	−0.291	6.138***	−4.064	5.231	9.428*	−3.755
	(1.688)	(1.904)	(3.112)	(1.650)	(1.758)	(2.954)	(3.678)	(4.743)	(6.289)
Round number	0.105*	0.179**	0.019	0.075	0.085	0.097	0.133	0.114	0.086
	(0.051)	(0.062)	(0.093)	(0.061)	(0.068)	(0.118)	(0.084)	(0.120)	(0.123)
Male	1.692	−0.425	5.129	1.303	−1.635	5.251	−0.155	−0.373	2.910
	(1.709)	(1.922)	(3.059)	(1.671)	(1.774)	(2.908)	(3.724)	(4.787)	(6.166)
Age	−0.080	−0.050	−0.091	−0.070	−0.072	−0.034	−0.070	−0.026	−0.090
	(0.055)	(0.058)	(0.102)	(0.053)	(0.054)	(0.097)	(0.119)	(0.145)	(0.207)
Education level	1.437*	1.785*	1.965	0.970	1.302	2.116	1.005	0.496	1.952
	(0.716)	(0.781)	(1.356)	(0.701)	(0.721)	(1.291)	(1.561)	(1.945)	(2.731)
Monthly income	−1.418	−1.673	−1.036	−0.929	−0.156	−2.174	−2.073	−3.402	0.484
	(0.825)	(0.895)	(1.543)	(0.807)	(0.826)	(1.469)	(1.796)	(2.229)	(3.104)
Used a calculator	−0.125	6.209	−12.778	2.391	0.200	−1.262	−22.667	−1.896	−49.755
	(8.405)	(9.610)	(13.835)	(8.239)	(8.880)	(13.173)	(18.302)	(23.899)	(27.912)
Mental calculation capacity	3.232*	2.828	2.220	3.331**	2.950*	1.472	2.039	2.345	2.659
	(1.306)	(1.485)	(2.299)	(1.277)	(1.371)	(2.184)	(2.848)	(3.700)	(4.642)
Number of errors at the comprehension questionnaire	−4.575***	−5.262***	−3.825*	−3.720***	−4.617***	−4.222**	−6.524**	−8.072**	−1.941
	(0.932)	(1.085)	(1.637)	(0.913)	(1.002)	(1.560)	(2.026)	(2.703)	(3.282)
Participated before to a search game	4.233	3.301	9.801	3.406	3.573	9.225	8.445	6.024	14.529
	(5.121)	(5.346)	(9.771)	(4.994)	(4.935)	(9.207)	(11.215)	(13.318)	(19.980)
Constant	25.598***	20.934***	29.657**	38.201***	31.057***	46.213***	68.719***	69.526***	57.374**
	(5.021)	(5.345)	(9.512)	(4.917)	(4.956)	(9.046)	(10.939)	(13.301)	(19.191)
Observations	7,205	4,110	2,755	7,205	4,110	2,755	7,205	4,110	2,755
Subjects	393	208	167	393	208	167	393	208	167
Log Likelihood	−31,992.190	−17,571.450	−12,724.060	−33,585.310	−18,502.400	−13,306.610	−31,757.930	−18,118.710	−12,178.220
Akaike Inf. Crit.	64,014.370	35,172.910	25,478.130	67,200.610	37,034.790	26,643.210	63,545.860	36,267.420	24,386.430
Bayesian Inf. Crit.	64,117.610	35,267.720	25,566.940	67,303.850	37,129.610	26,732.030	63,649.100	36,362.240	24,475.250

Note: *p<0.05; **p<0.01; ***p<0.001

Table B.11: Additional control variables for T2 – score

	Dependent variable:								
	Total Score (in %)			Score Search rule 1 (in %)			Score Search rule 2 (in %)		
	All	Oversearchers	Undersearchers	All	Oversearchers	Undersearchers	All	Oversearchers	Undersearchers
Communication of the optimal strategy	2.703	5.957*	−0.047	0.622	6.568**	−3.116	7.126	10.147	1.284
	(2.230)	(2.588)	(4.142)	(2.126)	(2.279)	(3.955)	(4.628)	(5.859)	(8.235)
Round number	0.077	0.189**	−0.064	0.011	0.103	−0.087	0.139	0.171	0.053
	(0.062)	(0.072)	(0.121)	(0.074)	(0.081)	(0.151)	(0.093)	(0.133)	(0.137)
Male	3.628	1.503	7.649	2.705	−0.104	8.482*	1.505	1.592	3.819
	(2.244)	(2.576)	(4.024)	(2.141)	(2.268)	(3.853)	(4.655)	(5.832)	(7.956)
Age	−0.138	−0.008	−0.294*	−0.105	−0.047	−0.070	−0.082	0.097	−0.459
	(0.072)	(0.079)	(0.140)	(0.069)	(0.069)	(0.134)	(0.150)	(0.178)	(0.277)
Education level	1.332	2.230*	0.750	1.520	1.823*	2.190	0.292	1.534	−1.492
	(0.909)	(1.026)	(1.728)	(0.868)	(0.904)	(1.658)	(1.884)	(2.323)	(3.406)
Monthly income	−1.294	−1.506	−2.213	−0.891	−0.381	−2.588	−1.049	−2.231	−0.179
	(1.110)	(1.376)	(1.831)	(1.060)	(1.212)	(1.754)	(2.296)	(3.116)	(3.618)
Constant	31.951***	19.221***	46.507***	43.101***	31.352***	49.229***	67.417***	56.384***	86.723***
	(5.017)	(5.581)	(9.890)	(4.800)	(4.944)	(9.501)	(10.393)	(12.633)	(19.547)
Observations	4,748	2,803	1,717	4,748	2,803	1,717	4,748	2,803	1,717
Subjects (errors <= 1)	257	142	103	257	142	103	257	142	103
Log Likelihood	−21,148.060	−12,036.540	−7,970.748	−22,122.810	−12,633.370	−8,308.486	−20,632.540	−12,224.110	−7,490.342
Akaike Inf. Crit.	42,318.120	24,095.080	15,963.500	44,267.610	25,288.750	16,638.970	41,287.070	24,470.210	15,002.680
Bayesian Inf. Crit.	42,389.250	24,160.400	16,023.430	44,338.730	25,354.070	16,698.900	41,358.190	24,535.540	15,062.610

Note: *p<0.05; **p<0.01; ***p<0.001

Table B.12: Sub-sample estimations for T2 – score

B.8. Learning over time

	Dependent variable:	
	Absolute distance to optimal total opened	Absolute distance to optimal depth of search
Communication of the optimal strategy	0.098 (0.587)	0.011 (0.129)
Oversearcher	1.975*** (0.543)	0.241* (0.119)
Late rounds	−0.102 (0.292)	−0.023 (0.071)
T2*Oversearcher	−2.394** (0.786)	−0.279 (0.172)
Oversearcher*Late rounds	−0.881* (0.433)	−0.022 (0.106)
T2*Late rounds	0.383 (0.467)	0.031 (0.115)
T2*Oversearcher*Late rounds	0.416 (0.625)	−0.028 (0.153)
Male	0.217 (0.317)	−0.029 (0.072)
Age	0.015 (0.010)	−0.001 (0.002)
Education level	0.061 (0.135)	−0.027 (0.030)
Monthly income	0.146 (0.154)	−0.051 (0.035)
Constant	4.433*** (0.754)	1.787*** (0.169)
Observations	7,500	7,500
Subjects	375	375
Log Likelihood	−24,428.840	−13,837.360
Akaike Inf. Crit.	48,889.680	27,706.720
Bayesian Inf. Crit.	49,000.440	27,817.490

Note: *p<0.05; **p<0.01; ***p<0.001

Table B.13: Interaction effects of learning behavior

B.9. Additional figures

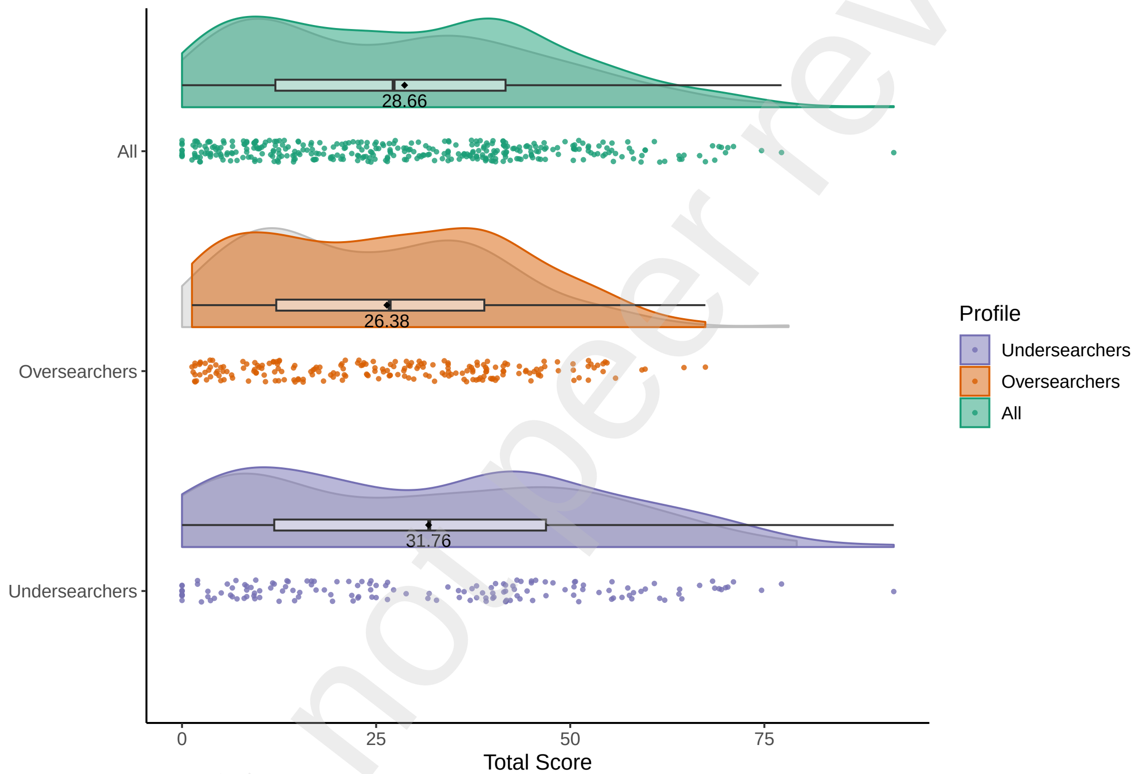

Figure B.1: Evolution of the distributions of the scores for the different profiles
Note: the shaded distribution represents the first ten rounds. The colored distributions, the boxplots, and the means represent the last ten rounds.

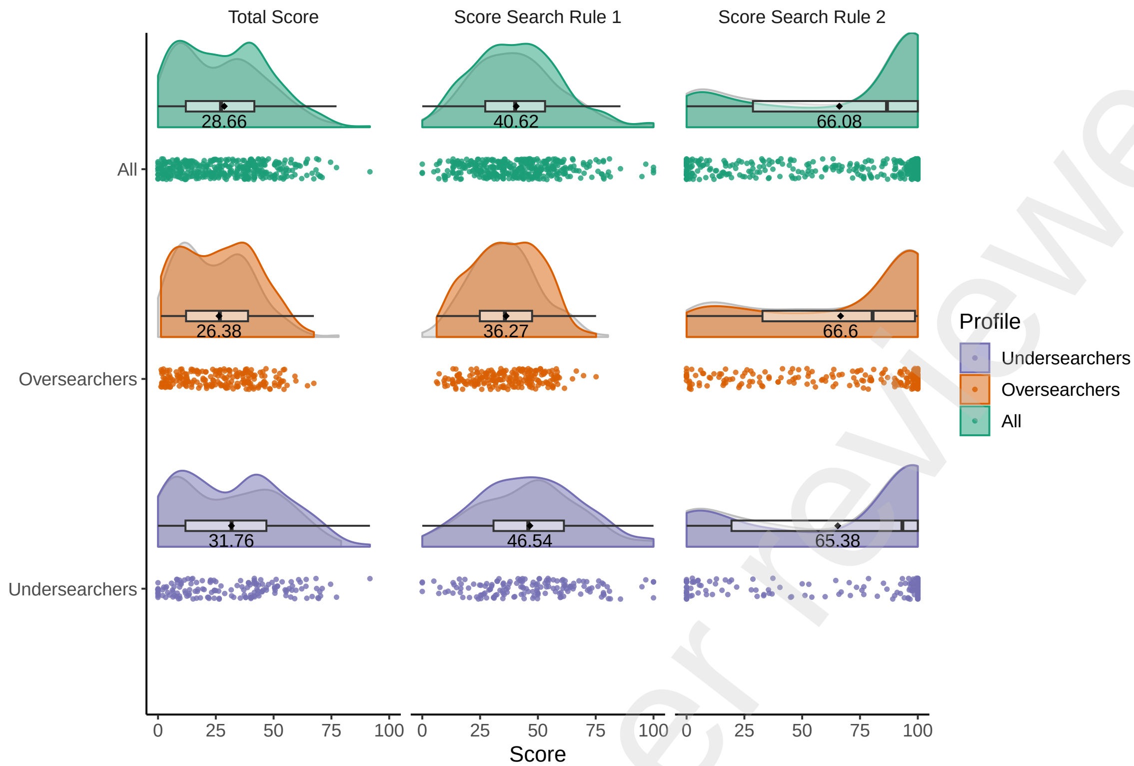

Figure B.2: Evolution of the distributions of the different scores for the different profiles
Note: the shaded distribution represents the first ten rounds. The colored distributions, the boxplots, and the means represent the last ten rounds.

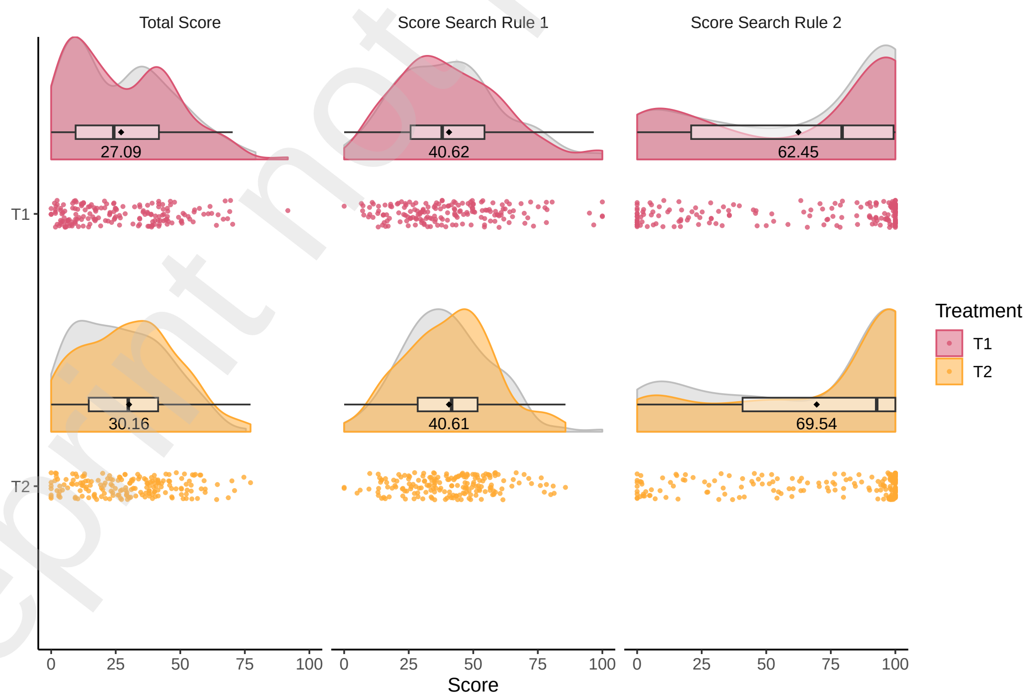

Figure B.3: Evolution of the distributions of the scores in the two treatments
Note: the shaded distribution represents the first ten rounds. The colored distributions, the boxplots, and the means represent the last ten rounds.

References

Bartoš V., Bauer M., Chytilová J., Matějka F., 2016. Attention discrimination: Theory and field experiments with monitoring information acquisition. American Economic Review 106, 1437–75.

Bearden J.N., Connolly T., 2008. On optimal satisficing: how simple policies can achieve excellent results. In: Decision modeling and behavior in complex and uncertain environments, 79–97. Springer.

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

Bouhlel I., Chessa M., Festré A., Guerci E., 2022. When to stop searching in a highly uncertain world? a theoretical and experimental investigation of "two-way" sequential search tasks. Journal of Economic Behavior & Organization 203, 80–92.

Brunyé T.T., Drew T., Weaver D.L., Elmore J.G., 2019. A review of eye tracking for understanding and improving diagnostic interpretation. Cognitive research: principles and implications 4, 1–16.

Camerer C.F., Johnson E.J., Rymon T., Sen S., 1993. Cognition and framing in sequential bargaining for gains and losses. MIT Press.

Caplin A., Dean M., 2015. Revealed preference, rational inattention, and costly information acquisition. American Economic Review 105, 2183–2203.

Caplin A., Dean M., Leahy J., 2019. Rational inattention, optimal consideration sets, and stochastic choice. The Review of Economic Studies 86, 1061–1094.

Chabris C.F., Morris C.L., Taubinsky D., Laibson D., Schuldt J.P., 2009. The allocation of time in decision-making. Journal of the European Economic Association 7, 628–637.

Chen D.L., Schonger M., Wickens C., 2016. otree—an open-source platform for laboratory, online, and field experiments. Journal of Behavioral and Experimental Finance 9, 88–97.

Chetty R., Looney A., Kroft K., 2009. Salience and taxation: Theory and evidence. American economic review 99, 1145–77.

Civelli A., Deck C., Tutino A., 2022. Attention and choices with multiple states and actions: A laboratory experiment. Journal of Economic Behavior & Organization 199, 86–102.

Costa-Gomes M., Crawford V.P., Broseta B., 2001. Cognition and behavior in normal-form games: An experimental study. Econometrica 69, 1193–1235.

DellaVigna S., 2009. Psychology and economics: Evidence from the field. Journal of Economic literature 47, 315–72.

Descamps A., Massoni S., Page L., 2022. Learning to hesitate. Experimental Economics 25, 359–383.

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

Eyal P., David R., Andrew G., Zak E., Ekaterina D., 2021. Data quality of platforms and panels for online behavioral research. Behavior Research Methods 1–20.

Festré A., Garrouste P., 2015. The 'economics of attention': A history of economic thought perspective. Œconomia. History, Methodology, Philosophy 3–36.

Gabaix X., 2019. Behavioral inattention. In: Handbook of Behavioral Economics: Applications and Foundations 1, volume 2, 261–343. Elsevier.

Gabaix X., Laibson D., Moloche G., Weinberg S., 2006. Costly information acquisition: Experimental analysis of a boundedly rational model. American Economic Review 96, 1043–1068.

Heinrich T., Arya B., Haering A., Horak S., 2022. Costly information acquisition: The influence of stakeholder earnings. Journal of Economic Psychology 90, 102504.

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

Lacetera N., Pope D.G., Sydnor J.R., 2012. Heuristic thinking and limited attention in the car market. American Economic Review 102, 2206–36.

Lieder F., Griffiths T.L., 2020. Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. Behavioral and Brain Sciences 43.

Mackowiak B., Matejka F., Wiederholt M., 2021. Rational inattention: A review .

Manzini P., Mariotti M., 2014. Stochastic choice and consideration sets. Econometrica 82, 1153–1176.

Mata A., Vaz A., Mendonça B., 2022. Deliberate ignorance in moral dilemmas: Protecting judgment from conflicting information. Journal of Economic Psychology 90, 102523.

Mischkowski D., Glöckner A., Lewisch P., 2021. Information search, coherence effects, and their interplay in legal decision making. Journal of Economic Psychology 87, 102445.

Momsen K., Ohndorf M., 2022. Information avoidance, selective exposure, and fake (?) news: Theory and experimental evidence on green consumption. Journal of Economic Psychology 88, 102457.

Payne J.W., Payne J.W., Bettman J.R., Johnson E.J., 1993. The adaptive decision maker. Cambridge university press.

Peer E., Brandimarte L., Samat S., Acquisti A., 2017. Beyond the turk: Alternative platforms for crowdsourcing behavioral research. Journal of Experimental Social Psychology 70, 153–163.

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

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

Rieger M.O., Wang M., Hausmann D., 2021. Pre-decisional information acquisition: Why do we pay too much for information? brief report. Journal of Economic Psychology 86, 102411.

Sanjurjo A., 2017. Search with multiple attributes: Theory and empirics. Games and Economic Behavior 104, 535–562.

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

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

Sims C.A., 2003. Implications of rational inattention. Journal of monetary Economics 50, 665–690.

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

Taubinsky D., Rees-Jones A., 2018. Attention variation and welfare: theory and evidence from a tax salience experiment. The Review of Economic Studies 85, 2462–2496.

Tinghög G., Barrafrem K., Västfjäll D., 2022. The good, bad and ugly of information (un) processing; homo economicus, homo heuristicus and homo ignorans. Journal of Economic Psychology 102574–102574.

Van Nieuwerburgh S., Veldkamp L., 2009. Information immobility and the home bias puzzle. The Journal of Finance 64, 1187–1215.

Woodford M., 2012. Inattentive valuation and reference-dependent choice .

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