Expert Systems With Applications 191 (2022) 116352

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

An information retrieval benchmarking model of satisficing and impatient users' behavior in online search environments

Debora Di Caprio [a,*,1], Francisco J. Santos-Arteaga [b,2], Madjid Tavana [c,d,3]

[a] Department of Economics and Management, University of Trento, Trento, Italy
[b] Departamento de Análisis Económico y Economía Cuantitativa, Universidad Complutense de Madrid, Madrid, Spain
[c] Business Systems and Analytics Department, Distinguished Chair of Business Analytics, La Salle University, Philadelphia, USA
[d] Business Information Systems Department, Faculty of Business Administration and Economics, University of Paderborn, Paderborn, Germany

ARTICLE INFO

Keywords:
Information retrieval
Satisficing
Benchmarking
Impatience
Click-through rate
Online search

ABSTRACT

This study analyzes the effects that the position of the alternatives ranked by a search engine and the relative impatience of users have on their information retrieval behavior. We design a stochastic information retrieval algorithm calibrated to mimic the click-through rates (CTRs) of users observed in real-life environments. We introduce two versions of the mimicking algorithm designed to demonstrate the importance of impatience as a determinant of CTRs conditioned by the alternatives' ranking position. The first version assumes that users proceed sequentially through the ranking until they find an alternative satisficing their expectations. Once they find a satisficing alternative, they continue retrieving information until they observe an alternative that violates their expectations. The second version increases users' impatience, who stop retrieving information as soon as an alternative does not satisfy their expectations – even if it is the top-ranked one. All three algorithmic structures are sufficiently malleable to incorporate any potential modification to users' beliefs and preferences. We simulate sets of 1,000,000 queries to illustrate how the CTRs of the top three ranked alternatives remain stable as users grow impatient, with differences widening as growingly impatient users proceed halfway through the ranking.

1. Introduction

Users tend to evaluate the alternatives obtained from a search in the order provided by the online search engine (Epstein & Robertson, 2015; Gao & Shah, 2020; Luo et al., 2011). Eye-tracking technology has validated this feature as well as the biased focus of users towards the highest-ranked alternatives (Lewandowski & Kammerer, 2020; Lorigo et al., 2008). Indeed, the first two alternatives composing the ranking receive a disproportionate number of clicks compared to the remaining ones within the first page of search results (Chitika, 2013; Dean, 2019).

The formalization of the information retrieval behavior observed through standard utility approaches must deal with the cognitive limits of users (Gupta et al., 2018; Lieder & Griffiths, 2020), whose behavior cannot be based on the almost four million permutations that can be computed from the ten results composing the initial page delivered by the engine (Basu, 2018; Victorelli et al., 2020). We are therefore left with the order implicit in the ranking provided by the engine as the only guideline available to replicate the information retrieval behavior of users (European Commission, 2016).

The satisficing approach to information retrieval and user behavior has gained considerable attention in later years, particularly within experimental settings. The empirical literature comparing the maximization and satisficing approaches to information retrieval emphasizes the difficulties faced when eliciting the utility derived from the outcomes of the search process (Misuraca & Fasolo, 2018). These analyses focus particularly on students, who are also found to generally follow a satisficing approach when identifying optimal information sources (List & Alexander, 2017). Note that, though not explicitly, the intuition behind the satisficing approach implies a certain degree of impatience from the user, who may decide to conclude the retrieval process as soon as he finds a suitable alternative.

The decision-theoretical literature has managed to identify the main

* Corresponding author.
E-mail addresses: debora.dicaprio@unitn.it (D. Di Caprio), fransant@ucm.es (F.J. Santos-Arteaga), tavana@lasalle.edu (M. Tavana).
[1] https://orcid.org/0000-0002-6900-3977.
[2] https://orcid.org/0000-0003-2385-4781.
[3] https://orcid.org/0000-0003-2017-1723.

https://doi.org/10.1016/j.eswa.2021.116352
Received 6 April 2021; Received in revised form 28 November 2021; Accepted 28 November 2021
Available online 9 December 2021

characteristics determining the behavior of impatient users (Ghafurian et al., 2020). The complexity of the interactions taking place between users and programs has been widely documented and ranges from reactions to different response times to connections within the emotional domain (Norman & Kirakowski, 2017; Victorelli et al., 2020). Initial studies concluded that the waiting time users were willing to tolerate the download of a Web page when retrieving information was approximately two seconds (Nah, 2004). The importance assigned by users to slow interactions and response times has evolved through time, with their demand for timely information adapting to the new technological paradigm (Lohr, 2012). Recently, impatient users have become the focus of analysis in queuing-related environments, providing fertile ground for the development of potential extensions of the algorithmic framework presented (Bolandifar et al., 2019; Li et al., 2018).

Credibility considerations have also gained considerable relevance, particularly in strategic and medical environments (Machackova & Smahel, 2018). The evaluation of information is a complex procedure where credibility is determined by the characteristics of users, contents, and information sources, encompassing also the task motivating the search and its relative difficulty (Lee & Pang, 2018). For instance, cognitive scientists have highlighted the reliance of users on information scent when evaluating the alignment of alternatives with their preferences (Karanam et al., 2016; Ong et al., 2017).

Regarding applicability, tourism research is one of the leading academic fields analyzing information retrieval processes conditioned by the output delivered by different search engines and recommender systems. The empirical findings from this research area describe decision-makers (DMs) overwhelmed by the amount of information that they must assimilate and process (Zillinger, 2020). Thus, DMs must rely on search engines without understanding the algorithmic mechanism delivering the results and even their own search strategies (Pirolli, 2018). As a result, this branch of the literature has emphasized the considerable confusion that exists regarding the actual search strategies of users (Lu & Gursoy, 2015).

1.1. Research objectives

The main contribution of the current paper is the design of a series of algorithmic benchmarks allowing researchers to extrapolate the behavior of DMs when facing different types of search frictions. The algorithms are built on basic behavioral assumptions so that further modifications to the incentives driving the retrieval process can be implemented. We aim at providing a reference framework of analysis for empirical studies when determining the consequences from modifications to the information retrieval incentives of DMs.

It is important to emphasize what is not being analyzed by the algorithms. The structural complexity of the algorithms contrasts with the basic characterization of the decision nodes. That is, each decision node is based on a simple command stating that

$$\text{if a random uniform realization} > \text{the cutoff value assigned to the alternative} \tag{1}$$

then the alternative is evaluated. The value of the stochastic realization reflects the characteristics of the alternative being evaluated – as observed by the DM –, which are compared to the subjective preferences defined by the DM as determinants of the cutoff value. The algorithms do not consider

- how the realizations observed follow from the cognitive capacities of DMs;
- how the cutoff values are defined based on the subjective preferences of DMs.

That is, we do not study how these features are determined but benchmark the response of DMs using the behavioral data provided by search engines. The literature dealing with these characterizations extends through different research fields, ranging from empirical decision-making (Doniec et al., 2020; Jankowski et al., 2016) and psychology (Khamitov et al., 2019) to cognitive sciences (Dou et al., 2010). Each of these areas has provided ample evidence regarding the factors that determine the alignment of the preferences of DMs with the characteristics observed. These factors are quite varied and encompass product features (Lu & Altenbek, 2021; Zhu & Zhang, 2010), and the subjective characteristics of DMs (Lauraéus, T., Saarinen, T., Öörni, A.: Factors affecting consumer satisfaction of online purchase. In: 48th Hawaii International Conference on System Sciences, Kauai, HI, 2015, 2015, 2015; Sadiq et al., 2021; Shafiq et al., 2015), ranging from gender differentials (Bae & Lee, 2011) to cognitive (Bartels & Johnson, 2015; Kimmel, 2012) and psychological frictions (Lerner et al., 2015).

1.1.1. Contribution

We design a benchmark information retrieval algorithm mimicking the click-through rate (CTR) behavior of users when deciding on which alternatives to click from the first page of results displayed by a search engine. The CTR of a given alternative is defined as the number of users who click on the link to the alternative divided by the total number of users performing a search. The decision-tree structure of this benchmark algorithm accounts for the 1,023 binary decision nodes that users may have to consider as they retrieve information from the first page of results and the 1024 final nodes describing the potential evaluation vectors generated through the different retrieval paths that may be followed by DMs. The only assumption imposed on the information retrieval process of users is that they observe alternatives in the order provided by the engine. In this regard, the benchmark algorithm assumes that users consider clicking on the different alternatives with a decreasing probability as they proceed through the ranking delivered by the engine. We equate the probability of clicking on an alternative to the empirical value of the CTR described in Dean (2019) and simulate a total of 1,000,000 queries per configuration to evaluate the ability of the algorithm to replicate the behavior observed.

After illustrating the capacity of the benchmark algorithm to replicate the CTR behavior observed, we analyze the effects from an increase in the impatience of users as they retrieve information on the alternatives provided by the search engine. As emphasized in the previous section, impatience represents a consistent characteristic defining the behavior of online users, a feature increasingly exacerbated at the mobile search level (Google, 2016; Varnali et al., 2012). We define two versions of the initial algorithm to illustrate the importance of impatience as a determinant of CTRs conditioned by the ranking position of the alternatives.

- The first version assumes that users proceed sequentially through the ranking until they find an alternative satisficing their expectations. Once they find a satisficing alternative, DMs continue retrieving information until they observe an alternative that violates their expectations.
- The second version increases the impatience of users, who stop retrieving information as soon as an alternative does not satisfy their expectations – even if it is the top-ranked one.

In all cases, the algorithms must incorporate the whole set of potential branches that may be generated through the information retrieval process. This request is relatively simple when accounting for impatient evaluation structures dealing with a total of 21 nodes, but its complexity increases considerably when dealing with the 2,047 nodes required to model the sequential behavior of the benchmark, i.e., patient, users. The resulting scenarios, including the intermediate satisficing setting – introduced to accommodate the bounded rationality approaches implemented in the economics and managerial literature –, are developed through the next section.

It should be emphasized that our approach differs completely from

the one generally implemented in the systems literature. We do not design an experimental framework whose results are compared to the actual behavior of users (Schneider et al., 2019; Speier-Pero, 2019), a methodology that also represents the standard procedure in the literature on electronic commerce (Sun et al., 2020; Yoo et al., 2016). We define a simulation benchmark determined by the search behavior of users online, whose modifications can be validated through experiments or empirical analyses (Bell & Mgbemena, 2018; Dunke & Nickel, 2020).

That is, the proposed algorithms constitute a malleable structure allowing to simulate a wide variety of behavioral phenomena, providing a benchmark evaluation framework for the systems literature (Hong et al., 2021; Mahony et al., 2016; Zhang et al., 2020). The algorithmic structures proposed also provide a benchmark counterpart to the big data analyses performed to elicit the behavior of users through the implementation of artificial intelligence techniques (Dell'Aversana & Bucciarelli, 2018).

The paper proceeds as follows. Sections 2 and 3 describe the intuition required to formalize the different information retrieval algorithms. The behavior of users under different willingness to search and impatience scenarios is analyzed in Section 4. Section 5 incorporates search frictions to the analysis and presents the main managerial implications. Section 6 concludes and suggests potential extensions.

2. Decision trees as evaluation techniques

Decision trees provide a formal structure sufficiently flexible to analyze most sequential information retrieval processes common to the management and operations research literature (Pei & Hu, 2018; Sagi & Rokach, 2020). Despite this fact, the systems literature has not generally considered the application of decision trees to formalize information retrieval incentives in online evaluation environments.

The apparent simplicity of the corresponding information retrieval process may have led researchers to disregard it as an object of analysis. From a computational perspective, it is not particularly demanding to simulate ten independent random processes and assign a cutoff value to each of them. However, the resulting framework would ignore the interactions across evaluations considered by users, DMs henceforth, when retrieving information. These relationships become particularly evident when defining the different algorithmic structures through the paper, since we must account for the whole set of potential paths that may arise from every decision node.

2.1. Benchmark algorithm: patient scenario

Consider the initial decision faced by a DM when evaluating the first alternative provided by a search engine. The DM must choose between clicking on the alternative to obtain additional information or proceeding with the second alternative. The decision is determined by the information displayed by the engine and its alignment with the subjective preferences of the DM. If both align, the DM clicks on the alternative and browses through the linked website.

The process follows in a similar manner through the rest of the $2^{10}-1$ binary nodes composing the tree, with two paths opening before each decision is made. That is, two outcomes must be considered per node, determined by the alignment of the corresponding observation with the preferences of the DM. The algorithm accounts for the two initial paths, the subsequent four, eight, sixteen, and proceeds through all the potential possibilities until reaching the end of the ranking. The algorithm increases in complexity as different paths arise from each subsequent node, each of them leading to binary choices that will lead to further paths and choices, and so on.

We must remark here an important feature of the retrieval process, implicitly assumed in the observed CTRs. The content of the website clicked by the DM may condition his willingness to continue retrieving information. That is, the link may lead to a website that does not align with his preferences. This misalignment may be assumed to impose frictions on his search process. It could also be assumed that the DM interrupts the search because the link leads to a website that fully aligns with his preferences.

The benchmark algorithm can be modified to incorporate both these features. We can increase the complexity of the algorithm by considering two stages per node, namely, an initial evaluation step and a second one where DMs browse through the link clicked. The subsequent CTRs would be determined by the product of the acceptance probabilities defining the evaluation and browsing stages. The algorithm could then be calibrated to match the behavior observed. We will however focus on the structural properties of the algorithm and set aside any potential extension that could be easily incorporated into the sequential evaluation framework.

A related retrieval structure can be found in the searches performed by academic scholars, who are able to observe several features of a paper, including some lines from the abstract, before clicking on a link. As fully rational agents, scholars may suffer a loss in the confidence subjectively assigned to the rankings – or an increase in impatience – after browsing through papers that are not related to their intended research line.

2.2. Satisficing evaluation algorithm

The satisficing evaluation decision tree presents a less complicated design and formalizes the following information retrieval structure. The DM observes the first alternative and must decide whether or not to click on it. If he does not click, he proceeds with the retrieval process, deciding whether or not to click on any subsequent alternatives. After clicking on an alternative, he continues retrieving information but stops as soon as one of the remaining alternatives does not align with his preferences.

In other words, the DM only considers ending the retrieval process after observing an alternative that satisfies his preferences. The intuition behind this algorithm follows from the bounded rationality hypothesis commonly considered in decision theory and economics (Cristofaro, 2017; Ren & Huang, 2018). Satisficing heuristics are defined to balance cognitive effort and the subsequent accuracy of the decisions being made (Kreye et al., 2012; Richardson, 2017). Thus, once the DM secures a satisficing level, he may decide to interrupt the retrieval process.

2.3. Impatient scenario

The impatient setting is the simplest one from a formal and computational perspective. In this case, the DM interrupts the retrieval process as soon as an alternative does not align with his preferences, even if it is ranked first. In other words, the DM proceeds through the retrieval process as long as the alternatives displayed by the engine align with his preferences. Impatience is a salient characteristic among online users, and their willingness to conclude the search at the very initial stages of the retrieval process has been consistently documented (Fishkin, 2019). The impatient algorithm incorporates these basic features into the analysis.

3. Designing the information retrieval algorithms

The information retrieval algorithms analyzed are intuitively illustrated in Figs. 1 to 3. The sequential evaluation structures described through these figures incorporate each alternative in the order delivered by the search engine. That is, the first element composing each initial node within all figures, denoted by 1°, corresponds to the first element ranked by the search engine. The same intuition applies to the remaining elements composing the nodes until the tenth and final one, denoted by 10°, is reached.

Random uniformly distributed values within [0, 1] are assigned to each alternative displayed on the first page of results per query. These

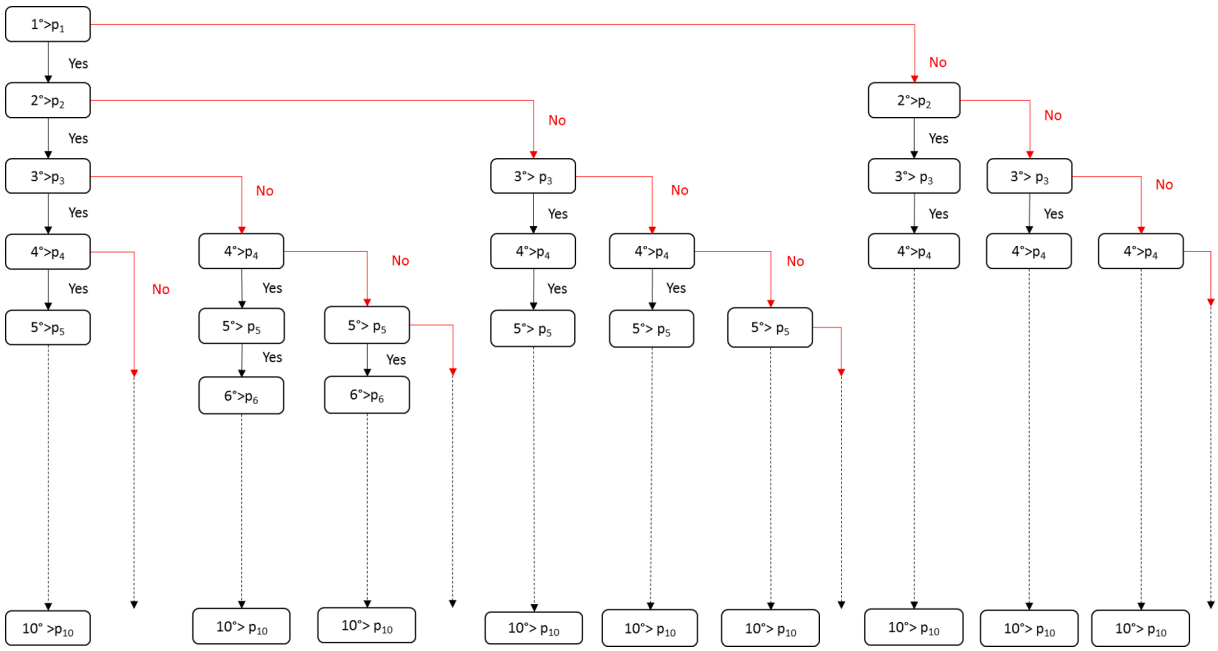


Fig. 1. Structure of the patient evaluation algorithm.

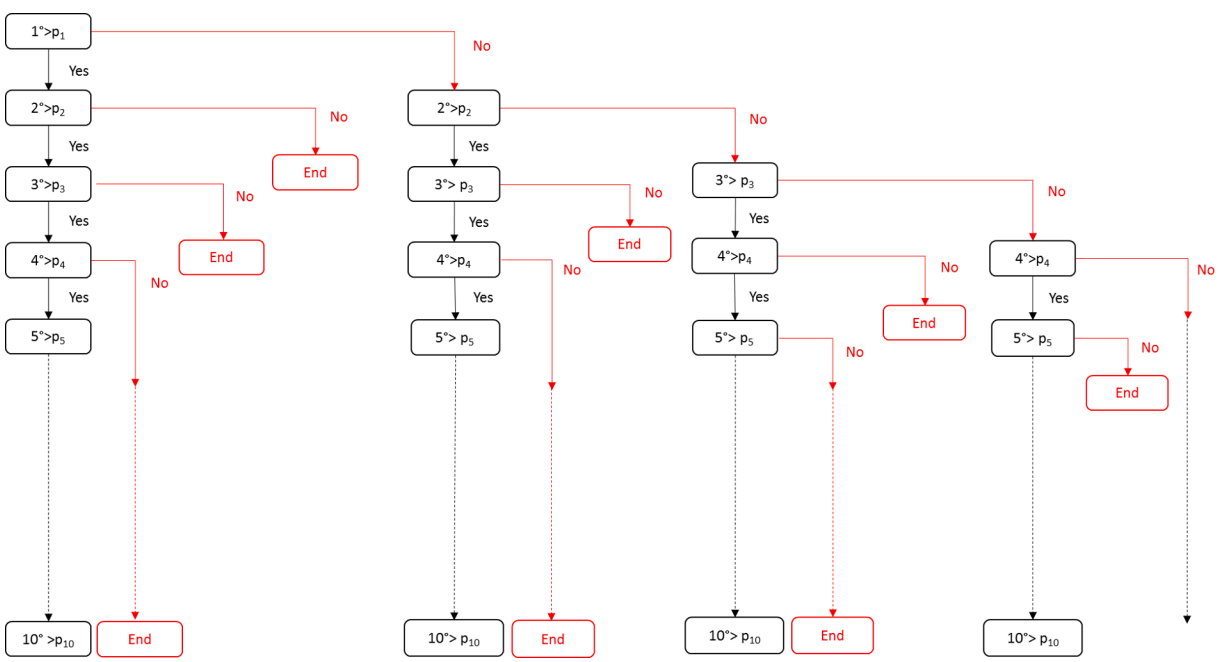


Fig. 2. Structure of the satisficing evaluation algorithm.

values reflect the uncertainty inherent to the evaluation and retrieval processes (Dimoka et al., 2012; Liu et al., 2011). Abusing notation, we represent these random realizations through the ranking positions of the different alternatives, that is, 1°, 2°,…, 10°, within all figures. Similarly, cutoff values, denoted by p_i, $i = 1,\ldots,10$, are assigned to each alternative according to their ranking positions.

The information retrieval process is determined by the value of the realizations assigned to each alternative relative to the corresponding cutoffs.

- If the value of the realization is higher, the user clicks on the alternative, evaluates it, and then proceeds with the next one in the ranking.
- If the value is lower, the user does not click on the alternative, but his decision regarding whether or not to proceed with the next alternative depends on the subjective degree of impatience assumed.

In all cases, each query leads to multiple potential paths depending on the alternatives evaluated by the user and their corresponding ranking positions.

Figs. 1 and 4 present the benchmark algorithmic structure designed to mimic the empirical behavior observed among online users. It is the most complex setting due to the willingness of users to consider the whole set of alternatives displayed by the engine. In this scenario, users evaluate the alternatives whose realizations are higher than their corresponding cutoff values. They do not evaluate those alternatives with

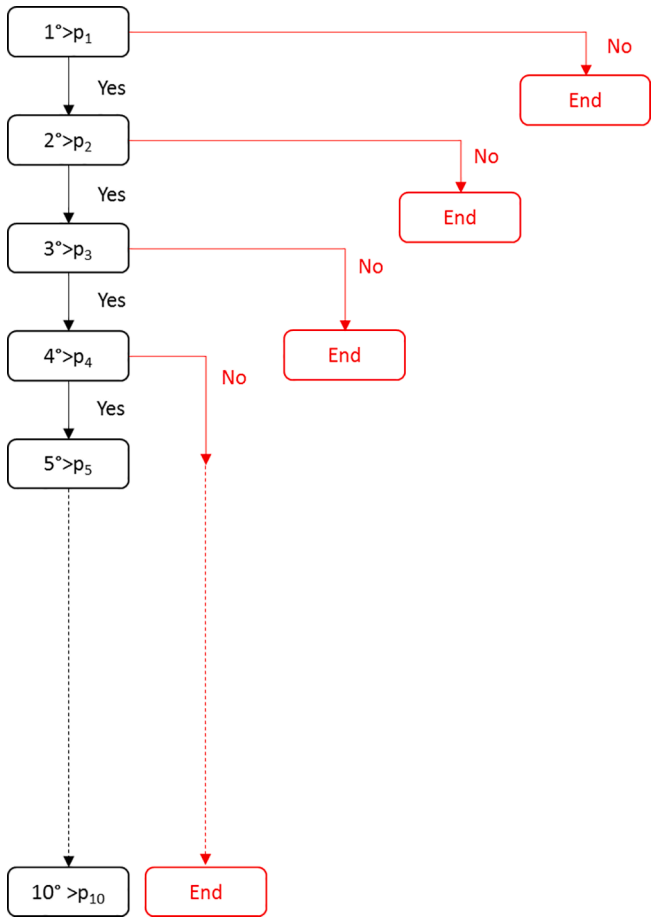


Fig. 3. Structure of the impatient evaluation algorithm.

lower realizations than the cutoff values but proceed through each and every alternative and observe all the realizations.

More precisely, consider Fig. 1. If the first realization defined within the initial node is located above the cutoff value, i.e., $1^\circ > p_1$, the DM evaluates the alternative and proceeds to observe the next one in the ranking through the 'Yes' branch of the tree. On the other hand, if the realization is located below the cutoff value, i.e., $1^\circ < p_1$, the DM does not evaluate the alternative but proceeds to observe the next one composing the ranking through the 'No' branch of the tree. This is the case for each and every alternative described within the initial page of search results.

The intuition defining Fig. 1 relies on the fact that each branch of the tree delivers a vector summarizing the evaluation behavior, i.e., retrieval path, of the DM. For instance,

- The first column composing the tree would lead to a vector consisting of ten alternatives being evaluated, namely, (1, 2, 3, 4, 5, 6, 7, 8, 9, 10).
- Assume now that the first realization is located below the cutoff value, leading the DM through the alternative path arising from the first decision node. In this case, the path corresponds to the seventh column composing the tree.
  - o If all additional alternatives are evaluated, the final vector would be given by (2, 3, 4, 5, 6, 7, 8, 9, 10, 0).
  - o However, if the second alternative is not evaluated, but the remaining ones are, the final vector would be given by (3, 4, 5, 6, 7, 8, 9, 10, 0, 0). The same logic applies to all the potential retrieval paths arising through the tree.

Note that we have described the pages evaluated within the initial entries of the vector, leaving the final ones composed by zeros. It should be emphasized that this notational choice does not have any influence on the results obtained but provides intuition regarding the length of the evaluation path arising from the different potential retrieval processes.

The algorithmic structure defined in Fig. 4 is simplified below, where the fact that the DM proceeds retrieving information independently of the realizations observed has been emphasized. The algorithm must, however, consider the entire set of potential outcomes, a fundamental feature of the benchmark structure when introducing frictions within the search process, as will be illustrated numerically in Section 5.

```
generate random realization
        if the cutoff condition is met
        page is evaluated
        proceed with the next page
                        if next realization meets cutoff condition
                        page is evaluated
                        proceed with the next page
                        ……
                        else
                        page is not evaluated
                        proceed with the next page
                        ….
        else
        page is not evaluated
        proceed with the next page
        ….
```

The structure of the satisficing algorithm scenario described in Figs. 2 and 5 defines the behavior of partially impatient users. These users allow realizations to be lower than the corresponding cutoff values before clicking on an alternative. Once they have clicked, users stop searching after they observe a realization lower than its cutoff value. The intuition behind Fig. 2 is similar to the one provided for Fig. 1. However, in this case, the DM considers ending the retrieval process as soon as he observes a suitable satisficing alternative. Thus, as in the benchmark scenario described in Fig. 1

- If the first realization defined within the initial node is located above the cutoff value, the DM evaluates the alternative and proceeds to observe the next one in the ranking.
  - o After evaluating the first alternative, the DM will consider ending the retrieval process whenever any of the remaining realizations is located below the corresponding cutoff value. That is, if the realization of the second alternative is located below the cutoff value, i.e., $2^\circ < p_2$, the final vector summarizing the retrieval process would be given by (1, 0, 0, 0, 0, 0, 0, 0, 0, 0). In Fig. 1, this result would have led the DM to proceed through the whole ranking, the final vector being given by (1, -, -, -, -, -, -, -, -, 0). The second alternative would not have been evaluated, leading to the final entry of zero, with the remaining values determined by the outcomes arising from the potential evaluation paths.
- If the first realization is located below the cutoff value, the DM does not evaluate the alternative but proceeds to observe the next one composing the ranking. This path leads to the fifth column within Table 2.

```
generate random uniform_i
        if random uniform_i > p_i
        page i is visited and evaluated;
        random uniform_{i+1};
                        if random uniform_{i+1} > p_{i+1}
                        page i+1 is visited and evaluated;
                        random uniform_{i+2};
                                        if random uniform_{i+2} > p_{i+2}
                                        ……
                        else
                        page i+1 is not visited;
                        if random uniform_{i+2} > p_{i+2}
                        page i+2 is visited and evaluated;
                        random uniform_{i+3};
                                        if random uniform_{i+3} > p_{i+3}
                                        page i+3 is visited and evaluated;
                                        random uniform_{i+4};
                                        ….

                                        else
                                        page i+3 is not visited;
                                        random uniform_{i+4};
                                        if random uniform_{i+4} > p_{i+4}
                                        page i+4 is visited and evaluated;
                                        random uniform_{i+5};
                                        ….
                        else
                        page i+2 is not visited;
                        if random uniform_{i+3} > p_{i+3}
                        page i+3 is visited and evaluated;
                        random uniform_{i+4};
                        ….
        else
        page i is not visited;
        if random uniform_{i+1} > p_{i+1}
        page i+1 is visited and evaluated;
        random uniform_{i+2};
                if random uniform_{i+2} > p_{i+2}
                page i+2 is visited and evaluated;
                random uniform_{i+2};
                ….
                        end
                                        end
                        end
        end
end
```

Fig. 4. Codified structure of the patient evaluation algorithm.

- If the next alternative within this path is not evaluated, i.e., $2^\circ < p_2$, the DM proceeds with the third one, as described in the eight column of Table 2.
- However, if the second alternative is evaluated, i.e., $2^\circ > p_2$, then the DM proceeds with the third one within the same path and considers ending the retrieval process after observing a realization below the corresponding cutoff value.
  - If the third alternative is evaluated, i.e., $3^\circ > p_3$, the final vector would be given by (2, 3, -, -, -, -, -, -, 0). The first alternative, which was not evaluated by the DM, leads to the final entry of zero. The remaining values are determined by the outcomes arising from the potential evaluation paths.
  - If the third alternative is not evaluated, i.e., $3^\circ < p_3$, then the final vector would be given by (2, 0, 0, 0, 0, 0, 0, 0, 0, 0). The remaining potential paths follow a similar evaluation pattern.

The algorithmic structure described in Fig. 5 is simplified below, where the impatience that results from satisficing arises as soon as the DM observes an alternative aligning with his preferences. Note how the alignment leading to the second 'if' command implies that as soon as an 'else' is observed, the DM ends the search process. However, if preferences do not align within the first 'if' command, then the DM proceeds with the next alternative, allowing for further potential misalignments until an 'if' command aligns with his preferences.

```
generate random uniform_i
        if random uniform_i > p_i
        page i is visited and evaluated;
        random uniform_i+1;
                if random uniform_i+1 > p_i+1
                page i+1 is visited and evaluated;
                random uniform_i+2;
                        if random uniform_i+2 > p_i+2
                        ……
                else
                page i+1 is not visited;
                end
                        else
                        page i+2 is not visited;
                        end
        else
        page i is not visited;
        if random uniform_i+1 > p_i+1
        page i+1 is visited and evaluated;
        random uniform_i+2;
                if random uniform_i+2 > p_i+2
                page i+2 is visited and evaluated;
                random uniform_i+2;
                ….
                else
                page i+2 is not visited;
                end
        else
        page i+2 is not visited;
        if random uniform_i+3 > p_i+3
        page i+3 is visited and evaluated;
        random uniform_i+4;
        ….
        end
        end
end
```

Fig. 5. Codified structure of the satisficing evaluation algorithm.

```
generate random realization
        if the cutoff condition is met
        page is evaluated
        proceed with the next page
                if next realization meets cutoff condition
                page is evaluated
                proceed with the next page
                ……
                else
                page is not evaluated
                end search
        else
        page is not evaluated
        proceed with the next page
        ….
```

The impatient users represented in Figs. 3 and 6 stop searching as soon as the realization assigned to an alternative is lower than the corresponding cutoff value. We must highlight here the main difference with respect to the satisficing scenario, where users allow realizations to be lower than the cutoff values before clicking on an alternative – and starting to consider ending the search –. Fig. 3 represents the simplest evaluation path, composed by a main branch of the tree and the potential decisions to end the retrieval process as soon as a realization is located below the cutoff value. Similarly, to the previous cases

- If the first realization defined within the initial node is located above the cutoff value, the DM evaluates the alternative and proceeds to observe the next one in the ranking.
  - o If the second alternative is located below the cutoff value, the DM ends the retrieval process, leading to the final vector (1, 0, 0, 0, 0, 0, 0, 0, 0, 0).
  - o If the second alternative is evaluated, the DM proceeds with the third one, delivering the final vector (1, 2, -, -, -, -, -, -, -, -), determined by the remaining realizations observed through the evaluation path.
- On the other hand, if the first realization is located below the cutoff value, the DM does not evaluate the alternative and ends the retrieval process, leading to the following final vector (0, 0, 0, 0, 0, 0, 0, 0, 0, 0).

```
generate random uniform_i
        if random uniform_i > p_i
        page i is visited and evaluated;
        random uniform_i+1;
                        if random uniform_i+1 > p_i+1
                        page i+1 is visited and evaluated;
                        random uniform_i+2;
                                        if random uniform_i+2 > p_i+2
                                        ......
                        else
                        page i+1 is not visited;
                        end
                                        else
                                        page i+2 is not visited;
                                        end
        else
        page i is not visited;
        end
end
```

Fig. 6. Codified structure of the impatient evaluation algorithm.

The algorithmic structure defined in Fig. 6 is summarized below, highlighting the fact that any observation that does not align with the preferences of the DM ends the retrieval process.

```
generate random realization
        if the cutoff condition is met
        page is evaluated
        proceed with the next page
                        if next realization meets cutoff condition
                        page is evaluated
                        proceed with the next page
                        ......
                        else
                        page is not evaluated
                        end search
        else
        page is not evaluated
        end search
```

The numerical analysis focuses on the ability of the mimicking algorithm representing patient users to replicate the empirical behavior observed, while the satisficing and impatient algorithms illustrate the consequences from decreasing the willingness of users to continue retrieving information and evaluating alternatives. The algorithms have been coded using Matlab software. We have provided the Matlab code corresponding to the benchmark algorithm as an online supplementary appendix to the manuscript.

4. Impatience and click through rates

The intuition on which the mimicking algorithm is built follows directly from the sequential behavior observed among online users. Dean (2019) analyzed a sample of five million queries and computed the CTRs of the organic results ranked within the first page of Google. The second column of Table 1 describes the CTRs reported by Dean (2019), which are then compared with those obtained from the different information retrieval algorithms determined by the relative impatience of users. The remaining columns composing Table 1 describe the CTRs obtained when simulating the behavior of DMs conditioned by the probabilities defining different types of evaluation frameworks.

More precisely, we have introduced several evaluation frameworks to study the CTR behavior of users as their impatience increases, determining their willingness to retrieve information from the alternatives displayed. The evaluation frameworks presented in Table 1 are based on these sets of cutoff values

Table 1
Comparing CTRs across different scenarios.

Framework		$[p_1,\ldots, p_{10}] = [68,\ldots, 97]$			$[p_1,\ldots, p_{10}] = [50,\ldots, 95]$			$[p_1,\ldots, p_{10}] = [10,\ldots, 95]$		
CTR	Dean (2019)	Patient	Satisficing	Impatient	Patient	Satisficing	Impatient	Patient	Satisficing	Impatient
Avg	–	1.19	0.90	0.42	2.75	1.55	0.86	4.55	2.89	2.66
1	31.7	32.02	32	32.06	50.07	50.06	49.98	90.00	89.97	90.03
2	24.7	24.98	24.96	7.97	45.07	45.06	22.50	80.00	79.96	72.04
3	18.7	19.03	14.42	1.54	40.05	29.1	8.98	69.96	57.34	50.39
4	13.6	13.98	7.84	0.22	34.96	15.93	3.14	60.07	34.74	30.20
5	9.5	9.99	4.32	0.02	30.01	7.98	0.95	50.00	17.48	15.11
6	6.2	6.02	2.18	0	25.03	3.86	0.24	39.97	7.06	6.06
7	4.1	3.99	1.29	0	20.05	1.88	0.05	30.01	2.15	1.80
8	3.1	2.98	0.89	0	15.04	0.95	0	19.96	0.43	0.35
9	3	3	0.88	0	10.02	0.47	0	10.02	0.05	0.03
10	3	3	0.83	0	4.96	0.2	0	5.01	0	0

All numbers indicate percentages except the third row, which refers to the average number of pages clicked per search.

$$[p_1, p_2, p_3, p_4, p_5, p_6, p_7, p_8, p_9, p_{10}] = \begin{cases} [0.68, 0.75, 0.81, 0.86, 0.90, 0.94, 0.96, 0.97, 0.97, 0.97]; \\ [0.50, 0.55, 0.60, 0.65, 0.70, 0.75, 0.80, 0.85, 0.90, 0.95]; \\ [0.10, 0.20, 0.30, 0.40, 0.50, 0.60, 0.70, 0.80, 0.90, 0.95], \end{cases}$$

Each corresponding column within the table reports the CTRs derived from simulating 1,000,000 runs encompassing 10 alternatives. That is, we have simulated one million queries for each type of DM within every evaluation framework. The number of runs has been chosen to illustrate the convergence of the algorithms to the cutoff values defined within their respective benchmarks.

In particular, the 'Patient' column within the $[p_1,\ldots, p_{10}] = [68,\ldots, 97]$ framework presents the calibration benchmark based on the results reported by Dean (2019). That is, the patient DMs within $[p_1,\ldots, p_{10}] = [68,\ldots, 97]$ have been designed to mimic the CTRs observed empirically. The 'Satisficing' and 'Impatient' columns define evaluation scenarios determined by the implicit decrease in the probability of clicking as increasingly impatient users proceed through the ranking. Note how the satisficing DMs exhibit a decrease in their CTRs after the second alternative. As described in the previous section, the first two observations trigger the same retrieval behavior from patient and satisficing DMs, delivering almost identical CTRs. Impatient DMs display a more radical retrieval behavior and only preserve the same CTR for the first observation. Indeed, these DMs do not consider evaluating two thirds of the alternatives composing the ranking.

The $[p_1,\ldots, p_{10}] = [50,\ldots, 95]$ and $[p_1,\ldots, p_{10}] = [10,\ldots, 95]$ frameworks represent simple benchmarks designed to illustrate the effects of impatience on the CTRs of users as the initial probability of clicking on an alternative increases. That is, the evaluation probabilities that follow from the empirical framework defined by Dean (2019) are quite low, particularly when dealing with alternatives located within the lower half of the ranking. As a result, satisficing and impatient agents barely retrieve information from these latter alternatives. These additional frameworks introduce higher evaluation probabilities throughout the alternatives, allowing us to smooth the decrease of the CTRs as DMs proceed through the lower half of the ranking. As can be observed in Table 1, the relations and patterns obtained are very similar across evaluation scenarios.

The $[p_1,\ldots, p_{10}] = [50,\ldots, 95]$ framework delivers the same trends as $[p_1,\ldots, p_{10}] = [68,\ldots, 97]$, based on higher evaluation probabilities. CTRs are larger among DMs, as well as the average number of pages clicked per query. This latter variable, denoted 'Avg', is described in the third row of Table 1. An almost identical relationship can be observed between patient and satisficing DMs. The same intuition applies to the impatient DMs, who, in this case, disregard the lower half of the ranking.

Finally, the $[p_1,\ldots, p_{10}] = [10,\ldots, 95]$ framework preserves the relations and patterns observed in the previous ones, while delivering higher CTRs and average number of pages clicked for all types of DMs. In this case, impatient DMs do not consider evaluating the last third of the alternatives composing ranking.

The main feature derived from the different CTRs is the smoother evaluation patterns exhibited by satisficing DMs relative to impatient ones, whose retrieval behavior tends to display abrupt changes throughout all frameworks.

4.1. The information retrieval algorithms at work

Table 2 presents samples from the simulations of all the evaluation scenarios within the $[p1,\ldots, p10] = [10,\ldots, 95]$ framework. This table – together with Figs. 1 to 6 – provides an intuitive description of the actual behavior of the different algorithmic structures simulated numerically. Indeed, the entries of the table have been taken from the output matrices derived from the implementation of the different algorithms. Each column represents a search query performed by a DM. That is, the columns composing Table 2 are subsamples from the three million queries simulated, a million per algorithmic structure.

The rows composing the "Stochastic Realizations" section of the table present the values of the uniformly distributed realizations assigned to each alternative as the DM proceeds through the ranking. These realizations are compared to the cutoff values defined for each alternative, determining the incentives of the DM to click on the corresponding link. The rows composing the "Pages Clicked" section of the table summarize the subsequent information retrieval process per query. That is, each column within this section describes the total number of pages clicked together with their position within the ranking. We

Table 2
Simulation samples from the $[p1,\ldots, p10] = [10,\ldots, 95]$ framework.

Cutoff Values		Search Queries											
		Patient				Satisficing				Impatient			
0.10	Stochastic Realizations	0.815	0.158	0.656	0.706	0.079	0.530	0.472	0.051	0.118	0.323	0.061	0.393
0.20		0.906	0.971	0.036	0.032	0.100	0.982	0.021	0.477	0.430	0.857	0	0.522
0.30		0.127	0.957	0.849	0.277	0.490	0.774	0	0.601	0.312	0.231	0	0.810
0.40		0.913	0.485	0.934	0.046	0.903	0.355	0	0.981	0.488	0	0	0.988
0.50		0.632	0.8	0.679	0.097	0.696	0	0	0.702	0.245	0	0	0.850
0.60		0.098	0.142	0.758	0.823	0.879	0	0	0.896	0	0	0	0.866
0.70		0.278	0.422	0.743	0.695	0.250	0	0	0.145	0	0	0	0.569
0.80		0.547	0.916	0.392	0.317	0	0	0	0	0	0	0	0
0.90		0.958	0.792	0.655	0.95	0	0	0	0	0	0	0	0
0.95		0.965	0.959	0.171	0.034	0	0	0	0	0	0	0	0
Pages Clicked		1	1	1	1	3	1	1	2	1	1	0	1
		2	2	3	6	4	2	0	3	2	2	0	2
		4	3	4	9	5	3	0	4	3	0	0	3
		5	4	5	0	6	0	0	5	4	0	0	4
		9	5	6	0	0	0	0	6	0	0	0	5
		10	8	7	0	0	0	0	0	0	0	0	6
		0	10	0	0	0	0	0	0	0	0	0	0
		0	0	0	0	0	0	0	0	0	0	0	0
		0	0	0	0	0	0	0	0	0	0	0	0
		0	0	0	0	0	0	0	0	0	0	0	0

provide below a comparison of the main differences arising across algorithms within each subset of queries.

Consider the first column within the patient scenario. The algorithm evaluates the whole set of alternatives composing the ranking but only accounts for clicks if the realizations are higher than the corresponding cutoff values. This is the case for the first and second alternatives, whose realizations are higher than 0.10 and 0.20, respectively. The algorithm acknowledges this fact and records the first and second page within the first two rows of the "Pages Clicked" section.

Satisficing and impatient DMs would conclude their retrieval process here since the realization of the third alternative, 0.127, is lower than 0.30. Patient DMs bypass the third alternative, which is not recorded among the pages clicked, and proceed with the fourth one. The realization assigned to the fourth alternative is higher than 0.40, leading to a click and its subsequent recording in the third row of the "Pages Clicked" section.

The retrieval process proceeds in the same way through the remaining part of the ranking, with the DM omitting alternatives 6 to 8 but clicking on the last two. The set of entries composing the "Pages Clicked" section summarizes the results from the retrieval process, describing the pages that have been clicked by the DM. Note how, independently of the number of preference misalignments arising through the process, the DM proceeds until the end of the ranking. That is, the ten rows composing the "Stochastic Realizations" section have all been assigned an actual realization. If the DM were to interrupt the retrieval process before reaching the end of the ranking, a value of zero would have been assigned to the corresponding entries.

A similar intuition, though leading to very different results, follows from the other two scenarios. The satisficing one requires the DM to observe an alternative aligning with his preferences before considering ending the search. The first column within the "Satisficing" subset illustrates how the realizations of the first two alternatives do not align with the preferences of the DM. Thus, both alternatives are omitted from the "Pages Clicked" section.

The first alternative aligning with the preferences of the DM is the third one, whose realization is above the corresponding cutoff value of 0.30. The alternative is recorder in the first row of the "Pages Clicked" section, allowing the DM to consider ending the search as soon as he observes a realization that does not align with his preferences. In this case, the seventh realization compels the DM to stop the search. The fourth to sixth alternatives align with the preferences of the DM and are therefore recorded in the corresponding rows of the "Pages Clicked" section.

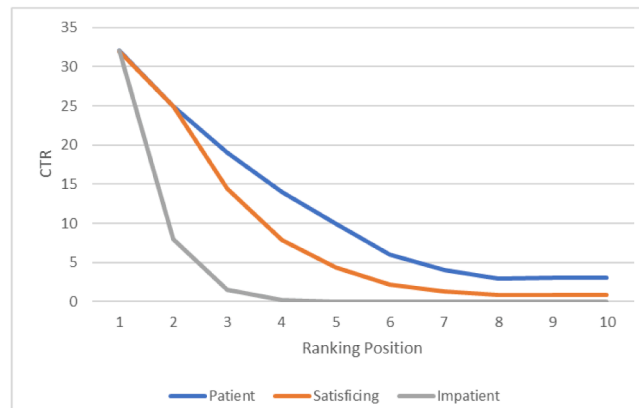
It should be emphasized that the DM does not reach the end of the ranking in any of the queries displayed. For this to be the case, the alternatives should keep on satisficing the preferences of the DM all the way to the end of the ranking. Note also that unless all the alternatives display realizations below the corresponding cutoff values, the DM finds at least one alternative to click. This conclusion applies also to the patient scenario described above. Thus, though satisficing DMs exhibit a higher degree of impatience, they click on at least one satisficing alternative if observed through the retrieval process. On the other hand, satisficing DMs may forego alternatives aligning with their preferences, while patient DMs identify all of them. The opposite end of the spectrum is defined by the impatient DMs, who may not click on any alternative even if the engine provides several of them aligning with their preferences.

The information retrieval processes of impatient DMs are described within the final subset of columns. They represent the stricter and, at the same time, simpler scenario to simulate. Impatient DMs end the search as soon as they observe an alternative that does not align with their preferences. As illustrated by the third column within the subset, this restriction implies that DMs may stop retrieving information after observing the first realization. The first column describes the retrieval of information from four alternatives, emphasizing the fact that for an impatient DM to proceed through the whole set of alternatives, all of them must align with his preferences.

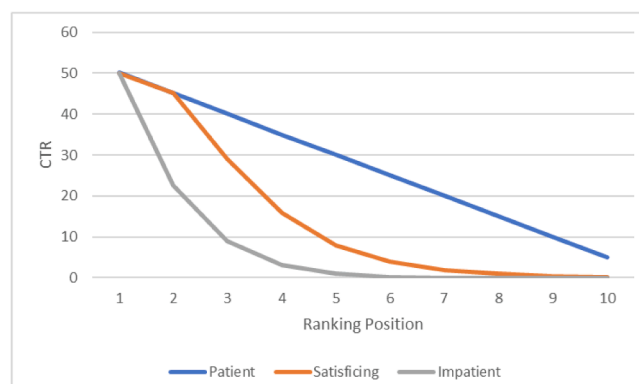
The algorithms deliver 1,000,000 search queries per retrieval structure. The CTRs obtained per page are the result of dividing the total number of times each page appears within the "Pages Clicked" section of the matrix by one million. We now proceed with the analysis of the main results obtained from the different sets of numerical simulations.

4.2. Numerical results

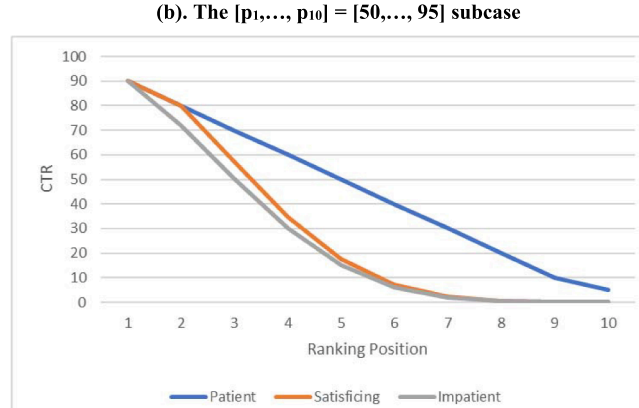
As intuition suggests, we observe that the $[p_1,\ldots, p_{10}] = [68,\ldots, 97]$ and $[p_1,\ldots, p_{10}] = [50,\ldots, 95]$ frameworks display lower CTRs than $[p_1, \ldots, p_{10}] = [10,\ldots, 95]$ due to their higher cutoff values. In this regard, Fig. 7 illustrates the results described in Table 1, allowing for an immediate evaluation of the main differences in CTRs arising among the different scenarios. The horizontal axis represents the ranking position of the alternatives, while CTRs are defined on the vertical axis.



(a). The $[p_1,\ldots, p_{10}] = [68,\ldots, 97]$ subcase



(b). The $[p_1,\ldots, p_{10}] = [50,\ldots, 95]$ subcase



(c). The $[p_1,\ldots, p_{10}] = [10,\ldots, 95]$ subcase

Fig. 7. Differences in CTRs across different scenarios.

Consider the $[p_1,\ldots, p_{10}] = [68,\ldots, 97]$ framework described in Fig. 7 (a) and note the divergent tendency exhibited by the satisficing and impatient users relative to the patient benchmark. The CTR differences between satisficing and impatient users tend to vanish as the cutoff values decrease, namely, as the acceptance probabilities increase when moving from Fig. 7(a) to 7(c). In this regard, note the substantial increase in CTR differences that takes place as impatience increases relative to the patient benchmark, especially when users proceed halfway through the ranking.

The numerical results obtained highlight the importance of being ranked among the first three alternatives, a feature particularly evident as impatience increases through the $[p_1,\ldots, p_{10}] = [68,\ldots, 97]$ and $[p_1, \ldots, p_{10}] = [50,\ldots, 95]$ frameworks. These results also underline the structural complexities inherent to the formalization of the information retrieval processes of DMs. Note that a substantial number of potential scenarios can be defined based on the algorithmic structures presented. For instance, the satisficing algorithm could be easily extended to incorporate users exhibiting a lower degree of impatience and requiring two misalignments to conclude the retrieval process.

Finally, we analyze the substantial differences arising across frameworks regarding the average number of clicks per search described in the third row of Table 1. Jansen et al. (1998) and Baeza-Yates (2005) found that users tend to click an average of two pages per query. The values obtained are close to two pages per query among satisficing and patient users within the $[p_1,\ldots, p_{10}] = [50,\ldots, 95]$ framework, and among satisficing and impatient users within $[p_1,\ldots, p_{10}] = [10,\ldots, 95]$. The results corresponding to the $[p_1,\ldots, p_{10}] = [68,\ldots, 97]$ framework are relatively lower, implying that in order for users to click on two pages per query, they should operate under a less restrictive evaluation framework than the one mimicking the CTRs reported by Dean (2019).

5. Managerial implications

The decision trees described through the paper are sufficiently flexible to incorporate any subjective constraints inherent to the preferences of DMs. In addition, the trees can be modified at each node depending on the complexity of the decision process that must be analyzed. This latter feature highlights the main implications of the current set of algorithms from a managerial viewpoint.

Managers are provided with a benchmark structure illustrating the main consequences derived from any potential modification to the information retrieval behavior of DMs – through their beliefs or any subjective characteristic –. Strategic environments involving the reception of signals or modifications in the learning capacities of DMs can be easily incorporated into any of the algorithms.

As a second implication, we must emphasize the capacity of the algorithms to simulate the effects from different types of shocks to the information retrieval incentives of DMs in terms of the relative ranking position of the alternatives. In this regard, the algorithms formalize a benchmark scenario built on the observed CTRs of DMs, constituting a result of considerable practical importance.

An empirical counterpart to the current analysis should evaluate the responses of groups of DMs within controlled experimental environments. An illustrative example is provided by the introduction of frictions through the information retrieval process. Frictions take place when DMs do not click on an alternative after observing a realization. The misalignment between the description provided by the engine and the preferences of DMs lowers their confidence in the ranking. The benchmark patient algorithm – as well as the satisficing one – allows for the introduction of frictions of varying intensity and duration at any point through the retrieval process.

We define a scenario with two frictions, denoted by $f_i, i = 1,2$, with $f_1 < f_2$. That is, when the preferences of the DM do not align with the characteristics of the alternatives for the first time, the DM adds friction f_1 to all the subsequent cutoff values. If a second misalignment occurs through the process, a higher friction f_2 is added to the remaining cutoff values. Additional misalignments do not increase the value of the frictions. We have thus imposed an upper ceiling on the number of times misalignments decrease the trust of users in the ranking. The duration of the frictions could also be modified by assuming that DMs do not maintain their distrust through the whole retrieval process.

Frictions can be interpreted as increments in the impatience of DMs, who become less willing to click on the alternatives displayed by the engine. In this case, the impatience of DMs increases through a maximum of two misalignments and remains constant afterwards, though it could be assumed that frictions continue to grow throughout the entire process as misalignments occur. The retrieval scenario with both frictions is summarized in Fig. 8, while Fig. 9 and Table 3 describe the numerical consequences from the incorporation of frictions to the analysis.

In particular, Table 3 and Fig. 9 describe two numerical settings, a simple one defined by a unique friction, $f_1 = 0.1$, and a more complex cumulative structure with $f_1 = 0.1$ and $f_2 = 0.2$. Both settings have been incorporated within the $[p_1,\ldots, p_{10}] = [10,\ldots, 95]$ benchmark patient algorithm. The incremental decrease in CTRs described in Table 3 is represented in Fig. 9, which illustrates the CTR differences between both settings and the patient frictionless scenario.

The non-linearity of the CTR differences highlights the concentration of the effects from an incremental shock on the alternatives located within the lower half of the ranking, namely, those endowed with relatively lower CTRs. Thus, a unique decrement in the willingness of DMs to retrieve information should trigger different responses depending on the relative position of the alternatives within the ranking. At the same time, a cumulative decrement exacerbates the response of DMs as they proceed through the middle and lower sections of the ranking.

We conclude by performing an intuitive exercise on the impatient DMs. We assume that instead of ending the search after observing an alternative that does not align with their preferences, impatient DMs require two misalignments through the retrieval process to end the search. The enhanced impatient algorithm relaxes the retrieval constraints imposed on the initial environment and brings the performance of the model closer to that of the satisficing scenario. The results from the corresponding simulations are presented in Table 4.

A direct comparison between the CTRs of the enhanced impatient and satisficing scenarios illustrates the substantial increment exhibited by the former through the different evaluation frameworks. That is, the CTRs of the enhanced impatient algorithm lag behind the satisficing ones within the $[p_1,\ldots, p_{10}] = [68,\ldots, 97]$ framework, catch up through $[p_1,\ldots, p_{10}] = [50,\ldots, 95]$ and overtake the satisficing CTRs within $[p_1, \ldots, p_{10}] = [10,\ldots, 95]$. Allowing for two misalignments increases the CTRs of impatient DMs, an effect that gains momentum as the constraints imposed by the cutoff values are relaxed, increasing the alignment probability through the retrieval process.

In addition, when comparing the impatient and enhanced impatient scenarios, we can identify the alternatives experiencing higher increments in their CTRs, providing the corresponding firms with the necessary incentives to foster the willingness of DMs to retrieve information.

All in all, the algorithms presented endow managers with the capacity to simulate a substantial number of scenarios determined by the assumptions imposed on DMs regarding their willingness to retrieve information from the alternatives.

6. Conclusion

We have designed a benchmark algorithm that mimics the information retrieval behavior of users when evaluating the initial page of alternatives ranked by a search engine. Several versions of the algorithm have been defined to account for different degrees of user impatience and their effects on the resulting CTRs. The main results obtained are conditioned by the structural differences existing among the impatient scenarios analyzed, which, together with the corresponding evaluation

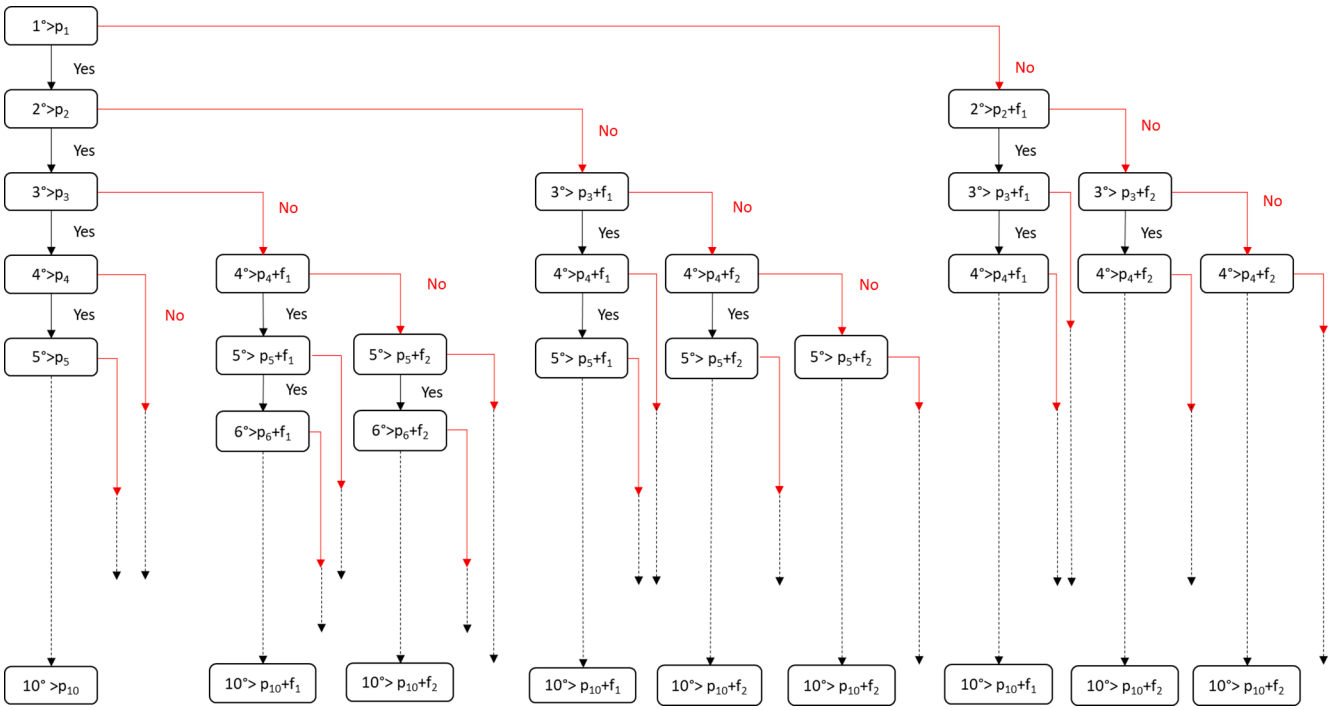


Fig. 8. Structure of the patient evaluation algorithm with two levels of friction.

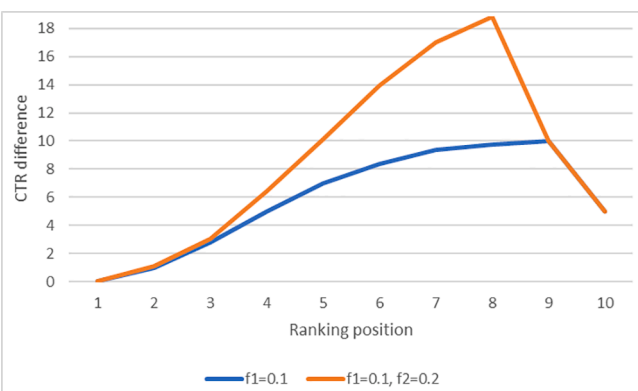


Fig. 9. CTR differences relative to the patient frictionless scenario.

Table 3
CTRs without and with one and two levels of friction.

Framework	$[p_1,\ldots, p_{10}] = [10,\ldots, 95]$		
CTR	Patient	$f_1 = 0.1$	$[f_1 = 0.1; f_2 = 0.2]$
Avg	4.55	3.97	3.70
1	90.00	89.95	89.98
2	80.00	79.01	78.91
3	69.96	67.15	66.91
4	60.07	55.08	53.66
5	50.00	43.04	39.87
6	39.97	31.58	26.04
7	30.01	20.66	13.01
8	19.96	10.20	1.15
9	10.02	0.04	0.03
10	5.01	0	0

Table 4
CTRs of the enhanced impatient scenario.

CTR	Enhanced Impatient Scenario		
	$[p_1,\ldots, p_{10}] = [68,\ldots, 97]$	$[p_1,\ldots, p_{10}] = [50,\ldots, 95]$	$[p_1,\ldots, p_{10}] = [10,\ldots, 95]$
Avg	0.69	1.48	3.63
1	31.96	49.94	89.98
2	25.07	45	80.04
3	9.34	28.94	68.53
4	2.19	14.91	54.11
5	0.35	6.23	37.15
6	0.03	2.11	20.93
7	0	0.57	8.99
8	0	0.11	2.63
9	0	0.01	0.42
10	0	0	0.04

probabilities, determine the set of alternatives that may be clicked throughout the retrieval process.

The satisficing algorithm displays a divergent behavior relative to the patient benchmark setting within all the evaluation frameworks analyzed. It is also intuitively clear that a simpler satisficing structure requiring DMs to discontinue their information retrieval processes as soon as they come across a suitable alternative would result in an even more divergent scenario relative to the behavior observed.

An important feature of the algorithmic structures presented is their malleability when dealing with frictions and the information retrieval incentives of users. For instance, impatience could be interpreted as decrements in the attention span of users as they proceed through the ranking. In this regard, the fact that strategically positioned stimuli can renew the attention of DMs through their search processes can be easily incorporated into the proposed algorithms (Ahn et al., 2018).

We conclude by emphasizing that even though the algorithms presented mimic rational decision-making processes, they are built using

empirical observations while leaving aside the design of formal counterparts. Future research should aim at validating the combinatorial optimality of the behavior observed, particularly when accounting for the evaluation of different subsets of alternatives within those composing the initial page of search results provided by an engine.

CRediT authorship contribution statement

Debora Di Caprio: Formal analysis, Validation, Writing – review & editing. **Francisco J. Santos-Arteaga:** Formal analysis, Investigation, Methodology, Writing – review & editing. **Madjid Tavana:** Supervision, Validation, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

Dr. Madjid Tavana is grateful for the partial financial support he received from the Czech Science Foundation (GACR 19-13946S).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.eswa.2021.116352.

References

Ahn, J. H., Bae, Y. S., Ju, J., & Oh, W. (2018). Attention adjustment, renewal, and equilibrium seeking in online search: An eye-tracking approach. *Journal of Management Information Systems, 35*, 1218–1250.

Bae, S., & Lee, T. (2011). Product type and consumers' perception of online consumer reviews. *Electronic Markets, 21*, 255–266.

Baeza-Yates, R. Applications of web query mining. In: Losada, D.E., Fernández-Luna, J. M. (Eds.). Advances in Information Retrieval. Information Systems and Applications, incl. Internet/Web, and HCI, Vol. 3408. Springer-Verlag, Berlin Heidelberg, 7-22 (2005).

Bartels, D. M., & Johnson, E. J. (2015). Connecting cognition and consumer choice. *Cognition, 135*, 47–51.

Basu, S. (2018). Information search in the internet markets: Experience versus search goods. *Electronic Commerce Research and Applications, 30*, 25–37.

Bell, D., & Mgbemena, C. (2018). Data-driven agent-based exploration of customer behavior. *Simulation, 94*, 195–212.

Bolandifar, E., DeHoratius, N., Olsen, T., & Wiler, J. (2019). An empirical study of the behavior of patients who leave the emergency department without being seen. *Journal of Operations Management, 65*, 430–446.

Chitika: The value of Google result positioning. Chitika Insights June 7, 2013. Chitika, Westborough (2013) Available at perma.cc/7AGC-HTDH.

Cristofaro, M. (2017). Herbert Simon's bounded rationality: Its historical evolution in management and cross-fertilizing contribution. *Journal of Management History, 23*, 170–190.

Dean, B. (2019). We analyzed 5 million Google search results. Here's what we learned about organic click through rate. Available at https://backlinko.com/google-ctr-stats.

Dell'Aversana, R., & Bucciarelli, E. (2018). Towards a natural experiment leveraging big data to analyse and predict users' behavioural patterns within an online consumption setting. In *International Symposium on Distributed Computing and Artificial Intelligence 2018 Jun 20* (pp. 103–113).

Dimoka, A., Hong, Y., & Pavlou, P. A. (2012). On product uncertainty in online markets: Theory and evidence. *MIS Quarterly, 36*, 395–426.

Doniec, A., Lecoeuche, S., Mandiau, R., & Sylvain, A. (2020). Purchase intention-based agent for customer behaviours. *Information Sciences, 521*, 380–397.

Dou, W., Lim, K. H., Su, C., Zhou, N., & Cui, N. (2010). Brand positioning strategy using search engine marketing. *MIS Quarterly, 34*, 261–279.

Dunke, F., & Nickel, S. (2020). Neural networks for the metamodeling of simulation models with online decision making. *Simulation Modelling Practice and Theory, 99*, Article 102016.

Epstein, R., & Robertson, R. E. (2015). The Search Engine Manipulation Effect (SEME) and its possible impact on the outcomes of elections. *Proceedings of the National Academy of Sciences, 112*, E4512–E4521.

European Commission: Online platforms. (2016). Special Eurobarometer 447- April 2016. European Union.

Fishkin, R. (2019). Less than half of Google searches now result in a click, Aug. 13, Available at https://sparktoro.com/blog/less-than-half-of-google-searches-now-result-in-a-click/.

Gao, R., & Shah, C. (2020). Toward creating a fairer ranking in search engine results. *Information Processing & Management, 57*, Article 102138.

Ghafurian, M., Reitter, D., & Ritter, F. E. (2020). Countdown timer speed: A trade-off between delay duration perception and recall. *ACM Transactions on Computer-Human Interaction (TOCHI), 27*, 1–25.

Google: The need for mobile speed: How mobile latency impacts publisher revenue. September (2016). Available at https://www.thinkwithgoogle.com/intl/en-154/insights-inspiration/research-data/need-mobile-speed-how-mobile-latency-impacts-publisher-revenue/.

Gupta, S., Kar, A. K., Baabdullah, A., & Al-Khowaiter, W. A. (2018). Big data with cognitive computing: A review for the future. *International Journal of Information Management, 42*, 78–89.

Hong, W., Xiong, Z., You, J., Wu, X., & Xia, M. (2021). CPIN: Comprehensive present-interest network for CTR prediction. *Expert Systems with Applications, 168*, Article 114469. https://doi.org/10.1016/j.eswa.2020.114469

Jankowski, J., Kazienko, P., Wątróbski, J., Lewandowska, A., Ziemba, P., & Zioło, M. (2016). Fuzzy multi-objective modeling of effectiveness and user experience in online advertising. *Expert Systems with Applications, 65*, 315–331.

Jansen, M. B. J., Spink, A., Bateman, J., & Saracevic, T. (1998). Real life information retrieval: A study of user queries on the web. *ACM SIGIR Forum, 32*, 5–17.

Karanam, S., van Oostendorp, H., & Fu, W. T. (2016). Performance of computational cognitive models of web-navigation on real websites. *Journal of Information Science, 42*, 94–113.

Khamitov, M., Wang, X., & Thomson, M. (2019). How well do consumer-brand relationships drive customer brand loyalty? Generalizations from a meta-analysis of brand relationship elasticities. *Journal of Consumer Research, 46*, 435–459.

Kimmel, A. J. (2012). *Psychological foundations of marketing.* Routledge.

Kreye, M. E., Goh, Y. M., Newnes, L. B., & Goodwin, P. (2012). Approaches to displaying information to assist decisions under uncertainty. *Omega, 40*, 682–692.

Lauraéus, T., Saarinen, T., & Öörni, A.: Factors affecting consumer satisfaction of online purchase. In: 48th Hawaii International Conference on System Sciences, Kauai, HI, 2015, 3364-3373. doi: 10.1109/HICSS.2015.406.

Lee, H., & Pang, N. (2018). Understanding the effects of task and topical knowledge in the evaluation of websites as information patch. *Journal of Documentation, 74*, 162–186.

Lerner, J. S., Li, Y., Valdesolo, P., & Kassam, K. S. (2015). Emotion and decision making. *Annual Review of Psychology, 66*, 799–823.

Lewandowski, D., & Kammerer, Y. (2020). Factors influencing viewing behaviour on search engine results pages: A review of eye-tracking research. *Behaviour & Information Technology, 1–31*. https://doi.org/10.1080/0144929X.2020.1761450

Li, G., Huang, J. Z., & Shen, H. (2018). To wait or not to wait: Two-way functional hazards model for understanding waiting in call centers. *Journal of the American Statistical Association, 113*, 1503–1514.

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*, E1. https://doi.org/10.1017/S0140525X1900061X

List, A., & Alexander, P.A. (2017) Text navigation in multiple source use. Computers in Human Behavior 75, 364-375.

Liu, Y., Miao, J., Zhang, M., Ma, S., & Ru, L. (2011). How do users describe their information need: Query recommendation based on snippet click model. *Expert Systems with Applications, 38*, 13847–13856.

Lohr, S. (Feb. 29, 2012). For impatient web users, an eye blink is just too long to wait. The New York Times, Available at https://www.nytimes.com/2012/03/01/technology/impatient-web-users-flee-slow-loading-sites.html.

Lorigo, L., Haridasan, M., Brynjarsdóttir, H., Xia, L., Joachims, T., Gay, G., … Pan, B. (2008). Eye tracking and online search: Lessons learned and challenges ahead. *Journal of the American Society for Information Science and Technology, 59*, 1041–1052.

Lu, W., & Altenbek, G. (2021). A recommendation algorithm based on fine-grained feature analysis. *Expert Systems with Applications, 163*, Article 113759. https://doi.org/10.1016/j.eswa.2020.113759

Lu, A. C. C., & Gursoy, D. (2015). A conceptual model of consumers' online tourism confusion. *International Journal of Contemporary Hospitality, 27*, 1320–1342.

Luo, W., Cook, D., & Karson, E. J. (2011). Search advertising placement strategy: Exploring the efficacy of the conventional wisdom. *Information & Management, 48*, 404–411.

Misuraca, R., & Fasolo, B. (2018). Maximizing versus satisficing in the digital age: Disjoint scales and the case for "construct consensus". *Personality and Individual Differences, 121*, 152–160.

Machackova, H., & Smahel, D. (2018). The perceived importance of credibility cues for the assessment of the trustworthiness of online information by visitors of health-related websites: The role of individual factors. *Telematics and informatics, 35*, 1534–1541.

Mahony, C., Sammon, D., & Heavin, C. (2016). Design guidelines for online resources: A longitudinal analysis of information processing. *Journal of Decision Systems, 25*, 329–342.

Nah, F. F. (2004). A study on tolerable waiting time: How long are web users willing to wait? *Behaviour & Information Technology, 23*, 153–163.

Norman, K. L., & Kirakowski, J. (Eds.). (2017). *The Wiley handbook of human computer interaction.* John Wiley & Sons.

Ong, K., Järvelin, K., Sanderson, M., & Scholer, F. (2017). Using information scent to understand mobile and desktop web search behavior. In *Proceedings of the 40th*

international ACM SIGIR conference on research and development in information retrieval (pp. 295–304).
Pei, S., & Hu, Q. (2018). Partially monotonic decision trees. *Information Sciences, 424*, 104–117.
Pirolli, B. (2018). Travel information online: Navigating correspondents, consensus, and conversation. *Current Issues in Tourism, 21*, 1337–1343.
Ren, H., & Huang, T. (2018). Modeling customer bounded rationality in operations management: A review and research opportunities. *Computers & Operations Research, 91*, 48–58.
Richardson, R. C. (2017). Heuristics and satisficing. In: W. Bechtel and G. Graham (Eds.) A companion to cognitive science. John Wiley & Sons, pp. 566-575. https://doi.org/10.1002/9781405164535.ch44.
Sadiq, S., Umer, M., Ullah, S., Mirjalili, S., Rupapara, V., & Nappi, M. (2021). Discrepancy detection between actual user reviews and numeric ratings of Google App store using deep learning. *Expert Systems with Applications, 181*, Article 115111. https://doi.org/10.1016/j.eswa.2021.115111
Sagi, O., & Rokach, L. (2020). Explainable decision forest: Transforming a decision forest into an interpretable tree. *Information Fusion, 61*, 124–138.
Schneider, M., Deck, C., Shor, M., Besedeš, T., & Sarangi, S. (2019). Optimizing choice architectures. *Decision Analysis, 16*, 2–30.
Shafiq, O., Alhajj, R., & Rokne, J. G. (2015). On personalizing Web search using social network analysis. *Information Sciences, 314*, 55–76.
Speier-Pero, C. (2019). Using aggregated data under time pressure: A mechanism for coping with information overload. *Journal of Decision Systems, 28*, 82–100.
Sun, H., Fan, M., & Tan, Y. (2020). An empirical analysis of seller advertising strategies in an online marketplace. *Information Systems Research, 31*, 37–56.
Varnali, K., Yilmaz, C., & Toker, A. (2012). Predictors of attitudinal and behavioral outcomes in mobile advertising: A field experiment. *Electronic Commerce Research and Applications, 11*, 570–581.
Victorelli, E. Z., Dos Reis, J. C., Hornung, H., & Prado, A. B. (2020). Understanding human-data interaction: Literature review and recommendations for design. *International Journal of Human-Computer Studies, 134*, 13–32.
Yoo, B., Jeon, S., & Han, T. (2016). An analysis of popularity information effects: Field experiments in an online marketplace. *Electronic Commerce Research and Applications, 17*, 87–98.
Zhang, J., Adomavicius, G., Gupta, A., & Ketter, W. (2020). Consumption and performance: Understanding longitudinal dynamics of recommender systems via an agent-based simulation framework. *Information Systems Research, 31*, 76–101.
Zillinger, M. (2020). The curious case of online information search. *Current Issues in Tourism, 23*, 276–279.
Zhu, F., & Zhang, X. (2010). Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics. *Journal of Marketing, 74*, 133–148.