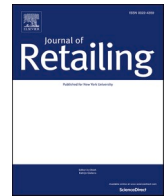




Contents lists available at ScienceDirect

Journal of Retailing

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

Consumer search patterns in online retailing

Anna Dreneva^{a,*} , Jacob L. Orquin^{b,c} 

^a Department of Psychology and Behavioural Sciences, Aarhus University, Denmark

^b Department of Management, Aarhus University, Denmark

^c Department of Business and Economics, Reykjavik University, Iceland

ARTICLE INFO

Editor: Dr Katrijn Gielens

Keywords:

Online retailing
Consumer search
Eye tracking
Visual attention
Scroll website

ABSTRACT

Online retailers routinely rely on sorting, filters, and personalized assortments, but the effectiveness of these design choices depends on how consumers search and decide in digital environments. This research develops a unified and behaviorally grounded framework that conceptualizes consumer search as a context-dependent screening process. Using a novel approach that integrates eye-tracking, mouse-scroll behavior, subjective product preferences, and choice data, we analyze sequences of search events across two studies. The findings show that consumers do not follow a single search strategy but instead adapt their search flexibly to product value and context. Search unfolds as an adaptive screening process in which consumers scan many low-value products, examine a smaller set of higher-value options, and compare only a few top candidates before choosing. Moreover, consumers search less when the overall value of seen products is higher. This framework integrates previously disconnected research streams on decision processes, stopping rules, and consideration set formation, and offers actionable guidance for design and decision-support tools in online retailing.

Online retailing offers several technological advantages over physical stores. Retailers can readily sort assortments by price or popularity, strategically place promotions, co-locate related products to facilitate comparisons, and personalize assortments for specific consumption occasions. Effectively leveraging these capabilities depends on how consumers search for information and make decisions in digital environments (Honka et al., 2024; Kahn, 2017). At the same time, online retailing differs from physical retailing in ways that fundamentally shape consumer search. Whereas physical stores provide a multi-sensory shopping experience, online shopping relies almost entirely on visual information (Gauri et al., 2021; Kahn, 2017; Szocs et al., 2023). Search costs also tend to be lower online (Ratchford et al., 2022), as consumers are not constrained by physical movement, carts, or baskets (Larsen et al., 2020). These differences imply that consumer search may unfold differently in online settings, with greater emphasis on visual screening, product ordering, and interface design. Although existing research on consumer search has yielded important insights, it has largely been developed in contexts that do not fully reflect these digital characteristics (e.g., Chandon et al., 2009; Chen et al., 2021; Larsen et al., 2020; Payne et al., 1988; Seiler & Pinna, 2017; Stüttgen et al., 2012; Van der Lans et al., 2008). As online retail environments continue to evolve, understanding how consumers search, evaluate, and select products under these conditions becomes increasingly important for both theory and practice.

This research advances understanding of consumer search in online retailing by examining how consumers engage in different types of search events (e.g., scanning, examining, and comparing) and how these events interact with stopping rules and screening

* Corresponding author.

E-mail address: anna.dreneva@psy.au.dk (A. Dreneva).

<https://doi.org/10.1016/j.jretai.2026.05.004>

0022-4359/© 2026 The Authors. Published by Elsevier Inc. on behalf of New York University. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Please cite this article as: Anna Dreneva, Jacob L. Orquin, *Journal of Retailing*, <https://doi.org/10.1016/j.jretai.2026.05.004>

mechanisms. Using a data-driven approach, we jointly examine three long-standing debates in consumer search, single versus multiple decision processes, fixed versus sequential search, and consideration set formation, by framing them in terms of observable behavioral patterns. This approach allows us to uncover search patterns not predicted by existing models and to study consumer behavior in naturalistic online environments without constraining analyses to specific theoretical assumptions.

To do so, we develop a novel method that integrates eye-tracking, mouse-scroll behavior, subjective product values, and choice data while consumers complete a shopping task on a live retail website. Search is classified into sequences of discrete events, such as scanning, examination, and comparison (Lohse & Johnson, 1996; Payne et al., 1988; Russo & Leclerc, 1994).¹ Each sequence begins when a product category becomes visible and ends with a product selection. We compare sequences using Optimal String Alignment (OSA), a method originally developed in bioinformatics to quantify similarity between sequences (Damerou, 1964). We then cluster these similarity measures to examine how search patterns relate to choices and subjective product values, which were elicited via individual ratings (1–100 in Study 1 and 1–7 in Study 2).

Our research makes several important contributions. First, it advances understanding of consumer search in online retailing by introducing an integrated framework that links decision processes, search dynamics, and consideration set formation. We conceptualize consumer search as a context-dependent screening process that unfolds through scanning, examination, and comparison. Consumers initially engage in broad scanning to efficiently filter out low-value options, examine promising alternatives in greater depth, and, when necessary, compare top candidates directly. These behaviors operate alongside two stopping rules: (a) terminating search upon identifying a satisfactory product or (b) stopping after reaching a fixed time limit. Together, these mechanisms make search highly context dependent: when high-value options are readily available, consumers terminate search early; when such options are scarce, they continue searching until time constraints bind.

Second, our research offers practical insights for retailers seeking to optimize consumer search and decision making. Our findings highlight how tools such as product sorting, filtering, and wishlists shape search, and how behavioral data, including eye-tracking and mouse-scrolling measures, can be used to infer consumers' decision processes. Finally, we outline directions for future empirical and theoretical research to further refine understanding of the screening process, with particular attention to how

1. Theories on consumer search and decision making

Prior research on consumer search and decision making in retail contexts has developed along three largely independent theoretical themes. The first theme concerns the nature and number of underlying decision processes. Some theories assume that consumers rely on a single decision process, whereas others propose that consumers can flexibly choose among multiple, qualitatively different processes depending on the task and context (for a discussion, see Krefeld-Schwalb et al., 2019; Söllner & Bröder, 2016). The second theme focuses on how consumers decide when to stop searching. In this stream, models differ in whether consumers commit in advance to evaluating a fixed number of products (Stigler, 1961) or instead make sequential stopping decisions, reassessing after each option whether to continue or terminate search (Weitzman, 1978). The third theme examines whether and how consumers form consideration sets that is, whether search unfolds through distinct and consecutive screening stages that progressively narrow the set of viable options (Hauser, 2014; Russo & Leclerc, 1994; Shocker et al., 1991). Although each of these themes has generated important insights, they have largely been studied in isolation. This fragmentation poses a challenge for online retailers, whose design choices, such as product sorting, filtering tools, and assortment structure, simultaneously influence decision processes, stopping behavior, and screening mechanisms. As a result, retailers lack a unified framework for predicting how consumers will respond to changes in the shopping environment. In the following sections, we review each theme in more detail and highlight the need for an integrated perspective on consumer search in online retailing.

1.1. Single vs multiple decision processes

A central debate in the consumer search literature concerns whether consumers rely on a single decision process or flexibly switch among multiple processes (Krefeld-Schwalb et al., 2019; Söllner & Bröder, 2016). Single-process models assume that search and choice are governed by a common mechanism, such as evidence accumulation models in which consumers sequentially sample information until a decision threshold is reached (Busemeyer & Townsend, 1993; Lee & Cummins, 2004). By adjusting this threshold, such models can account for adaptive changes in search intensity (Söllner & Bröder, 2016). In contrast, multiple-process models propose that consumers adapt their search by selecting among qualitatively different strategies depending on the decision environment (Gigerenzer & Todd, 1999; Payne et al., 1988). For example, consumers may rely on simple heuristics such as lexicographic choice in low-involvement settings or use elimination-by-aspect strategies when multiple attributes are relevant (Tversky, 1972). Related work in dual-systems theory similarly suggests that consumers combine fast, intuitive processes with more deliberate, analytical evaluation (Kahneman, 2011). Despite extensive investigation, there is no consensus on whether single- or multiple-process models better characterize consumer search, and both approaches continue to be applied to multi-option choice (Krefeld-Schwalb & Rosner, 2020; Krefeld-Schwalb et al., 2019; Söllner & Bröder, 2016; Stüttgen et al., 2012; Thomas et al., 2021). This unresolved debate has important implications for retail practice. If consumers flexibly shift among decision processes, predicting search and choice becomes highly context dependent, complicating website design and assortment optimization. Conversely, if a common decision process underlies

¹ Prior literature has used terms such as *operations* or *elementary information processes* to describe these events. aspiration levels and time limits are formed and interact during consumer search.

most consumer search, retailers can develop more parsimonious and accurate models to guide product presentation and decision-support tools.

1.2. Fixed vs sequential search

Another central theme in the consumer search literature concerns whether consumers engage in fixed sample search or sequential search (Honka et al., 2019, 2024). In fixed sample search, consumers commit in advance to evaluating a predetermined number of products, whereas in sequential search they evaluate options one at a time and decide after each whether to continue or stop. Classic economic models formalize these approaches by equating the marginal costs and benefits of search (Stigler, 1961) or by specifying optimal stopping rules based on reservation utilities (Weitzman, 1978). Related behavioral models emphasize satisficing, in which consumers stop searching once an option meets an aspiration level, which may itself adapt based on the quality of encountered options (Reutskaja & Hogarth, 2009; Simon, 1979; Stüttgen et al., 2012). Empirical evidence comparing fixed and sequential search is mixed. Field data from online retailing have been interpreted as supporting fixed sample search (Santos et al., 2012), whereas eye-tracking studies provide stronger evidence for sequential search (Lu & Hutchinson, 2024), and other work suggests that consumers combine elements of both strategies (Reutskaja et al., 2011). This distinction has important implications for retail practice. Under fixed sample search, product order should have little effect on search, as consumers will continue searching even if a high-value option appears early. Under sequential search, however, encountering a high-value option early can truncate search, consistent with the overall value effect observed in simpler binary choices (Shevlin et al., 2022).

1.3. Consideration set formation

A third central theme in the consumer search literature concerns whether and how consumers form consideration sets, that is, whether search unfolds through distinct screening stages (Hauser & Wernerfelt, 1990; Roberts & Lattin, 1991; Wright & Barbour, 1977). Seminal work conceptualizes consideration set formation as a multi-stage process in which consumers move from an initial awareness set to a smaller consideration set and ultimately to a final choice set (Shocker et al., 1991). This perspective frames consumer search as a screening process, whereby attention progressively narrows to fewer, higher-value options. Research suggests that early screening often relies on non-compensatory decision rules, such as lexicographic choice or elimination by aspect, to quickly discard inferior options, while later stages involve more compensatory evaluation of remaining alternatives (Bremer et al., 2017; Gilbride & Allenby, 2004; Hauser, 2014). Empirical support for this view comes from both choice modeling (e.g., Gilbride & Allenby, 2004) and direct observation of search. Eye-tracking studies in retail environments, for example, identify sequential stages of orientation, evaluation, and verification that mirror consideration set formation (Russo & Leclerc, 1994). More recent work combining choice and eye-tracking data shows that models allowing for consideration set formation explain behavior better than standard random utility models for a substantial share of consumers (Edenbrandt et al., 2022). Consideration set formation has important implications for online retailers, who can actively shape screening by designing effective sorting and filtering tools and by offering features such as wishlists that help consumers curate and manage a manageable set of options (Groening et al., 2021).

1.4. Connecting the themes

Prior research on consumer search and decision making has largely examined the three theoretical themes in isolation (Krefeld-Schwalb et al., 2025). This fragmentation limits our ability to derive coherent predictions about how consumers respond to retail environments in which decision processes, stopping behavior, and screening mechanisms necessarily interact. Yet, these themes

Table 1

Overview of the relationship between theoretical themes, expected, and observed consumer search patterns.

	Single process: sequential search	Single process: fixed sample search	Multiple processes: sequential or fixed	Multiple processes: consideration set formation	Study 1 & 2 Results
Search events	Single event (e.g., mainly examinations).	Single event (e.g., mainly examinations).	Multiple events (e.g., scanning, examination, comparison).	Multiple events possible.	Multiple events (scanning, examination, and comparison).
Clusters of search	No clusters.	No clusters.	Clusters.	Clusters.	Three clusters: Fast, Intermediate, Slow.
Search stopping rules	More likely after a high-value option is noticed.	After a fixed number of options are examined.	Both rules possible (e.g., SAT vs LEX).	Both rules possible.	After a high-value option is noticed or after scrolling the entire webpage.
Overall value effect	Less extensive search when overall value is higher.	More extensive search when overall value is higher.	Both directions possible (e.g., SAT vs EBA).	Both directions possible.	Less extensive search when overall value of seen options is higher.
Screening behavior	No screening.	No screening.	Screening via stages (e.g., EBA).	Screening via stages: awareness set → consideration set → choice set.	Screening via events: scanning → examination → comparison.

Abbreviations: SAT = Satisficing decision strategy; LEX = Lexicographic decision strategy; EBA = Elimination by aspect decision strategy.

are logically connected and generate distinct, empirically testable predictions when considered jointly (for an overview see Table 1).

Single-process models assume that search is governed by a homogeneous mechanism, implying that search is dominated by a single type of event (e.g., examinations), that consumers do not form distinct clusters of search patterns, and that no screening stages emerge during search (Stigler, 1961; Ursu et al., 2024; Weitzman, 1978). Depending on whether search is fixed or sequential, these models further imply opposing predictions about stopping behavior and the overall value effect: fixed sample search predicts more extensive search when overall value is higher, whereas sequential search predicts earlier stopping when high-value options are encountered. In contrast, multiple-process models allow consumers to switch among qualitatively different strategies, such as lexicographic choice, elimination by aspect, or satisficing, depending on the decision context (Gigerenzer & Todd, 1999; Payne et al., 1988). These models predict heterogeneous search patterns, the emergence of meaningful clusters of behavior, and the coexistence of both fixed and sequential stopping rules. However, they typically do not specify whether or how these strategies unfold over time. Consideration set models extend this view by explicitly proposing that multiple decision processes are organized into distinct screening stages, such that consumers move from broad scanning to focused evaluation and comparison (Shocker et al., 1991). This perspective predicts multiple search events whose prevalence changes over time, structured screening behavior, and flexible stopping rules that depend on both option value and search costs.

In the following section, we introduce a data-driven approach that allows us to jointly test these competing predictions. By examining search events, clustering patterns, stopping rules, overall value effects, and screening behavior within a unified framework, we empirically connect the three themes and assess which theoretical predictions best describe consumer search in online retailing (Table 1).

2. Study 1

Study 1 uses eye-tracking to examine consumer search in a realistic online retail environment. Eye-tracking provides fine-grained measures of visual attention and information acquisition, allowing researchers to identify how consumers navigate digital interfaces, evaluate products, and make purchase decisions (e.g., Boardman & McCormick, 2022; Hwang & Lee, 2018; Modi & Singh, 2023). Importantly, eye-tracking enables direct observation of the identity and sequence of products and attributes viewed, which is critical for inferring search dynamics and decision processes (Kahn, 2017; Martinovici et al., 2021; Ursu et al., 2024). To enhance ecological validity, we conducted the study on a live grocery retail website rather than using simulated interfaces or static screenshots, which are common in prior work (Chocarro et al., 2022; Schröter et al., 2021). We recruited a sample of consumers ($N = 121$) to complete a naturalistic shopping task across 12 product categories. During the task, we recorded eye movements, scrolling behavior, and mouse clicks. Immediately afterward, a custom Python script scraped product images, recorded on-screen product coordinates, and captured full-page screenshots. Participants then rated the products they encountered, presented exactly as they appeared during the shopping task. This procedure ensured consistent measurement of both search and subjective product value. Eye-tracking data were processed using the *eyeScrollR* package, which maps gaze coordinates onto full-page webpage coordinates using scroll information (Larigaldie et al., 2024). Based on these data, we automatically defined areas of interest (AOIs), yielding precise fixation sequences, product-level attention measures, subjective value ratings, and final choices for each category. To test the theoretical predictions summarized in Table 1, we classified search into discrete events: systematic scanning, non-systematic scanning, examination, comparison, scrolling up, scrolling down, and choice. Each search sequence begins when a category page becomes visible and ends with a product selection. We compared sequences using Optimal String Alignment (OSA) and conducted cluster analyses on the resulting distance measures.² If consumers rely on a single decision process, search sequences should be relatively homogeneous and dominated by a single event type. In contrast, multiple-process accounts predict heterogeneous sequences and meaningful clusters characterized by different combinations of events. We examined fixed versus sequential search by analyzing stopping behavior. Sequential search predicts that consumers stop shortly after encountering a satisfactory option, whereas fixed sample search predicts continued search even after the chosen product is identified. These accounts further generate opposing predictions for the overall value effect: fixed sample search predicts more extensive search when overall value is higher, whereas sequential search predicts earlier stopping when high-value options are encountered. Finally, we examined consideration set formation by assessing whether search events change systematically over time and whether scanning, examination, and comparison reflect differences in subjective product values. Specifically, consideration set formation predicts that early search is dominated by scanning, followed by examination and comparison of increasingly higher-value products later in the sequence. Together, these analyses allow us to test competing theoretical predictions within a unified, behaviorally grounded framework.

2.1. Method

Open science statement. Prior to data analysis, the analysis plan was pre-registered on the Open Science Framework (OSF). The preregistration, data, code, and materials used in the study, are available at OSF.³

² An alternative approach is to analyze the sequences using a Hidden Markov Model (HMM), which assumes a set of hidden (unobserved) states to explain the observed sequences. Previous research has used HMMs, for example, to test hypotheses about global versus local attention in consumer search (Liechty et al., 2003). However, given our data-driven approach, Optimal String Alignment offers the advantage of not requiring any assumptions about hidden states.

³ <https://osf.io/54b92/overview>

Participants. A market panel agency recruited a representative sample of 121 adults aged 24–64 between September and October 2023. Participants were active household grocery shoppers with experience in online grocery shopping. Inclusion criteria required participants to be at least 18 years old, have normal color vision, not rely on reading glasses (standard eyeglasses or contact lenses were permitted), be responsible for household grocery shopping (individually or jointly), have familiarity with online grocery shopping, and possess basic English proficiency. Demographic and personal characteristics are reported in the Supplementary Materials (Table S1). Nine participants were excluded due to technical issues or poor eye-tracking calibration quality, as assessed by the default iMotions algorithm, which considers factors such as offset means and standard deviations, screen resolution, and missing calibration samples. The final sample comprised 112 participants (61 females).

Stimuli and procedure. Eye-tracking was used to examine consumer search in online shopping under minimal restrictions. The study comprised two tasks: a shopping task and a product evaluation task. In the shopping task, participants were directed to a live online grocery website (see Figure S1) and asked to select one product from each of 14 webpages. The first two webpages served as training trials, followed by 12 main webpages presented automatically in randomized order. The product categories included milk, cheese, butter, yogurt, liver pâté, crisp bread, soft drinks, breakfast cereals, ground coffee, rye bread, meat toppings, pasta, ketchup, and cookies. Before each category page appeared, participants received the following instruction (in Danish): “Let’s switch to another category. Select ONE product that you would like to purchase, click the cart button, and then press W to continue.” Each category contained between 27 and 59 products, presented in the retailer’s default “most relevant first” order. Participants were instructed to use only the mouse for scrolling and selection and were not allowed to navigate to other categories or product pages, ensuring a consistent data structure. The task was completed by pressing the “buy” option for the chosen product. In the subsequent product evaluation task, participants provided subjective valuations for products from a subset of the previously viewed categories. Immediately after the shopping task, a custom Python script scraped product images, x–y coordinates, and webpage screenshots for each participant.⁴ This procedure ensured that the evaluation task reflected the exact products and layouts encountered during shopping, including temporary features such as stock status or special offers. Participants were instructed to select at least seven categories they frequently purchased and to rate all products within those categories on a scale from 1 to 100. During the evaluation task, the categories, as well as the products within categories, were presented in a randomized order. The seven most frequently selected categories were rye bread ($N = 95$), milk ($N = 92$), butter ($N = 90$), cheese ($N = 86$), breakfast cereals ($N = 85$), pasta ($N = 83$), and meat toppings ($N = 79$).

Hardware and software. The Aurora smart eye-tracker with a sample rate of 120 Hz was used to record eye movements. The online shopping task was performed on an HP EliteBook 850 G7 15.6” with a display resolution 1920 × 1080 using iMotions v. 9.3 software which recorded when participants scrolled up, scrolled down, and clicked the buy button for the final product choice. The distance from the participant’s eyes to the monitor was 60–65 cm. Data for the second task was gathered using PsychoPy software, version 2022.2. All statistical analyses were conducted using RStudio, version 2023.09.1+494.

Data pre-processing. Raw eye movement data, including x–y fixation coordinates, were exported from iMotions. Fixations were identified using the velocity-threshold identification (I-VT) algorithm, which distinguishes fixations from saccades based on eye-movement speed; fixations with zero duration were excluded. Because participants scrolled freely on a live website, screen-based gaze coordinates were transformed into webpage coordinates using the *eyeScrollR* package (Larigaldie et al., 2024), which deterministically remaps gaze data based on mouse-scroll information. Areas of interest (AOIs) were then automatically defined using product-specific webpage coordinates unique to each category. To avoid misattributing fixations to neighboring products, each AOI corresponded exactly to a single product box (Orquin et al., 2016). The terms AOI and product are used interchangeably. This process yielded a dataset containing corrected fixation coordinates, fixation durations, fixation counts per AOI, scroll events, and final choices.

Classification of search events. Following prior research on behavioral decision making (Glöckner & Herbold, 2011; Lohse & Johnson, 1996; Payne et al., 1988; Perkovic et al., 2018), we classified search into seven event types. *Systematic scanning* was defined as at least two consecutive fixations on adjacent AOIs within a row or column (up, down, left, or right), indicating search guided by webpage organization. This yielded 11,661 instances, consistent with previously observed structured search patterns in organized displays (Russo & Leclerc, 1994; Van der Lans et al., 2008). *Non-systematic scanning* comprised at least two consecutive fixations on non-adjacent AOIs, as well as single fixations not fitting other categories, reflecting less structured search. We identified 17,331 such instances. Rather than using fixation duration, which varies with stimulus properties (Rayner, 1998), we distinguished scanning from deeper processing using fixation counts. *Examination* was defined as at least two consecutive fixations within the same AOI; all fixations in such sequences were classified as examinations, yielding 64,786 events. Because attribute-level fixations were not distinguished, *comparison* was defined as at least three transitions between two AOIs (e.g., AOI1–AOI2–AOI1–AOI2), following prior sequence-based approaches (Russo & Leclerc, 1994). This resulted in 10,953 comparison events. Event types were assigned by priority to avoid overlap: comparisons had highest priority, followed by examinations, with both scanning types lowest. Thus, if fixations met multiple criteria, the higher-priority event was assigned (e.g., examination followed by comparison was coded as comparison). Examples of all four types of events are shown in Fig. 1A. The remaining three events, scrolling up, scrolling down, and choice, were automatically recorded, yielding 18,064 up-scrolls, 29,032 down-scrolls, and 1,287 choices.

Statistical methods. After classifying fixations into events, we encoded each event with a unique letter: non-systematic scanning (‘n’), systematic scanning (‘s’), examination (‘e’), comparison (‘c’), scroll down (‘d’), scroll up (‘u’), and choice (‘b’). This produced a sequence of letters for each participant and product category. For example, a sequence involving systematic scanning, examination,

⁴ Although products could in principle change availability or position between tasks, manual inspection confirmed identical products and locations across all cases.

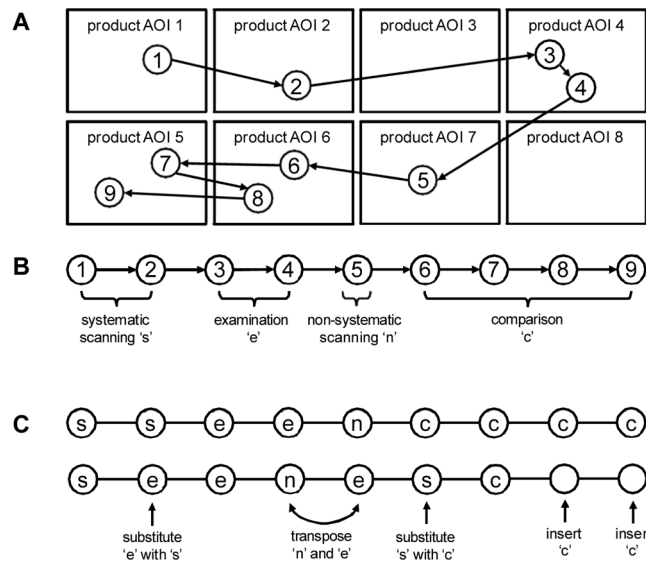


Fig. 1. A) An example of fixations on product AOIs. Circles 1 through 9 indicate the order of fixations. B) The sequence of fixations is classified into the events systematic scanning = 's', examination = 'e', non-systematic scanning 'n', and comparison = 'c', producing the sequence 'sseencccc'. C) The sequence 'sseencccc' is compared to another sequence 'seenesc' using OSA. It requires five operations (substitutions, transposition, and insertions) to make the sequences identical; that is, the distance value is 5.

non-systematic scanning, and comparison would be encoded as sseencccc (see Fig. 1B). We then applied Optimal String Alignment (OSA) to quantify similarity between sequences. OSA defines the distance between two sequences as the minimum number of operations—insertions, deletions, substitutions, or transpositions (all with equal cost)—required to transform one sequence into the other. For instance, transforming seenesc into sseencccc requires five operations, yielding an OSA distance of 5, whereas transforming sseencccc into sseencccc requires only one deletion, yielding a distance of 1 (see Fig. 1C). Larger OSA distances indicate more dissimilar search patterns, suggesting different search processes. We computed OSA distances for all pairwise combinations of sequences and clustered them using hierarchical clustering with Ward's method (*ward.D2* in the *stats* R package). The optimal number of clusters was determined using the elbow method based on the Within-Cluster Sum of Squared Errors (WSS), selecting the point at which improvements in WSS began to level off (Figure S2). This procedure grouped sequences with similar search patterns for subsequent analysis. To examine how search clusters related to stopping behavior, we modeled the percentile position (event order number) at which the chosen product was first viewed using a generalized linear mixed-effects model with a beta distribution, implemented via *glmmTMB*. The dependent variable was linearly transformed to the interval (0, 1) using the maximum event order number within each trial for each participant as the denominator. The model included cluster as a fixed effect, with random intercepts for participant and trial order number. Approximate degrees of freedom (ddf) for the fixed effects were computed using the Satterthwaite method.

2.2. Results

The final sample consisted of 112 participants (61 females), with an average age of 38.7 years ($SD = 11$). Detailed socio-demographic characteristics are provided in Table S1 in the Supplementary Materials. A total of 57 recordings (sequences) were excluded from further analysis due to instruction violations, such as unfolding additional products within the category, sorting or

Table 2

Descriptive statistics for the three clusters of event sequences in Study 1.

	Fast	Intermediate	Slow	Overall
Total number of sequences	344	613	330	1,287
Avg. sequence length	37	113	216	119
Avg. sequence time (ms)	9,212	23,013	45,856	27,034
Avg. number of seen products	11	28	40	26
Percentage (frequency) of events per sequence				
Choice	3% (1)	1% (1)	<1% (1)	1% (1)
Comparison	10% (3.7)	6% (6.5)	8% (17.3)	7% (8.5)
Examination	47% (17.3)	39% (43.6)	45% (97.3)	42% (50.3)
Non-systematic scanning	13% (4.8)	12% (13.7)	10% (22)	11% (13.5)
Systematic scanning	8% (2.8)	8% (8.7)	8% (16.3)	8% (9.1)
Scroll up	2% (0.7)	13% (14.6)	12% (26.9)	12% (14)
Scroll down	18% (6.7)	22% (24.5)	16% (35.6)	19% (22.6)

filtering products, or navigating to another category.

Single vs multiple decision processes. To determine whether participants relied on a single or multiple decision processes, we analyzed search sequences using OSA. Our dataset included 1,287 sequences from 112 participants. On average, each sequence contained 119 events, with a range from 6 to 656 events. We conducted $1,287 \times 1,287 = 1,656,369$ pairwise sequence comparisons. After removing redundant comparisons (e.g., A vs. B and B vs. A) and self-comparisons (e.g., A vs. A), we obtained 827,541 unique distance values. The mean distance value was 111, with a range from 1 to 650.

Next, based on the OSA distance values across sequences, we conducted a hierarchical cluster analysis, which identified three distinct clusters of search sequences (see Figure S2). Table 2 presents the descriptive statistics for each cluster. It is important to note that we used sequences, not participants, as the unit of analysis. This means that a single participant could have sequences belonging to multiple clusters. In fact, nearly all participants had sequences belonging to more than one cluster (see Figure S3). The propensity of Fast, Intermediate, and Slow clusters was not explained by the product category or by the order of the trials (see Table S2 and Table S3). The first cluster, "Fast" was named for its shorter sequence duration and fewer events compared to the other two clusters, Intermediate and Slow, with approximately three and six times fewer events, respectively (see Table 2). This cluster contained 344 sequences, with a mean length of 37 events, ranging from 6 to 89 events. The average duration of these sequences was 9,212 ms, with an average of 11 observed products per sequence. The second cluster, "Intermediate," was so named because its descriptive statistics fall between those of the other two clusters. It consisted of 613 sequences, with a mean length of 113 events, ranging from 51 to 196 events. The average duration of these sequences was 23,013 ms, with participants observing an average of 28 products per sequence. The third cluster, "Slow," contained 330 sequences with a mean length of 216 events, the longest among the three clusters. Sequence lengths in this cluster ranged from 133 to 656 events. These sequences had a significantly longer processing time, averaging 45,856 ms, with participants observing an average of 40 products per sequence.

A visual inspection of Fig. 2A reveals both similarities and differences in the distribution of events across the three clusters. To reduce noise due to a low number of observations, the graphs are truncated on the x-axis. The largest event number is defined as the median event number plus 1.5 times the interquartile range for each cluster. The resulting cutoffs are 79 events for the Fast cluster, 178 for the Intermediate, and 305 for the Slow. Across all clusters, examinations account for the largest share of events, while scanning and comparing occur much less frequently. As expected, choosing represents only a small proportion of events, increasing gradually toward the end of sequences.

Although the overall distribution of events remained relatively stable throughout the search sequences, some systematic temporal changes were observed. Most notably, scanning behaviors were more prevalent early in the sequence and declined rapidly thereafter, while examinations showed a slight upward trend in the Fast cluster sequences. A notable difference in scrolling behavior emerges between clusters. In the Fast cluster, there is a distinct wave of scrolling down. In contrast, in the Intermediate and Slow clusters, this initial downward scroll is followed by a subsequent wave of scrolling up. Given that participants in these clusters observed an average of 28 and 40 products, respectively, this suggests that they may have scrolled to the bottom of the page before scrolling back up.

Identifying three clusters of search sequences does not necessarily imply that participants rely on multiple decision processes. Table 2 shows that the Fast, Intermediate, and Slow sequences are structurally similar in terms of the percentage of event types, differing primarily in length. Thus, the differences in sequence length across clusters could stem from a single decision process that varies in duration. However, Fig. 2A suggests that Fast sequences are not simply truncated versions of Intermediate and Slow sequences. In the Intermediate and Slow clusters, search is characterized by a pronounced wave of downward scrolling followed by a wave of upward scrolling, consistent with participants scrolling toward the bottom of the webpage before returning upward to make a choice. In contrast, Fast sequences are dominated by downward scrolling, with upward scrolling occurring less frequently and never exceeding the probability of downward scrolling, suggesting more limited revisiting of previously viewed products.

Fixed vs sequential search. In a fixed-sample search framework, consumers decide in advance how many products to examine, choosing a sample size that maximizes expected utility by balancing the benefits of finding more preferred options against the costs of continued search (Santos et al., 2012). If search costs are similar across categories, this framework predicts that consumers should examine more products when average product value is higher. To test this prediction, we examined the correlation between the number of products viewed and the average subjective value of all products within each category. The results showed no evidence for this relationship, $r = 0.01$, $p = .703$. We also tested whether the average subjective value of all products differed across search clusters. The analysis revealed no significant differences in average subjective value among the three clusters, $F(2, 790.14) = 0.653$, $p = .521$.

Next, we examined whether search was associated with the value of the products consumers viewed. The average value of seen products differed significantly across clusters, $F(2, 801.64) = 19.090$, $p < .001$. Sequences in the Fast cluster showed the highest average values ($M = 53.70$, $SD = 19.80$), followed by the Intermediate cluster ($M = 48.36$, $SD = 17.55$), and the Slow cluster ($M = 46.58$, $SD = 17.32$). We also found a negative correlation between the number of products viewed and the average value of the products seen, $r = -.21$, $p < .001$, indicating that participants tended to search less when they encountered higher-value options. Taken together, these results are inconsistent with fixed-sample search based on optimal sample size calculations for two reasons. First, there is no evidence that participants rely on a priori expectations about the value of products within a category to determine how much to search. Second, the findings reveal a strong overall value effect: higher values among the products viewed are associated with fewer products searched. This overall value effect has been documented in binary choice settings, where higher values lead to shorter decision times (Fontanesi et al., 2019; Hunt et al., 2012; Pirrone et al., 2018; Polanía et al., 2014; Smith & Krajbich, 2018). Our results extend this phenomenon to multi-option choice, showing that higher product values reduce the number of products consumers search.

Another key distinction between fixed sample search and sequential search is how consumers finalize their choices. Under fixed sample search, consumers recall and select the best option they have encountered. In contrast, sequential search suggests that consumers stop searching either immediately upon encountering an option of particularly high value or shortly thereafter (similar to

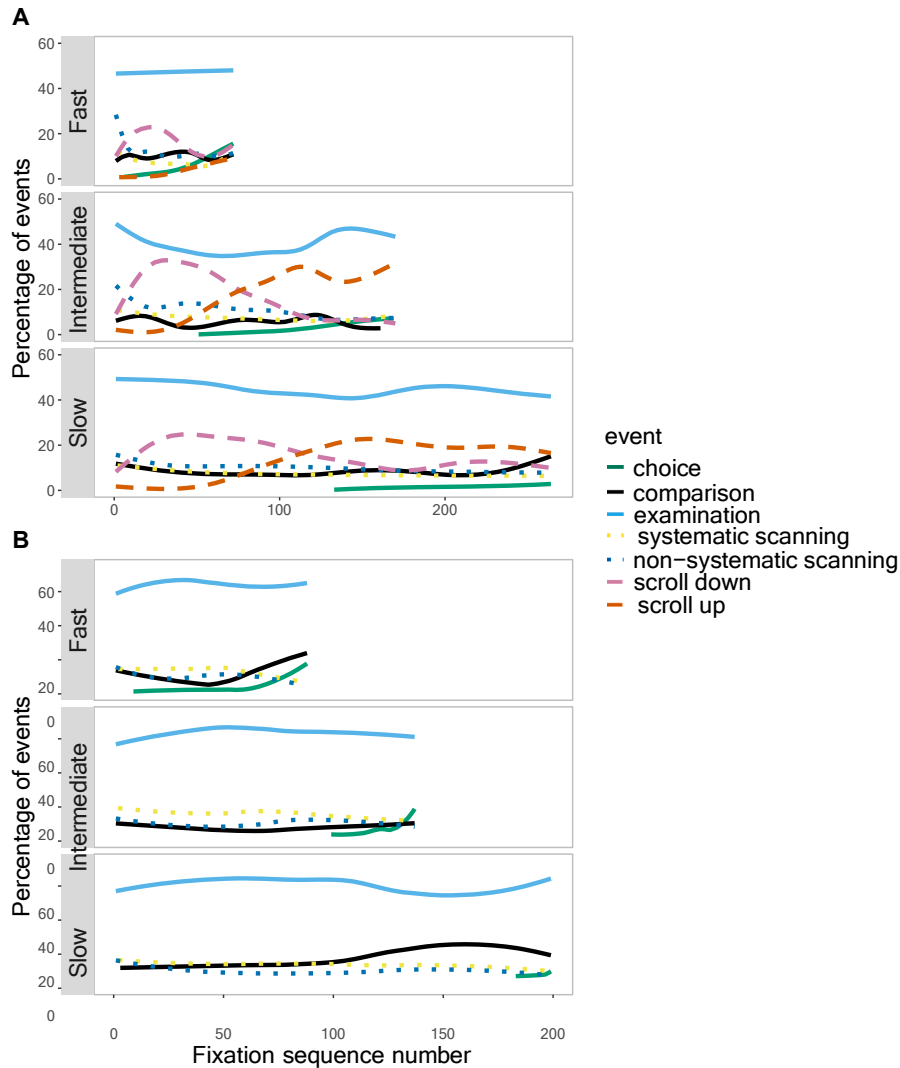


Fig. 2. A) Percentage of events within sequences for the three clusters in Study 1. The percentages of events are relatively stable during the sequences except for scrolling down and up. To remove noise due to a low number of observations, the graphs are truncated on the x-axis. B) Percentage of events in Study 2. Note that scroll events were not recorded in Study 2.

Reutskaja et al., 2011), without relying on recall. Prior studies have shown a significant increase in stopping probability after participants encounter their first satisfactory option, suggesting that recall may not be commonly used (Reutskaja & Hogarth, 2009; Stüttgen et al., 2012). To investigate this, we analyzed when the chosen product was first noticed and ran a generalized linear mixed-effects model, with cluster as a fixed effect and participant and trial order number as random intercepts. The model revealed a significant difference between the Fast and Intermediate cluster sequences, $B = -0.486$, $t = -5.655$, $ddf = 4989.7$, $p < .001$, as well as between the Fast and Slow cluster sequences, $B = -0.726$, $t = -6.818$, $ddf = 2730.8$, $p < .001$. Model convergence checks (examination of Hessian properties, optimization procedures, etc.) confirmed that despite the low variance of the trial order random effect, the model converged on the global maximum likelihood of the data, and the parameter estimates were reliable and trustworthy. The median event number when the chosen product was first noticed was 18 for the Fast cluster, 26 for the Intermediate cluster, and 50 for the Slow cluster. To interpret these findings, it is important to consider the average sequence length in each cluster: Fast sequences averaged 37 events, while Intermediate and Slow sequences contained 113 and 216 events, respectively. Fig. 3A illustrates that in Fast cluster sequences, participants were most likely to first notice the chosen product toward the very end of the sequence. In contrast, in Intermediate and Slow cluster sequences, participants tended to first notice the chosen product early in the sequence. With the exception of a subset of Fast cluster sequences, the data suggest that participants rarely stop searching immediately after first encountering their final choice. Instead, the findings indicate that recall is commonly used in Intermediate and Slow cluster sequences, whereas sequential stopping is more prevalent in Fast cluster sequences.

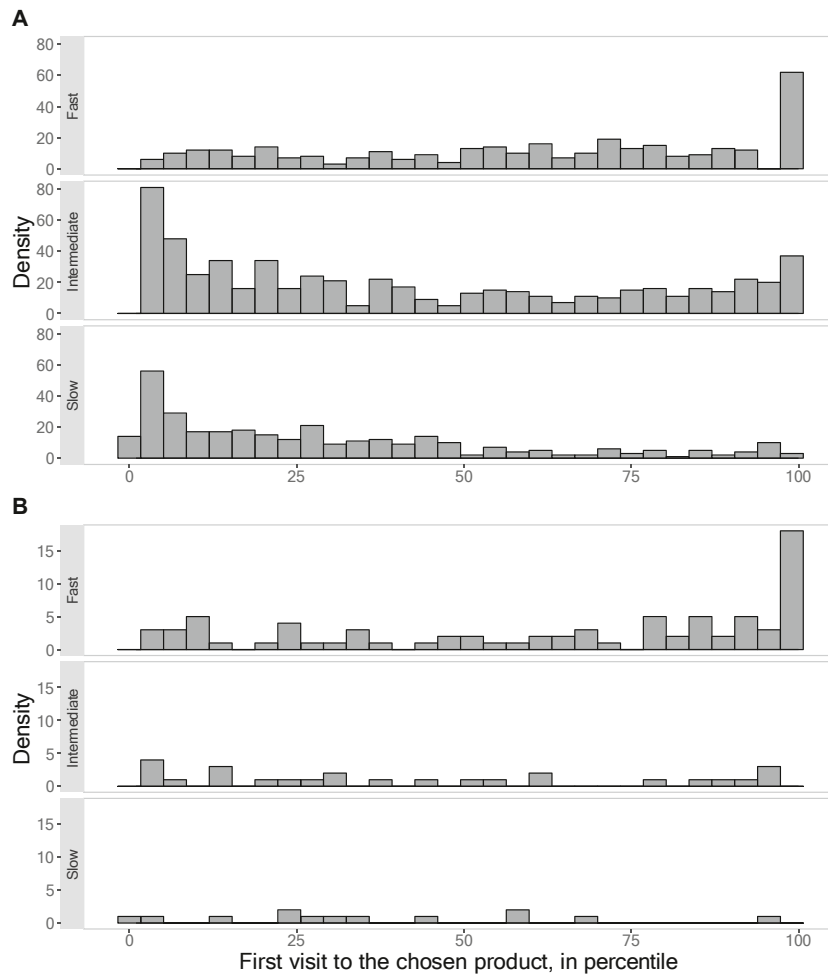


Fig. 3. A) The order number of the first visit to the chosen product, in percentile in Study 1. In Fast cluster sequences, participants are likely to first notice the chosen product at the end of the sequence while in the Intermediate and Slow cluster sequences, participants are likely to first notice the chosen product at the beginning of the sequence and still continue searching. B) The order number of the first visit to the chosen product, in percentile in Study 2.

Table 3

Percentage of products seen and mean value of products across search events and clusters for Study 1 and Study 2. Note that in Study 1 mean values are on a scale from 1 to 100 while in Study 2 on a scale from 1 to 7.

	Fast		Intermediate		Slow	
	Share	Mean Value	Share	Mean Value	Share	Mean Value
Study 1						
Choice	2%	86	2%	83	2%	81
Comparison	5%	66	7%	63	8%	55
Examination	10%	57	28%	52	45%	49
Scanning (sys)	8%	49	18%	45	28%	44
Scanning (non-sys)	11%	49	28%	46	36%	45
Non-Visited	80%	42	44%	44	24%	42
Study 2						
Choice	2%	4.9	4%	5.8	2%	5.5
Comparison	3%	4.5	6%	4.5	13%	4.5
Examination	9%	3.8	31%	3.7	35%	3.7
Scanning (sys)	8%	3.7	28%	3.8	27%	3.8
Scanning (non-sys)	7%	3.7	20%	3.5	20%	3.8
Non-Visited	71%	3.9	12%	4.0	3%	4.8

Consideration set formation. Next, we examined the relationship between events, choices, and subjective values. The chance-level probability of selecting any product across categories was 2.3%, increasing to 3.8% if at least one event occurred. The distribution of events varied significantly between chosen and non-chosen products. For instance, among chosen products, 26% were systematically scanned and 36% were non-systematically scanned, compared to only 16% and 24%, respectively, for non-chosen products. Nearly all chosen products (92%) were examined, whereas only 25% of non-chosen products were examined. Similarly, 28% of chosen products were compared, in contrast to just 3% of non-chosen products. Across all clusters, the subjective value of chosen products averaged 84 (on a scale of 100), followed by 60 for compared products, 51 for examined products, 46 for non-systematically scanned products, 45 for systematically scanned products, and 43 for non-visited products. The proportion of products scanned, examined, or compared varied across clusters. Fast sequences had the highest percentage of non-visited products, while Slow sequences exhibited the largest percentage of examined products. In general, the Slow cluster demonstrated a more thorough search strategy, scanning, examining, and comparing significantly more products than the Intermediate and Fast clusters.

Table 3 illustrates the relationship between the percentage of products that were non-visited, scanned, examined, and compared, and the average subjective value of products for each event type. Although these percentages vary across clusters, the table reveals key similarities. Consistent with consideration set theories (Hauser, 2014; Shocker et al., 1991), the results suggest that search operates as a screening process in which consumers allocate shallow attention to many lower-value products while examining a smaller set of more promising options in greater depth. Table 3 largely supports this pattern and further reveals that non-visited products tend to have the lowest subjective values across clusters.

2.3. Discussion

In Study 1, we found strong evidence that participants use multiple events, including scanning, examinations, and comparisons. Using a strict criterion, requiring at least three transitions between two products to classify a comparison, we nevertheless found that search sequences contained 6–10 % comparisons. To analyze search, we performed OSA and hierarchical clustering, identifying three distinct patterns: Fast, Intermediate, and Slow. Fast sequences had fewer events, shorter durations, and often resulted in participants selecting a product near the top of the webpage. Slow and Intermediate sequences involved longer searches and choices of lower-positioned products. Consumer heterogeneity, product categories, and category order did not explain these clusters. Instead, subjective product value emerged as a key correlate, with higher overall values of viewed products associated with less extensive search. Finally, we found evidence of a screening process: participants initially scanned many products of low or average value, examined a smaller subset of higher-value products, and compared only a few of the highest-value options.

The study raises an important question about causality: Does the overall value of products drive search patterns, or does search influence the subjective value of products? In Study 1, subjective values were measured after the search task, meaning that search patterns may have influenced value estimates. Study 1 had other limitations, such as the lack of incentivized choices and the inability of participants to defer their choice which we address in Study 2.

3. Study 2: reanalysis of Lu And Hutchinson, 2024

To overcome the limitations of Study 1 regarding causality, incentivized choice, and the possibility of choice deferral, we reanalyzed data from Experiment 2 in Lu & Hutchinson, 2024.⁵ Their study was an incentive-compatible experiment that manipulated the layout of an online clothing retailer based on subjective product values. The experimental conditions, best-first, worst-first, and random sorting, enabled us to test whether subjective values have a causal effect on search patterns.

3.1. Method

The sample consisted of female undergraduate participants, each compensated \$10 USD for browsing the online store of the clothing retailer Forever 21. Participants completed a prescreening survey five to seven days before the eye-tracking task, in which they rated their liking for 60 Forever 21 products on a seven-point Likert scale and ranked products within each rating level to resolve ties. These personalized preference ratings were used to construct individualized landing pages for the experimental conditions.

During the eye-tracking task, participants accessed a modified Forever 21 landing page and were given five minutes to shop. They were incentivized with a 1-in-20 chance to receive one item from their cart (valued up to \$50 USD) or the remaining budget in cash; all displayed products were within budget. The landing page included a category area of interest (AOI) and a grid of 40 product AOIs selected from the prescreened set. A total of 132 participants were randomly assigned to one of three layout conditions: random ($N = 41$), worst-first ($N = 45$), or best-first ($N = 46$). In the latter two conditions, product order was personalized based on prescreening rankings. Participants were unaware of their experimental condition.

Eye movements were recorded using SensoMotoric Instruments and mapped to webpage AOIs using proprietary software. Mouse scroll events were not recorded. Although the original experiment was designed to test counterfactual simulations (Lu & Hutchinson, 2024), our focus was on the experimental manipulation of product order based on subjective preferences and its implications for consumer search. We applied our event methodology to assess whether search clusters identified in observational data could be

⁵ The dataset and original code provided by Lu and Hutchinson (2024) are publicly available at <https://services.informs.org/dataset/mnsc/download.php?doi=mnsc.2022.02983>.

replicated in a more controlled experimental setting.

The experimental design shares key features with Study 1, including the use of a live retail website, remapped gaze fixations, and subjective product ratings. Differences in data structure (e.g., absence of scroll events, between-subjects design, and fewer fixations) required minor adaptations to transform fixations into events. Despite these differences, the dataset was sufficiently compatible to permit re-analysis using the same classification procedures, statistical methods, and visualization techniques as in Study 1.

3.2. Results

The initial sample consisted of 132 female undergraduate participants and included a total of 12,412 fixations. Of these, 657 fixations on the category AOI were excluded to avoid misclassification. Six participants were excluded due to an insufficient number of fixations, resulting in a final sample of 11,755 fixations from 126 participants.

Single vs. multiple decision processes. We analyzed 126 search sequences using OSA. Sequences contained an average of 93 events (range: 9–428). A total of 15,876 pairwise comparisons were computed, yielding 7,875 unique distance values (mean = 82, range = 2–419). Hierarchical clustering identified three distinct clusters. As shown in Table 4, the Fast cluster comprised 86 sequences (mean length = 58 events, range = 9–135) with an average of 18 observed products. The Intermediate cluster included 26 sequences (mean = 134 events, range = 99–181) with 37 observed products, and the Slow cluster contained 14 sequences (mean = 232 events, range = 183–428) with 39 observed products. Panel B of Fig. 2 shows that the distribution of events across clusters closely mirrors Study 1. To reduce noise, the x-axis is truncated at the median plus 1.5 interquartile ranges for each cluster.

Fixed vs. sequential search. We next examined whether Study 2 replicates the overall value effect observed in Study 1. As an initial test, the number of products viewed was not correlated with the average subjective value of all products, $r = 0.03$, $p = .123$. In contrast, the average value of viewed products differed significantly across layout conditions, $F(2, 3015) = 102.7$, $p < .001$, with the highest values in the best-first condition ($M = 4.34$, $SD = 1.88$), followed by the random ($M = 3.81$, $SD = 1.88$) and worst-first ($M = 3.21$, $SD = 1.72$) conditions. Across conditions, the number of products viewed was moderately and negatively correlated with the average value of viewed products, $r = -.51$, $p < .001$. This overall value effect was strongest in the best-first layout, $r = -.84$, $p < .001$, followed by the random layout, $r = -.43$, $p = .007$. In the worst-first layout, the relationship reversed, $r = 0.38$, $p = .012$. To further illustrate the similarity between studies, we constructed a comparable dataset from Study 1 by binning sequences into lower, middle, and upper thirds based on average subjective value, corresponding to the worst-first, random, and best-first conditions. Fig. 4A shows that in Study 1 the number of Fast sequences increases with overall value, whereas Intermediate and Slow sequences decrease. Fig. 4B shows an analogous pattern in Study 2. We further examined when participants first noticed the chosen product. The median event number of first notice was 28 for the Fast cluster, 43 for the Intermediate cluster, and 63 for the Slow cluster. After standardizing by sequence length, Fig. 3 shows that Fast sequences are most likely to involve choosing the last-noticed product, consistent with sequential search. No clear pattern emerged for the Intermediate and Slow clusters, likely due to smaller sample sizes.

Consideration set formation. Finally, we examined the relationship between events, choices, and subjective values. Across all clusters, the average subjective value of chosen products was 5.3 (out of 7), compared to 4.5 for compared products, 3.8 for examined products, 3.8 for systematically scanned products, 3.7 for non-systematically scanned products, and 3.9 for non-visited products. Search depth varied markedly across clusters. Fast sequences had the highest proportion of non-visited products (71%), whereas Slow sequences had the lowest (3%) and exhibited significantly more scanning, examination, and comparison than the other clusters. Table 3 summarizes the relationship between event type, product inclusion, and subjective value across clusters. Consistent with consideration set theories (Hauser, 2014; Shocker et al., 1991), search reflects a screening process in which lower-value products are scanned, higher-value products are examined, and only the highest-value products are compared. Unlike Study 1, however, non-visited products in Study 2 did not exhibit the lowest subjective values.

3.3. Discussion

Study 2 reanalyzes data collected by an independent research team to address several limitations of Study 1, including concerns about causality, incentivization, the timing of subjective preference measurement, and the possibility of deferring choice during search. Despite methodological differences, most notably the absence of mouse-scroll data and differences in eye-tracking pre-processing, Study 2 replicates the key findings from Study 1. Consistent with Study 1, we identify three distinct clusters of search patterns and show that variation in these patterns is systematically related to product value. The experimental manipulation of product order supports the causal interpretation that encountering higher-valued products earlier in the search process increases the likelihood of Fast search sequences. At the same time, the results indicate the coexistence of sequential and fixed-sample search processes, as well as screening behavior, reinforcing the conclusion that consumer search is best understood as a context-dependent screening process.

4. General discussion

This research advances understanding of consumer search in online retailing by introducing a unified, behaviorally grounded framework that conceptualizes search as a screening process. Using a data-driven approach, we integrate eye-tracking, mouse-scroll behavior, subjective product values, and choices into sequences of discrete events (e.g., scanning, examination, comparison), which are analyzed using Optimal String Alignment (OSA) and clustering. Across two studies, this approach reveals robust and replicable patterns in how consumers search, evaluate, and select products online.

Across both studies, consumers engaged in multiple types of events, including systematic and non-systematic scanning,

Table 4

Descriptive statistics for the three clusters of event sequences in Study 2.

	Fast	Intermediate	Slow	Overall
Total number of sequences	86	26	14	126
Avg. sequence length	58	134	232	93
Avg. number of seen products	18	37	39	24
Percentage (frequency) of events per sequence				
Choice	2% (1)	1% (1)	<1% (1)	1% (1)
Comparison	11% (7.2)	7% (10)	16% (40.2)	11% (11.5)
Examination	64% (40.9)	67% (90.2)	66% (163.2)	65% (65.5)
Non-systematic scanning	10% (6.5)	9% (12.6)	6% (16.1)	9% (8.9)
Systematic scanning	13% (8.2)	15% (20.4)	12% (28.8)	13% (13.2)

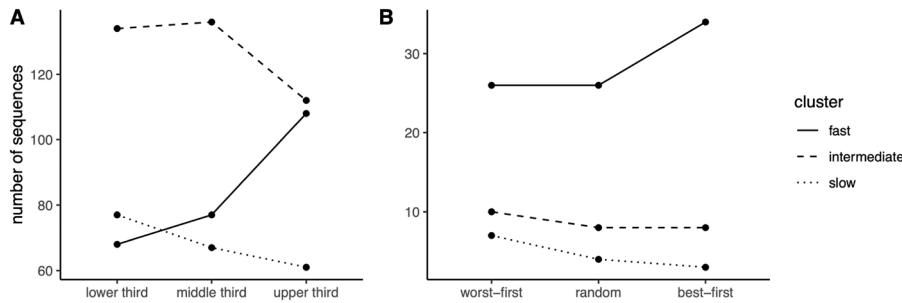


Fig. 4. A) Number of sequences in Study 1 as a function of cluster and overall value (lower, middle, and upper third), B) Number of sequences in Study 2 as a function of cluster and overall value (worst-first, random, best-first). Note that the total number of sequences in Study 1 is lower than in the other analyses because not all categories were selected by the participants for evaluation.

examinations, and comparisons. Even under conservative classification criteria, comparisons accounted for approximately 6–10% of events, underscoring their role in online decision making. Sequence analysis identified three distinct and interpretable search patterns. Fast sequences involved fewer events, shorter search whereas Slow sequences reflected more extensive search and deeper navigation on the webpage. Intermediate sequences exhibited characteristics between these extremes. Despite differences in context, data structure, and event availability, Study 2 replicated the same cluster structure, supporting the robustness of these search patterns. The difference in cluster size distributions between Study 1 and Study 2 could be due to the shopping contexts (grocery versus clothing) which has been shown to influence visual search strategies and overall behavior (e.g., Terblanche & Boshoff, 2004).

Variation in search was systematically related to subjective product value rather than to consumer characteristics, product categories, or category order. In both studies, higher average subjective value of viewed products was associated with less extensive search and a greater likelihood of Fast sequences. Study 2 further provided causal evidence that manipulating product order based on subjective value directly influences search. Analyses of stopping behavior showed that Fast sequences were more likely to end shortly after encountering the chosen product, whereas Intermediate and Slow sequences often involved continued search and eventual return to a previously seen high-value option.

Finally, analyses of events and subjective values provided consistent evidence of screening. Across both studies, consumers initially scanned many lower-value products, examined a smaller subset of higher-value products, and compared only a few top-valued options prior to choice. Although scanning decreased and comparisons increased slightly over time, search did not unfold in strictly separated stages, suggesting a flexible rather than rigidly sequential screening process.

4.1. Theoretical implications

Rather than resolving isolated theoretical debates, our findings advance a unified understanding of consumer search in online retailing. We found robust evidence that participants relied on various events, including scanning, examinations, and comparisons as predicted by multi-process models (e.g., Gigerenzer & Todd, 1999; Payne et al., 1988; Schoemann et al., 2019). In comparison, single-process models assume a homogeneous search process consisting primarily of scanning or examining products (e.g., Honka et al., 2024; Krajbich et al., 2010; Lu & Hutchinson, 2024). Furthermore, cluster analysis identified three distinct search patterns, Fast, Intermediate, and Slow, highlighting meaningful differences in consumer search.

Participants appear to use both sequential and fixed-sample search strategies. Evidence for sequential search includes the overall value effect and the tendency of participants in Fast cluster sequences to stop searching immediately or shortly after encountering their chosen product. This suggests a satisficing stopping rule (Simon, 1979), where participants search until they find a product that meets or exceeds their aspiration level — the probability of which increases with higher overall values. However, in Intermediate and Slow cluster sequences, participants often encountered their chosen product early but continued searching before ultimately returning to a previously seen high-value product. This stopping behavior could be explained by sequential search under two conditions: (1) if search

was exhaustive, requiring participants to stop and recall the best-seen product, or (2) if the maximum realized value exceeded the reservation utilities of the remaining unseen products (Honka et al., 2024; Weitzman, 1978). Given the number of seen products in Intermediate and Slow cluster sequences, we can rule out exhaustive search (see Table 2). If participants relied on reservation utilities, these could be informed either by prior expectations about the distribution of values in a product category or through observations of products. We found no evidence that prior expectations influenced search (i.e., the average value of all products in a category did not influence search). If participants, on the other hand, based their expectations on observation, higher observed values should lead to higher expected values. This in turn should result in higher reservation utilities (or higher aspiration levels if one assumes that these are adjusted) for the remaining unobserved products, resulting in more extensive search (Stüttgen et al., 2012; Weitzman, 1978). However, we observe the opposite pattern: higher overall observed values lead to less search. For the same reason, we can also rule out a fixed-sample search strategy based on computing the marginal value of search against search costs (Santos et al., 2012; Stigler, 1961). However, participants may have followed a fixed-sample search rule, such as stopping after reaching a time or cost limit (Bhatia & Mullett, 2016; Jessup et al., 2009). Variability in this limit would create variability in search extent, which, depending on the product sorting in Study 2 (best-first or worst-first), should produce either a positive or negative overall value effect. We observe exactly this pattern: When products were sorted from best to worst, more extensive search led consumers to notice lower-value products, while the opposite occurred when products were sorted from worst to best.

Our findings indicate that participants engage in screening behavior consistent with consideration set formation (Hauser & Wernerfelt, 1990; Roberts & Lattin, 1991; Shocker et al., 1991; Wright & Barbour, 1977). However, unlike standard models that suggest consideration sets are formed in distinct consecutive stages, we found that all events, scanning, examinations, and comparisons, occurred throughout the search process. While there was a slight decrease in scanning and an increase in comparisons over time, no strict stage-wise progression was observed.

Overall, the findings suggest that consumer search is best understood as a screening process in which different events help identify high-value products. Initially, consumers scan a large number of products which is a fast but relatively imprecise method of assessing value. The most promising products are then examined in greater depth to reduce uncertainty, and if necessary, the best of these are compared to establish dominance relationships. This screening process has two stopping rules: a) terminate search when a satisfactory product is encountered or b) terminate search when a time limit is reached. The two stopping rules make the process context dependent. When overall product value is high, consumers quickly converge on one or more satisfactory products and stop searching immediately or shortly after encountering a suitable product. However, when overall value is low, the process may not yield a satisfactory product and consumers continue searching for a predetermined duration before stopping. If consumers are compelled to make a choice (as in Study 1) they will return to a previously encountered high-value product otherwise they may decide to defer choice (as in Study 2).

4.2. Practical implications

Our findings suggest that consumer search functions as a screening process, and online retailers have various tools at their disposal to support this process. While most retailers already implement default product sorting (e.g., prioritizing the most popular items), other features could play a more significant role. For instance, filtering options allow consumers to exclude products with undesirable attributes, while wishlist features enable them to compile a set of highly preferred options. Additionally, any visual design that facilitates product comparisons, such as side-by-side displays or highlighting key differences, may further enhance the screening process.

Beyond supporting search, online retailers can also shape consumer search and choices. The way products are sorted influences how consumers perceive the distribution of product values. If the goal is to reduce search, retailers should present the most preferred products at the top of the webpage. Minimizing search may also reduce price sensitivity, as consumers who browse less are less likely to encounter cheaper alternatives (Seiler & Pinna, 2017). Retailers may also adjust search depending on the shopping context. For instance, consumers shopping on weekdays under time constraints may benefit from less search, whereas those shopping for leisure on weekends may find value in a more extensive search experience that exposes them to a broader range of options (Chung et al., 2023).

Furthermore, online retailers can leverage search data to make inferences about consumer preferences and engagement with product assortments. When available, eye-tracking data can be used to characterize how consumers allocate attention through scanning, examination, and comparison. In our studies, products that were examined or compared tended to have higher subjective value than those that were only briefly scanned or ignored, suggesting that visual attention patterns provide informative signals about relative preferences during search. Mouse scroll data, which are routinely collected through website analytics, can further inform retailers about search intensity and the overall subjective value of assortments. Our findings show that scroll depth (i.e., how far consumers scroll down a page) is systematically related to search patterns: deeper scrolling is associated with Intermediate and Slow search sequences, which in turn are linked to lower overall values among the products encountered. Conversely, limited scrolling is associated with Fast search sequences, which are more likely when overall product values are higher. Together, these behavioral measures offer retailers scalable ways to infer how consumers experience product assortments and how effectively current designs support search and choice.

4.3. Limitations and future research

Although Study 2 replicated the main findings of Study 1 using independent experimental data and addressed endogeneity concerns, the two studies differed in design and retail context, which likely contributed to differences in the distribution of search clusters. In particular, the Fast cluster dominated in the clothing context of Study 2, whereas the Intermediate cluster was most prevalent in the

grocery context of Study 1. Rather than constituting a limitation, these differences point to an important opportunity for future research. Retailer context, such as grocery versus apparel, may systematically shape the prevalence of Fast, Intermediate, and Slow search patterns due to differences in involvement, goal orientation, and the experiential versus functional nature of the shopping task. Future research could build on this insight by examining how cluster distributions vary across retail contexts and how these differences moderate the effectiveness of product organization, sorting, and interface design. For example, retailers with a higher proportion of Fast searchers may benefit from prominently displaying high-value products early, whereas retailers with more Intermediate or Slow searchers may benefit from richer comparison tools, filtering options, or extended assortments. Follow-up studies could experimentally test targeted interventions that adapt website design to the dominant search patterns within a given category or retailer type.

More broadly, future work should examine consumer search in more naturalistic settings and across a wider range of product categories, including luxury and high-involvement goods. Additional research is also needed to refine understanding of aspiration levels, time limits, and the role of comparisons in shaping preferences.

Investigating whether early screening relies on overt versus covert attention, and whether comparisons causally enhance product value, would further deepen insight into how consumers search and decide in online retail environments.

5. Conclusion

This research demonstrates that consumer search in online retailing operates as a screening process shaped by product values and interface design. By integrating behavioral data into a unified framework, we provide both theoretical clarity and practical guidance for designing more effective digital retail environments. Our findings highlight how retailers can support and influence consumer search through assortment design, product ordering, and decision-support tools, while pointing to promising avenues for future research on online consumer behavior.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the authors used ChatGPT and Grammarly to improve the article's readability and language. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Funding

This research was supported by the Independent Research Fund Denmark, grant number 8046–00014A and the Icelandic Research Fund, grant number 2511673–051.

Research data and code

All data sets, as well as the programming code used in the study, are available at [OSF](https://osf.io/54b92/).⁶

Declaration of competing interest

The authors declare no conflict of interest.

Acknowledgments

We thank Joy Lu and J. Wesley Hutchinson for permission to re-analyze their data, which enabled us to address limitations of Study 1 and assess the robustness of our findings. We are also grateful to the anonymous reviewers for their careful reading and constructive feedback, which substantially improved the quality of this manuscript. We further thank Dr. Denis Federiakin for his valuable comments on the statistical analyses.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.jretai.2026.05.004](https://doi.org/10.1016/j.jretai.2026.05.004).

References

Bhatia, S., & Mullett, T. L. (2016). The dynamics of deferred decision. *Cognitive Psychology*, 86, 112–151.

⁶ <https://osf.io/54b92/overview>

- Boardman, R., & McCormick, H. (2022). Attention and behaviour on fashion retail websites: An eye-tracking study. *Information Technology & People*, 35(7), 2219–2240.
- Bremer, L., Heitmann, M., & Schreiner, T. F. (2017). When and how to infer heuristic consideration set rules of consumers. *International Journal of Research in Marketing*, 34(2), 516–535.
- Bussemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review*, 100(3), 432.
- Chandon, P., Hutchinson, J. W., Bradlow, E. T., & Young, S. H. (2009). Does in-store marketing work? effects of the number and position of shelf facings on brand attention and evaluation at the point of purchase. *Journal of Marketing*, 73(6), 1–17.
- Chen, M., Burke, R. R., Hui, S. K., & Leykin, A. (2021). Understanding lateral and vertical biases in consumer attention: An in-store ambulatory eye-tracking study. *Journal of Marketing Research*, 58(6), 1120–1141.
- Chocarro, R., Cortiñas, M., & Villanueva, A. (2022). Attention to product images in an online retailing store: An eye-tracking study considering consumer goals and type of product. *Journal of Electronic Commerce Research*, 23(4), 257–281.
- Chung, J., Lee, L., Lehmann, D. R., & Tsai, C. I. (2023). Spending windfall (“found”) time on hedonic versus utilitarian activities. *Journal of Consumer Research*, 49(6), 1118–1139.
- Damerau, F. J. (1964). A technique for computer detection and correction of spelling errors. *Communications of the ACM*, 7(3), 171–176.
- Edenbrandt, A. K., Lagerkvist, C.-J., Lüken, M., & Orquin, J. L. (2022). Seen but not considered? awareness and consideration in choice analysis. *Journal of Choice Modelling*, 45, Article 100375.
- Fontanesi, L., Gluth, S., Spektor, M. S., & Rieskamp, J. (2019). A reinforcement learning diffusion decision model for value-based decisions. *Psychonomic Bulletin & Review*, 26(4), 1099–1121.
- Gauri, D. K., Jindal, R. P., Ratchford, B., Fox, E., Bhatnagar, A., Pandey, A., Navallo, J. R., Fogarty, J., Carr, S., & Howerton, E. (2021). Evolution of retail formats: Past, present, and future. *Journal of Retailing*, 97(1), 42–61.
- Gigerenzer, G., & Todd, P. M. (1999). *Simple heuristics that make us smart*. USA: Oxford University Press.
- Gilbride, T. J., & Allenby, G. M. (2004). A choice model with conjunctive, disjunctive, and compensatory screening rules. *Marketing science*, 23(3), 391–406.
- Glöckner, A., & Herbold, A.-K. (2011). An eye-tracking study on information processing in risky decisions: Evidence for compensatory strategies based on automatic processes. *Journal of Behavioral Decision Making*, 24(1), 71–98.
- Groening, C., Wiggins, J., & Raoofpanah, I. (2021). Wish list thinking: The quasi-endowment effect’s impact on online wish lists outcomes. *Journal of Consumer Behaviour*, 20(2), 412–425.
- Hauser, J. R. (2014). Consideration-set heuristics. *Journal of Business Research*, 67(8), 1688–1699.
- Hauser, J. R., & Wernerfelt, B. (1990). An evaluation cost model of consideration sets. *Journal of consumer research*, 16(4), 393–408.
- Honka, E., Hortaçsu, A., & Wildenbeest, M. (2019). Empirical search and consideration sets. *Handbook of the Economics of Marketing*, 193–257.
- Honka, E., Seiler, S., & Ursu, R. (2024). Consumer search: What can we learn from pre-purchase data? *Journal of Retailing*, 100(1), 114–129.
- Hunt, L. T., Kolling, N., Soltani, A., Woolrich, M. W., Rushworth, M. F., & Behrens, T. E. (2012). Mechanisms underlying cortical activity during value-guided choice. *Nature Neuroscience*, 15(3), 470–476.
- Hwang, Y. M., & Lee, K. C. (2018). Using an eye-tracking approach to explore gender differences in visual attention and shopping attitudes in an online shopping environment. *International Journal of Human-Computer Interaction*, 34(1), 15–24.
- Jessup, R. K., Veinott, E. S., Todd, P. M., & Bussemeyer, J. R. (2009). Leaving the store empty-handed: Testing explanations for the too-much-choice effect using decision field theory. *Psychology & Marketing*, 26(3), 299–320.
- Kahn, B. E. (2017). Using visual design to improve customer perceptions of online assortments. *Journal of Retailing*, 93(1), 29–42.
- Kahneman, D. (2011). *Thinking, fast and slow*. macmillan.
- Krajibich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience*, 13(10), 1292–1298.
- Krefeld-Schwalb, A., Donkin, C., Newell, B. R., & Scheibehenne, B. (2019). Empirical comparison of the adjustable spanner and the adaptive toolbox models of choice. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 45(7), 1151.
- Krefeld-Schwalb, A., Martinovici, A., Orquin, J. L., Aribarg, A., Erdem, T., Johnson, E. J., Reeck, C., Schley, D. R., Smith, S. M., Spekenbrink, M., et al. (2025). Search in service of choice: An agenda for comparing, extending, and integrating consumer search models. *International Journal of Research in Marketing*.
- Krefeld-Schwalb, A., & Rosner, A. (2020). A new way to guide consumer’s choice: Retro-cueing alters the availability of product information in memory. *Journal of Business Research*, 111, 135–147.
- Larigaldie, N., Dreneva, A., & Orquin, J. L. (2024). Eyescroll: A software method for reproducible mapping of eye-tracking data from scrollable web pages. *Behavior Research Methods*, 1–16.
- Larsen, N. M., Sigurdsson, V., Breivik, J., & Orquin, J. L. (2020). The heterogeneity of shoppers’ supermarket behaviors based on the use of carrying equipment. *Journal of Business Research*, 108, 390–400.
- Lee, M. D., & Cummins, T. D. (2004). Evidence accumulation in decision making: Unifying the “take the best” and the “rational” models. *Psychonomic Bulletin & Review*, 11(2), 343–352.
- Liechty, J., Pieters, R., & Wedel, M. (2003). Global and local covert visual attention: Evidence from a bayesian hidden markov model. *Psychometrika*, 68(4), 519–541.
- Lohse, G. L., & Johnson, E. J. (1996). A comparison of two process tracing methods for choice tasks. *Organizational Behavior and Human Decision Processes*, 68(1), 28–43.
- Lu, J., & Hutchinson, J. W. (2024). Information search within a web page: Modeling the full sequence of eye movement decisions, subjective value updating, and first clicks. *Management Science*.
- Martinovici, A., Pieters, F., & Erdem, T. (2021). Attention trajectories predict brand choice. *Tulin, Attention Trajectories Predict Brand Choice. February 14, 2021*.
- Modi, N., & Singh, J. (2023). Understanding online consumer behavior at e-commerce portals using eye-gaze tracking. *International Journal of Human-Computer Interaction*, 39(4), 721–742.
- Orquin, J. L., Ashby, N. J., & Clarke, A. D. (2016). Areas of interest as a signal detection problem in behavioral eye-tracking research. *Journal of Behavioral Decision Making*, 29(2–3), 103–115.
- Payne, J. W., Bettman, J. R., & Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of experimental psychology: Learning, Memory, and Cognition*, 14(3), 534.
- Perkovic, S., Bown, N. J., & Kaptan, G. (2018). Systematicity of search index: A new measure for exploring information search patterns. *Journal of Behavioral Decision Making*, 31(5), 673–685.
- Pirrone, A., Wen, W., & Li, S. (2018). Single-trial dynamics explain magnitude sensitive decision making. *BMC Neuroscience*, 19, 1–10.
- Polanía, R., Krajibich, I., Grueschow, M., & Ruff, C. C. (2014). Neural oscillations and synchronization differentially support evidence accumulation in perceptual and value-based decision making. *Neuron*, 82(3), 709–720.
- Ratchford, B., Soysal, G., Zentner, A., & Gauri, D. K. (2022). Online and offline retailing: What we know and directions for future research. *Journal of Retailing*, 98(1), 152–177.
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124(3), 372.
- Reutskaja, E., & Hogarth, R. M. (2009). Satisfaction in choice as a function of the number of alternatives: When “goods satiate”. *Psychology & Marketing*, 26(3), 197–203.
- Reutskaja, E., Nagel, R., Camerer, C. F., & Rangel, A. (2011). Search dynamics in consumer choice under time pressure: An eye-tracking study. *American Economic Review*, 101(2), 900–926.
- Roberts, J. H., & Lattin, J. M. (1991). Development and testing of a model of consideration set composition. *Journal of Marketing Research*, 28(4), 429–440.
- Russo, J. E., & Leclerc, F. (1994). An eye-fixation analysis of choice processes for consumer nondurables. *Journal of Consumer Research*, 21(2), 274–290.

- Santos, B. D.I., Hortaçsu, A., & Wildenbeest, M. R. (2012). Testing models of consumer search using data on web browsing and purchasing behavior. *American Economic Review*, 102(6), 2955–2980.
- Schoemann, M., Schulte-Mecklenbeck, M., Renkewitz, F., & Scherbaum, S. (2019). Forward inference in risky choice: Mapping gaze and decision processes. *Journal of Behavioral Decision Making*, 32(5), 521–535.
- Schröter, I., Grillo, N. R., Limpak, M. K., Mestiri, B., Osthold, B., Sebti, F., & Mergenthaler, M. (2021). Webcam eye tracking for monitoring visual attention in hypothetical online shopping tasks. *Applied Sciences*, 11(19), 9281.
- Seiler, S., & Pinna, F. (2017). Estimating search benefits from path-tracking data: Measurement and determinants. *Marketing Science*, 36(4), 565–589.
- Shevlin, B. R., Smith, S. M., Hausfeld, J., & Krajbich, I. (2022). High-value decisions are fast and accurate, inconsistent with diminishing value sensitivity. *Proceedings of the National Academy of Sciences*, 119(6), Article e2101508119.
- Shocker, A. D., Ben-Akiva, M., Boccara, B., & Nedungadi, P. (1991). Consideration set influences on consumer decision-making and choice: Issues, models, and suggestions. *Marketing Letters*, 2, 181–197.
- Simon, H. A. (1979). *Models of thought*, 352. Yale university press.
- Smith, S. M., & Krajbich, I. (2018). Attention and choice across domains. *Journal of Experimental Psychology: General*, 147(12), 1810.
- Söllner, A., & Bröder, A. (2016). Toolbox or adjustable spanner? a critical comparison of two metaphors for adaptive decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 42(2), 215.
- Stigler, G. J. (1961). The economics of information. *Journal of Political Economy*, 69(3), 213–225.
- Stüttgen, P., Boatwright, P., & Monroe, R. T. (2012). A satisficing choice model. *Marketing Science*, 31(6), 878–899.
- Szocs, C., Kim, Y., Lim, M., Mera, C. A., & Biswas, D. (2023). The store of the future: Engaging customers through sensory elements, personalized atmospherics, and interpersonal interaction. *Journal of Retailing*, 99(4), 605–620.
- Terblanche, N. S., & Boshoff, C. (2004). The in-store shopping experience: A comparative study of supermarket and clothing store customers. *South African Journal of Business Management*, 35(4), 1–10.
- Thomas, A. W., Molter, F., & Krajbich, I. (2021). Uncovering the computational mechanisms underlying many-alternative choice. *Elife*, 10, Article e57012.
- Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological Review*, 79(4), 281.
- Ursu, R. M., Erdem, T., Wang, Q., & Zhang, Q. (2024). Prior information and consumer search: Evidence from eye tracking. *Management Science*.
- Van der Lans, R., Pieters, R., & Wedel, M. (2008). Eye-movement analysis of search effectiveness. *Journal of the American Statistical Association*, 103(482), 452–461.
- Weitzman, M. (1978). *Optimal search for the best alternative*, 78. Department of Energy.
- Wright, P., & Barbour, F. (1977). *Phased decision strategies: Sequels to an initial screening*. Graduate School of Business, Stanford University.