Recommenders' Influence on Buyers' Decision Process

Sylvain Castagnos
EPFL - HCI Group
IC IIF Station 14
CH-1015 Lausanne
sylvain.castagnos@epfl.ch

Nicolas Jones
EPFL - HCI Group
IC IIF Station 14
CH-1015 Lausanne
nicolas.jones@epfl.ch

Pearl Pu
EPFL - HCI Group
IC IIF Station 14
CH-1015 Lausanne
pearl.pu@epfl.ch

ABSTRACT

Online stores offer an increasingly large set of products. Interactive decision aids are becoming indispensable tools assisting users as they search for an ideal product to purchase. For an e-commerce website, adopting the correct tools can affect its survival: effective product recommender tools are increasingly recognized by online stores as effective means to sell more products; on the other hand, sites that do not employ intelligent tools will not only see poor purchase volumes but also experience less traffic because consumers are more likely to return to a site employing recommender systems.

This paper presents ongoing research in understanding the impact of various decision aids on users' interaction behaviors and their subjective perceptions of these aids. In the current experiment, we employed an eye tracker in an in-depth user study to understand the influence of recommenders on how users select items for the basket set. We collected more than 20,300 fixation data points in 3,648 areas of interest. Our studies show that while users still rely on product filtering tools, the use of recommenders is becoming more prominent in helping them construct the basket set and is monotonically increasing as time goes on.

Categories and Subject Descriptors

H.1.2 [**Models and Principles**]: User/Machine Systems—*Human Information Processing*; H.5.2 [**Information Interfaces and Presentation**]: User Interfaces—*Evaluation / methodology*.

General Terms

Human Factors, Design, Experimentation.

Keywords

Recommender Systems, Decision Process, Interaction Design, Eye-Tracker, User Study, User Modeling, Usage Patterns.

1. INTRODUCTION

The retail e-commerce sector in the US is a 210 billion dollar industry; Europe and Asia are catching up at an increasingly significant speed. A crucial technological element in such sites is an effective product search tool to help consumers identify the right items and provide them with the necessary confidence to purchase. Consider an average user who is looking to buy a digital camera through an online product comparison site. She is unlikely to have an exact idea of the manufacturer's name, the model or other technical specifications of the eventual camera and thus perform a keyword search. With a catalog containing more than 1,000 items, she will not sift through the list one by one either. Instead, she is more likely to interact with an online catalog of products with the help of a decision aid.

Haubl and Trifts [4] maintain that there are two main steps during the decision process in an online product search environment. In the first step, the active user identifies a subset of products she wants to compare (called the consideration set or the basket). We will refer to this step as *product brokering* to be consistent with concepts used in [5]. During the second step, the user compares the different features and details of these products in order to make a decision. We will refer to this step as *product comparison*. In their work, two interaction decision aides were investigated for their roles in helping users make better decisions. The first tool, a recommender agent, assists users in the initial screening of the alternatives and establishing the basket set. The second one, a comparison matrix, helps users make an in-depth feature-by-feature comparison of the items in the basket. Their empirical studies showed that the use of a recommender agent leads to a significant reduction in the number of alternatives seriously considered for purchase, and increases the quality of consumers' decisions. The constructive influence of recommenders was consequently established. However, we observed one major outdated finding in their experiment. Even though a recommender was conceptualized in their setup, its actual implementation, thus the tool that their users experienced, was provided by personalogic[1]. In the more current literature of recommenders, personalogic is considered more as a multi-criteria product filtering tool rather than a recommender. The main difference is that in the former, users obtain a set of items after they have actively specified their preferences, whereas in the latter, users get recommended items without asking for them.

In our present work, we seek to understand users' decision behaviors under the influence of both multi-criteria filtering tool (MCF) and recommender system (RS), thus continuing along the line of research initiated by Haubl and Trifts [4] and our own previous work [7]. Ho pioneered a study to measure the impact of RS on users' clicks at different decision making stages [2]. Nevertheless, the absence of selection does not mean that the RS does not play a role. We thus chose to focus on users' behavior such as task time spent on browsing, filtering the catalog, looking up product details, and taking recommendations from the system. At the same time,

---

[1] www.personalogic.com does not exist anymore

we measure users' subjective perception of the decision aids, such as satisfaction with the interface and their intention to buy and to return to the site. We report new findings based on an in-depth user study which collected on average one hour of eye movements as well as action logs from each of the 7 subjects using the Tobii eye tracking system. This data set comprised of more than 20,300 fixation points and 3,648 areas of interest traces. The results of our current work constitute a first major step towards the formalization of the product brokering process involving recommender systems within the purchase decision model.

The rest of the paper is organized as follows. We describe the experiment setup in detail and present the participant's background information. The paper continues by presenting our main research goal before reporting the results of the study. We then discuss the results and their implications. Finally, we present our conclusions and ideas for future work.

2. THE EYE TRACKING EXPERIMENT

2.1 The Material

The eye tracker used in our experiment was a `Tobii 1750` device consisting of a computer screen with embedded cameras and a software capable of capturing a person's point of gaze. After an initial calibration, the setup allows users to look at the screen in a natural way without any need for a head mount.

The two interactive decision tools being investigated are the multi-criteria filtering tool (MCF) and the critiquing-based recommender system (CRS). The product catalog in consideration is a collection of 3,500 popular perfumes of usual brands for men and women. Most users recruited in our study had a consistent level of understanding of the product domain.

The MCF tool consists of a search panel `M` and a list view `L` of all products in multiple pages. The display of the first few items is illustrated in Figure 1. Users can review all of the products in the catalog by choosing a sorting criterion, such as brand or price, and examine the entire catalog page by page. To view a subset of the products, which most users will do, a user has to select a specific item under the multi-criteria search box, such as a price range. A MCF tool allows users to apply a mixture of several criteria to narrow the product list by filtering out the undesirable products. However, users can also delete any filtering criterion, thus enlarging the available product space. When the user becomes interested in a product, she can visit the product's detailed description page `D` (Figure 1). This page also presents a column of recommendations `R` computed by the CRS on the right hand side. Most e-commerces such as www.amazon.com use a single category of recommendations, such as "Users who bought this also bought". In critiquing based recommenders, this recommendation set is further categorized into groups of products, each of which provides explicit improvements of the current product. The comparison of a single category of recommended items (thus the list view approach) versus an organized view has been extensively studied in our previous work (see for example [7]). Our research showed that users found the organized view to be more user friendly, and they perceived that less effort was expended when comparing products. Even though such organized views are not readily used in mainstream e-commerce sites, we consider them to be superior to the list view. The CRS used in our experiment organizes the recommendation set into five categories: "more popular and cheaper", or if this former category does not contain any products we use "more popular but more expensive", "same brand and cheaper" or "same brand but more expensive", "just as popular and cheaper", "same price range and just as popular", and "people who like this also like".

2.2 Experiment Procedure

Each participant in our study went through the following steps.

Step 1. The experiment's administrator (hereafter admin) first debriefs each participant on the nature of the experiment. Background information on the user is collected (age, sex, etc.).

Step 2. The admin calibrates the eye tracker based on the location of the user's pupils. He then launches the tracking session and encourages the user to explore the system before fully launching into her first task.

Step 3. The user's first task is to find and put up to three perfumes that she has never heard of or used before into the shopping list. She was informed that she could ultimately select more than three items and delete some at the end, but that she must be prepared to purchase one of them for herself. We refer to this task as *Session 1*.

Step 4. Once this task is finished, the admin launches a new tracking session with a modified task. The user is now asked to search for one perfume she would like to offer to someone, preferably of the opposite sex (reducing the potential bias as a result of product familiarity). This will be called *Session 2*.

Step 5. To conclude the study, fourteen preference questions are asked in order to assess the user's overall perception of the system.

The preference questions are statements to which a user can indicate her level of agreement on a five-point Likert scale, ranging from -2 to $+2$, where -2 means "strongly disagree" and $+2$ is "strongly agree". The post-stage questions' results are shown in Figure 4. The questions were asked in random order, in order to eliminate ordering bias.

2.3 Participants' Background

The study was carried out over a period of one week and proposed an incentive to ensure that users behaved candidly. A 100 Swiss franc voucher was given out in a lottery to the winning user. to purchase one of her chosen perfumes. A total of 7 volunteers were recruited as participants. They were from 4 different countries (France, Germany, Serbia and Switzerland), with different professions (student, worker, Ph.D. student) and educational backgrounds (high school, graduate school).

For all of them, *fragrance* was an important feature when describing a perfume. Other important aspects include price, brand, quantity and design. All users had strong web experience, although their online shopping experience remained limited to standard items such as books, music or travel.

Background questions surveyed users' predisposition towards perfumes. Five participants said they bought perfumes about once a year, one a few times a year and one nearly monthly. When questioned about how they discovered new perfumes, 57% of users admitted that they encountered problems, mainly because they are not able to explicitly describe their perfume-preferences (lack of knowledge and difficulty of describing a fragrance). 70% said that they preferred to just test perfumes alone in a shop. 40% of them would also accept suggestions from friends. Most users also told us that they were prepared to reveal information such as previously liked & disliked perfumes, and price, in order to obtain recommendations. Interestingly, all users were prepared to reveal information about smells that they like, but explained that it is difficult to describe smells and hence relied on other aspects.

3. EXPERIMENT OBJECTIVES

We expect that the CRS will have an impact on two general aspects of consumer decision making in an online shopping environment: (1) choice strategies and (2) consideration sets. *Choice strategies* can be thought of as methods (sequences of operations) for searching through the decision problem space [6]. As explained
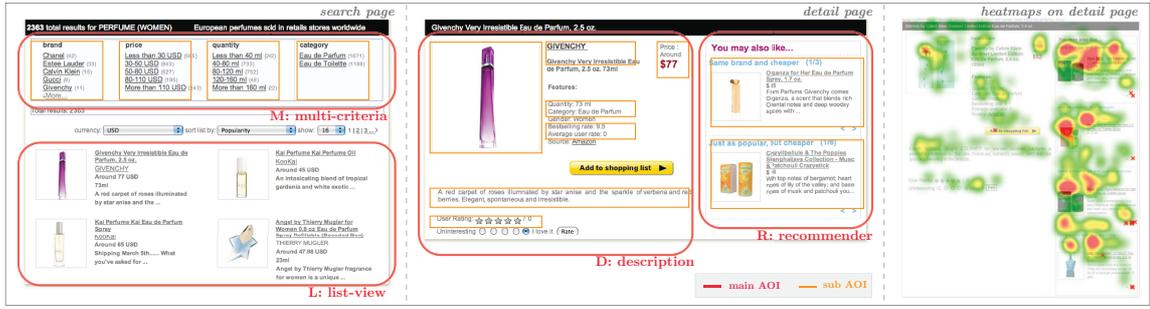
Figure 1: Snapshots of the main interfaces (including AOIs) and an example of resulting heatmap.

in Section 1, other research shows that these strategies can be broken down into two steps: discovery of their needs and comparison of products which fit to their criteria. *Consideration sets* are conceptualized as sets of alternatives (one for each user) that consumers seriously consider for purchase [1]. In this paper, since our framework does not include a comparative tool of items' features, we will suppose that an item is in the consideration set as soon as the user focuses attention on a possible selection by intently looking at it.

> RESEARCH QUESTIONS.
> *What is the influence of the recommender system? Is this influence continuously increasing over time at the product brokering stage?*

We expect that the influence of the recommender system will help avoid dividing choice strategies into two steps, since it helps users to both: (1) refine criteria to use with lexicographic ordering and (2) find valuable alternatives to a product in the consideration set. Moreover, this influence should last from one product search to another made consecutively by a user.

4. DATA ANALYSIS AND RESULTS

We collected gaze plots and heat maps, but strongly relied on *Areas Of Interest* (hereafter AOIs) as they are an objective data about users' actions on the website.

Using the eye tracking system, we aimed at measuring how users' interest for the different parts of the website evolves over time. We recorded an average of 1,450 fixations per user. We defined areas of interest (AOIs) spread in two separate types of pages: the search pages, and the detail pages (cf. Figure 1). Table 1 synthesizes the average number of pages seen and the average session times.

Table 1: Statistics of sessions for the overall set of users

	Session 1	Session 2	Both
Average Number of Search Pages	9.71	4.86	14.57
Average Number of Detail Pages	10.86	7.57	18.43
Average Session Time (Minutes)	12.72	6.18	9.45

Then, we computed the total fixation durations for each user $\in U$ on the different AOIs over time t. We paid attention to durations for four variables: the MCF tool M, the lexicographic ordered list L (list-view), the description of perfumes D and the CRS agent R. Usages of the MCF tool and CRS agent over time for the overall set of users U are made explicit in Figure 2. We summed the cumulative fixation durations of these two AOIs. Consequently, the curves flatten when the users stop looking at the corresponding AOIs.

We defined the variables $M_u(t)$ and $R_u(t)$ as the fixation durations of user u respectively on M and R in $(t-1; t]$. These have been
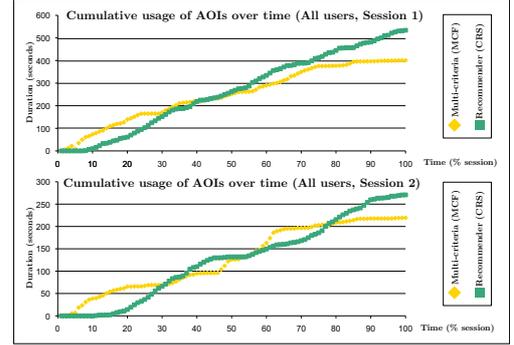

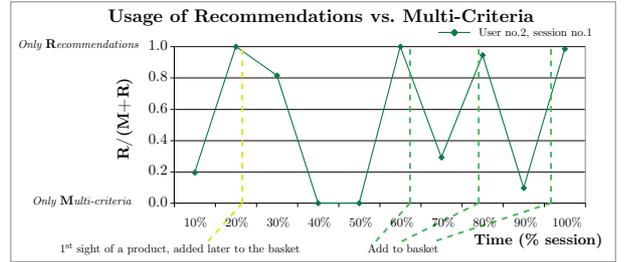**Figure 2: Evolution of MCF and CRS over time.**

Figure 3: Usage of Recommendations vs. Multi-Criteria.

normalized, as AOIs have different sizes. Then, we measured the usage of recommendations in comparison with the multi-criteria box for each user u and each session s, by computing:

$$f_{u,s}(t) = \frac{R_u(t)}{M_u(t) + R_u(t)}$$

We also analyzed the users' actions from the implicit access logs to determine the time at which users added products in the basket. Thus, we were able to cross data from the eye tracking system and the access logs. A representative example of the function f is displayed in Figure 3 (User 2, Session 1). Vertical dashed lines correspond to times at which the user added a product to the basket.

The results show that in 85.71% of the sessions, the user adds the product to the basket in the neighborhood of the intersection point with a horizontal tangent, when the use of the recommender
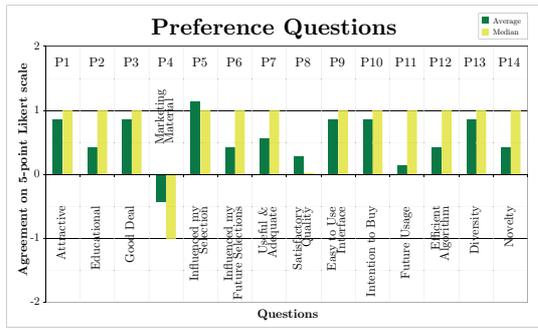
Figure 4: Answers of the Assessment Questionnaire.

is maximal. This means a decision is imminent at time t_x, when:

$$\frac{df_{u,s}}{dt}(t_x) = 0 \ and \ \frac{df_{u,s}}{dt}(t < t_x) > 0 \ and \ \frac{df_{u,s}}{dt}(t > t_x) < 0$$

After the experiment, we asked users to fill out an assessment questionnaire. The responses of this post-study survey are summarized in Figure 4.

5. DISCUSSION

The results show that for all users tested the use of the recommender system (CRS) progressively replaces the use of multi-criteria search (MCF) over time according to fixations (cf. Figure 2). We successfully reproduced hypotheses of [3] about the purchase decision making process' division into two steps. Thereby, we noticed that each participant started the experiment by identifying adequate search criteria. Thanks to users' access logs we know that in 85% of the sessions this step required several search cycles of choosing some criteria, reading examples of corresponding perfumes' detail pages, and going back to the search page to refine the search criteria. In the other 15%, users proceeded to a decision with only one attempt at defining some adequate search criteria. The results reflect the fact that users spend less time looking at choices provided in lexicographic order (list-view) by the MCF box in favor of detailed pages. Then, all interviewees used the selected search criteria in order to consider alternatives, before making a decision. Despite this notable division within the decision process, the impact of the recommender can be observed in both steps of 93% of the sessions, according to eye tracking data. Only one session of one user does not follow this scheme. We believe that this case is not meaningful, since the user did not notice the recommender system and consequently did not use it during this session.

Even when the influence of the recommender is noticeable within both decision making steps, our experiment highlights that this effect is greater at the end of the decision process (see Figure 2). In nearly all product searches we can see that the function $f_{u,s}$ is continuously increasing, from the beginning of the search to the moment he added it to the basket. Only two sessions of two different users are exceptions to this rule: in the first case, the user has not noticed the recommender. In the second case, the product search aimed at finding a perfume for a gift (Session 2). After the experiment, this user mentioned his incapability to come to a decision because he did not know the preferences of the person he had in mind for the gift. Consequently, he chose a product based on a brand that he liked. At the same time, we also noticed that the usage of recommendations in comparison with multi-criteria search, was increasing faster for the second and third perfumes added to the basket during Session 1, as expected in Section 3.

In order to ensure the veracity of decisions to add products in the basket, we asked users if they would buy the chosen perfumes given the opportunity - keeping in mind that one of the participants was going to win a 100 Swiss franc voucher to buy one of the perfumes he/she had added to the basket - or at least go in a perfume shop to smell them and learn more about them (P10 in Figure 4). Four users agreed to buy given the opportunity; two were not sure, but agreed to smell them before making a final decision. One user had a neutral opinion. This data supports the idea that participants took the decision seriously, and that the influence of the recommender constantly increases over time to reach a final decision.

6. CONCLUSIONS AND FUTURE WORK

In this paper, we examined the impact of a recommender system on customers' decision process. We followed the two-steps model of Haubl *et al.* [4] with the aim of discovering new usage patterns and decision subprocesses, thanks to an eye tracking system. In particular, we paid attention to how the influence of recommender systems integrates into the purchase decision making model. We based our experiment on our group's e-commerce simulation website selling perfume and other fashion goods. We collected more than 20,300 fixation data points. The analysis of this data, cross-checked with the users' actions and assessment questions, leads to the major conclusion that the recommender system is involved in the two steps of the purchase process: it both helps the users to refine criteria search (to sort items in a lexicographic order) and to choose alternatives to consider for purchase. We showed that the influence of the recommender increases over time. The closer the purchase decision is to being made, the more users turn towards the recommender to assist them in their decision process.

The study constitutes a first major step towards the formalization of subprocesses involving recommender systems within the purchase decision model. The perspectives will also consist of measuring the impact of recommendation categories and how they offer opportunities to discover new and interesting alternatives by providing diversity.

7. REFERENCES

[1] J. Hauser and B. Wernerfelt. An evaluation cost model of consideration sets. *Journal of Consumer Research*, 16:393–408, March 1990.

[2] S. Ho. Web personalization and its effects on users' information processing and decision making. PhD Thesis of the Hong Kong University of Science and Technology, 2004.

[3] G. Häubl and K. Murray. Preference construction and persistence in digital marketplaces: The role of electronic recommendation agents. *Journal of Consumer Psychology*, 13(1):75–91, 2003.

[4] G. Häubl and V. Trifts. Consumer decision making in online shopping environments: The effects of interactive decision aids. *Marketing Science*, 19(1):4–21, 2000.

[5] P. Maes, R. Guttman, A. Moukas, and R. Moukas. Agents that buy and sell: Transforming commerce as we know it. *Communications of the ACM*, 42:81–91, 1999.

[6] J. Payne, J. Bettman, and E. Johnson. *The Adaptive Decision Maker*. Cambridge University Press, 1993.

[7] P. Pu and L. Chen. Trust-inspiring explanation interfaces for recommender systems. *Knowledge-Based Systems Journal*, 20:542–556, 2007.