Classification of Goal-Directed Search and Exploratory Search Using Mobile Eye-Tracking

Completed Research Paper

Jella Pfeiffer
Karlsruhe Institute of Technology (KIT)
Englerstr. 14, 76133 Karlsruhe, Germany
Jella.Pfeiffer@kit.edu

Martin Meißner
University of Southern Denmark
Niels Bohrs Vej 9, 6700 Esbjerg, Denmark
(and) Monash University
Melbourne, Australia
Meissner@sam.sdu.dk

Jascha Prosiegel
Johannes Gutenberg-Universität Mainz
Jakob-Welder-Weg 9, 55128 Mainz, Germany
jprosieg@gmail.com

Thies Pfeiffer
CITEC, Bielefeld University
Inspiration 1, 33619 Bielefeld, Germany
Thies.Pfeiffer@Uni-Bielefeld.de

Abstract

In this paper, we investigate the visual attention of consumers with the help of mobile eye-tracking technology. We explore attentional differences between goal-directed search and exploratory search used when consumers are purchasing a product at the point-of-sale. The aim of this study is to classify these two search processes based solely on the consumers' eye movements. Using data from a field experiment in a supermarket, we build a model that learns about consumers' attentional processes and makes predictions about the search process they used. Our results show that we can correctly classify the search processes used with an accuracy of nearly 70% after just the first nine seconds of the search. Later on in the search process, the accuracy of the classification can reach up to 77%.

Keywords: Mobile commerce, Human Information Behavior, goal-directed search, exploratory search, Decision Support Systems (DSS), Neuro-IS

Introduction

Today, smartglasses such as Google Glass allow local information (e.g., information about a sightseeing spot or nearby shops) to be displayed in the user's field of view. Recent advances in mobile eye-tracking systems indicate that, in the near future, eye-tracking technology will probably become pervasive in the field of mobile consumer information technology, such as smartglasses. Identifying the objects that users are looking at, by integrating mobile eye-tracking equipment into the smartglasses, would provide unmatched fine-grained data on human attention processes, which in turn would enable the design of innovative real-world applications: For example, a mobile recommendation agent for purchase decisions.
in stores could provide recommendations based on the classifications of the wearers’ ongoing search processes and inferences about their preferences and intentions.

A widespread availability of mobile eye-tracking data in all kinds of everyday situations would allow research on human information behavior to flourish. We have therefore developed and used mobile eye-tracking technology to study, inter alia, the specific patterns of human information behavior, as recently suggested by Hemmer and Heinzl (2011). In doing so, we are following a research trajectory that tries to explain these patterns and use them to make general predictions about users’ current information requirements, grounded in extensive data on covert visual attention.

In this study, we focused on two different search situations: Goal-directed search and exploratory search (Janiszewski et al. 1998). These two search situations are of particular interest for both marketing researchers and practitioners. Research by Moe (2003), for instance, suggests that goal-directed search and exploratory search situations are linked to very different consumer motivations and different consumer responses to various marketing messages. In goal-directed situations, consumers need specific information, such as product alternatives, in order to find the alternative that best matches their individual preferences. For example, a consumer might search for muesli that includes almonds and chocolate but is low in calories. In an exploratory situation, however, a consumer will be primarily scanning and gathering general information about the available products.

The aim of this paper is to build a classifier that determines whether consumers are in goal-directed or exploratory search situations. One major challenge when classifying different search situations is that very little is known about the consumer in the early stages of the search processes. This paper addresses this problem by taking advantage of the large amount of fine-grained data that eye tracking provides already in the first seconds of the attentional search process. In particular, we investigated the following two research questions:

Are there measurable differences in eye movements that help predict whether a person is engaging in a goal-directed information search or an exploratory information search?

Which attentional variables are best suited for differentiating between goal-directed searches and exploratory searches relatively early in the search process?

According to the taxonomy developed by Gregor (2006), our research approach can be classified as being type 3 IS theory (Prediction). We followed an exploratory research approach because we found no theories or empirical studies on goal-directed searches and exploratory searches in real-world in-store settings. We therefore did not explicitly formulate research hypotheses, since the contexts in which these search processes were previously studied are very different from the in-store setting context we investigated in our paper. Instead, based on a review of the literature on the two search processes in other decision contexts, we formulated questions about the differences between both search situations, more specifically in terms of the consumers’ attentional processes. We considered a couple of potential measures and explored which attentional measures were the most suited to being used as predictor variables in a logit regression model. By taking mobile eye-tracking data from 20 participants put into either one of the search situations in a real-world setting, we trained our classification model and tested its predictive quality.

The results showed that consumers consider a smaller number of products in goal-oriented situations than in exploratory situations. Furthermore, they spend more time acquiring detailed product information and consider products that are located closely to previously-examined products. Our model can classify the purchase situations with an accuracy of nearly 70% after just nine seconds of observation, using only the number of products considered by the user. In the later stages of the attentional process, the distance of consecutive fixations on different products helps predict the situation and our model achieves an accuracy of up to 77%. In sum, we are able to satisfactorily detect the purchase situation early on in the search process, using only one predictor variable: the number of products that the user has considered so far. This is a promising result since the number of different products considered can be determined quite easily in real-time.

In the next section, we will elaborate on studies that investigated attentional processes in lab-like experiments, in online stores as well as in brick and mortar stores such as supermarkets. We will then describe our field experiment and how the data was used to calibrate a model that predicts whether the
consumer searches in a rather goal-directed manner or an exploratory manner. We end the paper with a discussion of our results and an outline of the relevance of our findings to the development of different applications. We also discuss avenues for future research.

**Literature Review on Information Needs and Attentional Processes in Purchase Situations**

**Investigating the Information Needs of Consumers**

Stores often have up to hundreds of products in the same product category, and these products are most likely different with respect to their characteristics, such as price, brand, materials, or ingredients in the case of food. As a consequence, the product-specific information available in a purchase situation goes well beyond a consumer’s bounded cognitive capacity. Therefore, marketers have been especially interested in assessing what captures consumers’ in-store attention (Clement et al. 2013), with the aim of winning the competitive race of brand salience on the product shelf (Van der Lans et al. 2008). Since consumers look only at a small fraction of the product alternatives in a store, retailers and marketers direct an increasing amount of their marketing budgets towards point-of-sale marketing. This strategy likely works because many consumers enter a store without yet knowing the exact product they are going to buy. In a situation where only a fraction of the products is looked at, in-store displays can influence consumers to buy certain products. Chandon et al. (2009) use the term “visual equity” to describe the impact of attention on sales. They argue that visual equity has to be taken into account in addition to the well-known “memory-based equity” that captures learned product and brand preferences — i.e., preferences that have been formed through the influence of advertisement.

In our study, we are especially interested in investigating search processes that have diverging information needs. Goal-directed and exploratory search processes fulfill this criterion and are of specific relevance to real-world purchase environments (Janiszewski et al. 1998). If a consumer’s behavior is directed and guided by some specific purchase in mind, i.e., a specific strategy or plan, we call this goal-directed search. In contrast, exploratory search describes a more stimulus-driven undirected search process during which respondents simply monitor the decision environment. In a similar vein, Novak et al. (2003) distinguished between goal-directed behavior that involves a search for specific product information such as price, and experiential behavior, which involves a less structured search. The authors describe these less structured searches as non-directed and affective; they also stress the hedonic component involved in browsing. Furthermore, exploratory search has been characterized as the default behavior of consumers. Indeed, whenever we are not actively involved in looking up pieces of information, our visual system continues to screen the environment (Janiszewski 1998).

In decision environments such as supermarkets, we expect consumers’ visual attention to be fragmented. Physical and design features of the product packages such as the shape, colors and contrast are supposed to dominate the initial phase, and should be most influential when the user screens the products for information (Clement et al. 2013). Therefore, in the next section, we will review previous empirical studies that have investigated attentional processes in lab-like experimental settings.

**Change of Information Acquisition during the Decision Process**

Several studies have investigated how information acquisition changes during the decision process. Russo and Leclerc (1994) were the first to examine this question empirically by analyzing eye movements. The authors suggested that the decision process consists of three stages: orientation, evaluation, and verification. In the orientation stage, consumers screen the products for decision-relevant information and try to reduce the complexity of the decision situation by eliminating alternatives based on important product characteristics. In the evaluation stage, consumers compare the remaining alternatives by looking at the product characteristics in greater detail. In the final verification stage, consumers check the characteristics of their favorite alternative before choosing it. While this study initiated a lot of follow-up research on decision processes, the validity of the findings can be called into question: the supermarket shelves were placed in a laboratory against a one-way mirror and the researchers themselves tracked the eye movements of participants by observation from behind the mirror. Thus, no eye-tracking equipment was used. The rules used to define the decision stages have been criticized since then, and it is unclear
whether these rules can be applied to mobile eye-tracking data gathered in real-world environments (Gidlöf et al. 2013).

Researchers have been especially interested in finding out what happens in the final few seconds before a decision is made. The so-called gaze-cascade effect (Shimojo et al. 2003) describes the finding that attention to the chosen alternative strongly increases when a consumer is close to making a decision. Though the explanation for this effect is controversial and has generated much discussion in the literature (see Bird et al. 2012), the effect was replicated in a several different settings. Shi et al. (2012) stressed that the gaze-cascade effect is important because information concerning a user’s visual attention can be used to predict the brand that a person will choose and is an indicator of how close a respondent is to making a purchase. To the best of our knowledge, the gaze-cascade effect has not yet been replicated in the context of a supermarket.

Meißner et al. (2014) investigated information acquisition processes in multi-alternative conjoint-analytic choice tasks using eye-tracking. The authors found that attention gradually shifted to high utility alternatives and important attributes. These findings are in line with previous research on decision stages: Indeed, the results suggest that early on in the decision process, decision-makers start by exploring the decision environment, but later focus on the most important information in order to make their decision. Based on the same dataset as Meißner et al. (2014), Pfeiffer et al. (2014) defined decision stages based on the number of fixations, and thus showed that respondents process information more alternative-wise in later stages of the decision process. This result is also in line with the idea of a verification stage. In order to verify the finally chosen alternative, respondents have to look at the different features of one alternative, which results in a high number of alternative-wise transitions.

To the best of our knowledge, the study by Gidlöf et al. (2013) was the first to investigate attentional processes of consumers in a real in-store context. These authors were particularly interested in verifying the different stages proposed by Russo and Leclerc (1994) because they thought that the structure of the task environment would strongly affect eye movements. In their empirical study, the authors compared decision-making tasks with search tasks. In the search task, respondents had to find a specific pasta on the supermarket shelf and bring it to the research assistant. In contrast, in the decision-making task, participants were asked to buy a pasta product of their choice. Gidlöf et al. (2013) found that the two tasks were substantially different with respect to how information was processed. In particular, the second stage of the decision-making task contained a significantly higher number of re-fixations than the second stage of a comparable search task. While the decision-making task is similar to our understanding of an exploratory search task, the search task described by the authors is very different from a goal-directed task. The search task as defined by Gidlöf et al. (2013) describes the search for one particular product and thus resembles an image detection task where sizes, colors and shapes are compared. A goal-directed search task, however, describes the search for a product that fulfills certain search criteria. Thus, in a goal-directed search, the participant looks for particular product information such as prices, brands and ingredients. Moreover, Gidlöf et al. did not separate or classify the attentional processes being used during their tasks.

Our review of the literature on attentional processes indicates that the manner in which information is processed in different decision situations leads to significant differences in eye movements. Indeed, previous findings suggest that different eye tracking measures could be relevant at different stages of the search process. In the next section, we will discuss which attentional measures can be expected to differ between the two search processes under investigation.

Information Acquisition in Goal-Directed and Exploratory Search

In a previous study on goal-directed and exploratory search, Moe (2003) investigated the two search processes (goal-directed search and exploratory search) by using the clickstream data of consumers in an online shop — not in a real-world in-store setting. A major aim of this study was to find indicators that would help distinguish between both search processes. The author argued that exploratory search most likely occurs when consumers have little previously stored knowledge, i.e., information that could help direct the search process. In particular, Moe (2003) suggested that four different shopping strategies could be distinguished from one another:
First, “directed buying” describes a situation in which the consumer has already gathered enough decision-relevant information to be able to make a decision. The search process in this case is rather short and focuses on single pieces of information that are essential to finishing the search process. Second, “search and deliberation” is a search strategy that is used to “build the consideration set and evaluate the items in the set”, in preparation of a decision. Third, “hedonic browsing” is a search process motivated by the pleasure that results from the in-store experience. Fourth, in a “knowledge-building” search process, it is the shopper’s objective to increase the product and/or marketplace expertise. While the former two search processes are goal-directed, the latter two processes are exploratory. From a marketing perspective, “search and deliberation” and “knowledge building” seem to be the most interesting search processes. “Directed buying” is less interesting because the consumer has already gathered relevant information and is close to making a decision, i.e., the consumer is less likely to be influenced by marketers’ recommendations and is less likely to be in need of help. “Hedonic browsing” is also less interesting to study because it is unclear when this search behavior occurs, or why it might need to be supported. Helping implement this kind of behavior might boost the enjoyment and fun of consumers, but would likely not be directly useful to consumers.

Building on the empirical results of the study by Moe (2003) and comparing “search and deliberation” (a goal-directed search) with “knowledge building” (an exploratory search), we expect to see a number of differences with respect to how information is processed in the two search processes. Indeed, Moe (2003) looked at the number of different product categories and products that were investigated during the web search. Respondents in the “search and deliberation” condition looked at a significantly higher number of different products than in the “knowledge building” condition. In line with this finding, we expect respondents in the goal-oriented search condition in a brick and mortar store to also look at a higher number of different products. Moreover, eye tracking will enable us to understand which product characteristics are being processed at a given time. For each product, we distinguished between attention to brand, price and product details. The study by Moe (2003) suggested that the sum of webpages viewed with detailed product specific information is higher in the “search and deliberation” situation than in the “knowledge building” situation. In line with this result, we expect consumers to more intensely consider detailed product information and spend more time doing this in goal-directed tasks than they would in exploratory tasks.

Finally, eye tracking allows us to gain information on the distance between two consecutive fixations on different products. Contrary to product presentations on the web, which are restricted in terms of screen space, and where the complete product set is usually not displayed on a single screen, there are no such restrictions when consumers search on supermarket shelves. To the best of our knowledge, empirical studies have not yet investigated or tested fixation distances (saccadic length). Thus, the question of whether the distance of two consecutive fixations is greater in goal-oriented search processes or in exploratory search processes remains an open research question.

**Research Method**

**Experimental Design**

We conducted an experiment in a real world setting, i.e., a medium-sized supermarket, using mobile eye-tracking equipment. Twenty shoppers were recruited directly after entering the store and they received 10€ as incentive for participation. We used a within-subject design where each participant had to complete the goal-oriented task (GT) and the exploratory task (ET) twice in alternate order and for four different product categories: muesli, cereals, marmalade and tea. The sequence of product categories was fixed and participants were randomly assigned to one of two experimental groups (see Table 1). Twelve out of the 80 observations (4 tasks times 20 participants) are missing because of technical problems with the USB-port during recording of the eye movements; an additional observation is missing because the person in question did not fixate on any product on the shelf in the first 30 seconds of the task. Thus, we have 16 observations for muesli and cereals, 17 for marmalade and 18 for tea. The number of products available for each category was 116, 76, 202, and 190, respectively.
At the beginning of the experiment, participants were given the task description and the experimenter ensured that participants had understood the task. In the GT and the muesli condition, participants were told to select muesli for a friend who would come over for a visit. In this scenario, the friend liked to have muesli which (1) contained chocolate, (2) contained almonds, (3) and as low in calories as possible. In the GT for cereals, the required criteria were (1) contains cinnamon, (2) package smaller than 400 grams, (3) low in calories. In the marmalade GT, participants were told to look for a marmalade that (1) contained oranges, (2) contained at least one additional other flavor, and (3) was low in price/kg. In the tea group, the requirements were: (1) is organic, (2) contains peppermint, (3) has the lowest price per 100 grams. In each category, the number of products fulfilling the first two criteria was eight, two, six, and two respectively. There was exactly one optimal product per category that fulfilled all three criteria. In the ET, participants were asked to gain a fairly good overview of the muesli (cereal, marmalade or tea) assortment and to silently determine criteria that were important to them when buying the product. Afterwards, they had to choose one product they would potentially buy themselves.

During the tasks, participants wore the SMI Eye Tracking Glasses (first model, 30Hz gaze data with an accuracy of 0.5 degrees, scene camera gaze overlay with 24Hz at 1280x960 and a field of view of 60°x46°), to record their overt visual attention. 80% of the participants regarded the equipment as non-distracting, while 20% regarded it as partially distracting; and none viewed it as distracting.

### Coding of Eye-Tracking Data

The problem with mobile eye-tracking data in comparison to eye-tracking data in front of stationary interfaces (for example using remote eye-tracking systems) is the time needed for annotating the data. Whereas on a stationary screen, software can automatically assign gazes to areas of interest, no satisfying solution yet exists for mobile eye-tracking data. This is because in mobile situations, users go through a number of head and body movements that result in constantly changing fields of vision. Thus in real-world scenarios, the position of areas of interest is constantly changing in the recorded video. Other studies have addressed this problem by building classifiers based on computer vision techniques that can detect the products fixated by the user in real-time (Harmening and Pfeiffer 2013). This was achieved by matching an area around the fixated point with product images from different angles that are stored in a database. However, this method is not yet accurate enough, so when we analyzed our data, we relied solely on the manual annotations of two coders.

For each tenth of a second, we determined which product was fixated by the participant. A fixation is a gaze that lasts at least 100 milliseconds, while the average fixation duration during a visual scene perception lasts about 330 milliseconds (Rayner 1998). For each product, we furthermore annotated whether the price tag, detailed product information or the packaging (including brand names and pictures) were looked at. In addition to this, we computed the distance in millimeters between two fixations on different products. Table 2 provides an overview of the three measures that we determined based on the annotations. A value for each of these measures was determined/updated every second of the decision process.

The coding was done after the experimental task was over. After a definition of the coding scheme, the annotators both coded four example videos. Following this, the annotations of the two different annotators were compared, divergences were discussed and the coding manual was updated. This was repeated until a common understanding of coding could be ensured. The final coding manual was consequently evaluated with four annotators on a two-minute sample with 200 fixations. For the
Identification of the fixated products, the four annotators achieved a Krippendorff alpha of 0.854 (87% pairwise agreement) and for the distinction in functional areas (price, packaging, product detail information) a Krippendorff alpha of 0.815 (91% pairwise agreement). These ratings are above the common threshold of alpha > 0.800 for drawing conclusions suggested by Krippendorff (2004, p. 241). Differences were primarily found for very short fixations during head movements, which were not relevant for the current analysis. Finally, two annotators coded the full set of videos. The coding of the aspects relevant for the analysis took about 80 hours for the 2.24 hours of videos.

<table>
<thead>
<tr>
<th>Table 2. Measures.</th>
</tr>
</thead>
<tbody>
<tr>
<td>NUMBERDIFFPROD</td>
</tr>
<tr>
<td>AVGLENGTHDETAIL</td>
</tr>
<tr>
<td>AVGDISTPERCENT</td>
</tr>
</tbody>
</table>

**Results**

*Differences in the Acquisition Behavior*

The mean age in the sample was 31.3 (standard deviation (std.)=13.27, maximum 53 years) and 70% were female. Participants took less time in the exploratory task (mean=76.06 seconds, std.=49.2756, min=10.65 seconds, max=269.46 seconds) than in the goal-oriented task (m=162.52 seconds, std.=113.09, min=24.69 seconds, max=458.10 seconds). Consequently, when comparing the means of the three measures NUMBERDIFFPROD, AVGLENGTHDETAIL and AVGDISTPERCENT at different points in time, we had to keep in mind that some participants had already possibly stopped their search for information. Figure 1 shows the number of participants still executing the task after x seconds for the four different product categories. From the 67 observations with complete datasets, only 56 remained after 30 seconds, and only 35 remained after 100 seconds. Because roughly only half of the participants were left after 100 seconds and thus there were only eight participants left in the exploratory search task, we decided to focus our analysis on the search behavior within the first 100 seconds.

We were interested in determining the extent to which the three measures differed over time in the two different search situations. Indeed, this would indicate the potential of these measures to differentiate the two situations. The results for the first 100 seconds are shown in Figures 2, 3 and 4. Contrary to the findings by Moe (2003) in the online shopping context, we found that the total number of different products fixated (NUMBERDIFFPROD) is larger in exploratory tasks than in goal-oriented tasks (t-tests indicate significant differences from second 4 on). This difference increased during the search process and decreased later on. The decreasing distance between the two curves partly resulted from the higher dropout rate of participants in the exploratory task. Indeed, the means in Figures 2-4 were calculated across all participants, even if some of the participants had dropped out. Hence, when a participant completed the task, the value of the last second of her task was extrapolated to future seconds. For example, assume a participant had fixated on average 40 different products in second 50 for the two exploratory tasks and finished the task in second 51. NUMBERDIFFPROD would then be set to 40 for this participant for all seconds >=50. Keeping such information in the data set even if the information came from participants who had already dropped out was necessary in order to smooth the function; otherwise, we would have probably seen large jumps in the measures.
Figure 3 displays the average time that participants processed detailed product information (AVGLENGTHDETAIL). Overall, the difference in AVGLENGTHDETAIL is as expected, with larger values appearing in the goal-oriented task (t-tests indicate significant differences from second 12 onwards). We furthermore observed that participants spent more time on average looking at the detailed information on products in goal-oriented tasks than in exploratory tasks. Furthermore, we learned that there is a steady but slow increase of AVGLENGTHDETAIL in exploratory situations. Apparently, participants analyzed products in greater detail the later they were in their decision-making process. In goal-oriented tasks, the increase stopped at around 40 seconds; this increase thus stopped earlier than in exploratory tasks. It seems that in goal-oriented tasks, participants analyzed products in detail in earlier stages of the process.

The AVGDISTPERCENT, which is plotted in Figure 4, also differs in the two situations, with larger values appearing in the exploratory tasks (t-tests indicate significant differences from second 39 onwards). Interestingly, in the first seconds, AVGDISTPERCENT increased for exploratory tasks and decreased for goal-oriented tasks. This observation could be explained as follows: In exploratory tasks, users first tried to get an overview of the products, whereas users in the goal-oriented task quickly started comparing products that were close to each other on the shelf. Overall, with 5%-7% AVGDISTPERCENT for both situations, we see that distances between consecutive products fixated are rather small. Thus, it is more likely that consumers consider products in close vicinity of each other. This might also be caused by strong preferences for certain brands or price ranges, since consumers know that in a typical supermarket, products are oftentimes sorted by brands horizontally and by price vertically on the product shelf.
Figure 2. Mean Values of NUMBERDIFFPROD for the different decision situations.

Figure 3. Mean Values of AVGLENGTHDETAIL for the different decision situations.

Figure 4. Mean Values of AVGDISTPERCENT for the different decision situations.
**Classification Model**

Using NUMBERDIFFPROD, AVGLENGTHDETAIL and AVGDISTPERCENT, we built a classification model that identifies whether a person is in a goal-oriented or exploratory search situation. We applied a binary logistic regression with the type of purchase situation being a dependent variable predicted by NUMBERDIFFPROD, AVGLENGTHDETAIL and AVGDISTPERCENT.\(^1\) One regression was computed for every second, because the predictive validity of the three variables could differ at different points in time.

Figure 5 plots the standardized beta-coefficients for the three explanatory variables for each of the 99 regression models (one for each second, starting from second two) and Figure 6 shows when each of the beta-coefficients is significant. Wald tests indicated that the three coefficients differed significantly from one another from second 9 onwards. NUMBERDIFFPROD appeared to be the best variable for predicting the search situation. From second 3 onwards, the coefficient was significant and positive, which means that the higher the NUMBERDIFFPROD, the higher the probability that the user is in an exploratory situation. Figure 5 shows that the beta-coefficient is much larger than for the other two explanatory variables, especially in the beginning of the search process. However, the predictive validity of NUMBERDIFFPROD decreased over time compared to AVGLENGTHDETAIL and AVGDISTPERCENT. Indeed, from second 64 onwards, the coefficient for AVGDISTPERCENT also became significant and, from second 39 onwards, a Wald test showed that there is no longer a significant difference between AVGDISTPERCENT and NUMBERDIFFPROD. Overall, we can conclude that the higher AVGDISTPERCENT, the higher the probability that the user is in an exploratory situation. The coefficient for AVGLENGTHDETAIL was negative most of the time, meaning that the higher AVGLENGTHDETAIL, the higher the probability that the user is in a goal-oriented purchase decision. However, the coefficient was not significant for most of the regression models.

In sum: For roughly the first sixty seconds, the number of different products that the consumer looks at (NUMBERDIFFPROD) is sufficient to predict the purchase situation. In contrast, later in the search process, measures that capture how long detailed product information is processed (AVGLENGTHDETAIL) and how close the products of interest are on the shelf (AVGDISTPERCENT) should also be used as predictors.

\(^1\) We also determined the relevance of product categories by computing a second version of the models with product categories as controls. Because the product categories were not significant, we refrained from using this more complex model and did not include this variable.
To evaluate our classification model, we used a 10-fold cross-validation (Kohavi 1995) to evaluate the accuracy of the prediction models. Cross-validation randomly divides the data set into equal-sized subsets called folds and repeatedly uses them for performance testing. Thus in our case, we randomly divided the participants into ten pairs and predicted the search process for each pair with an estimation based on a model that was calibrated with the remaining eighteen participants, respectively. This procedure was repeated five times in order to determine the mean prediction accuracy of our model. Figure 7 plots the prediction accuracy when using the logistic regression with all three explanatory variables (FULL-model). It also plots the prediction accuracy when using only NUMBERDIFFPROD as predictor. We compared this result with the baseline of randomly guessing one of the two situations. The FULL-model achieved a prediction accuracy of about 60% from second six onwards, increasing to values of 66-67% from second nine onwards and rising to about 74.67% percent at second 57, which is roughly 1.5 times as good as random guessing. In sum, after just eight seconds, our approach can correctly classify at least 2/3 of respondents. A model that takes only NUMBERDIFFPROD as a predictor compares better than our FULL-model at the beginning of the process. However, from second 50 onwards, its performance decreases tremendously. Thus, when researchers are interested in determining the type of search process of a new observation at the beginning of the search process, our simple model can be recommended since it has an accuracy of 77% after 45 seconds. However, we cannot recommend this model in later phases of the process, i.e., after second 45, as its accuracy drops significantly.

Discussion

The differences in information acquisition behavior at the point-of-sale that we found in our study partly contradict previous empirical results in other decision environments (see Moe (2003)). In contrast to the behavior of online-shoppers, consumers in a brick and mortar store look at a higher number of different products (NUMBERDIFFPROD) in an exploratory search task. In line with the result by Moe (2003), they spend less time acquiring information about the different products than when they are in a goal-oriented task. In particular, they look less intensely at detailed information (AVGLENGTHDETAIL). As a new contribution to literature, we also find that it is slightly likely that consumers consider products in close vicinity of each other since they jump back and forth between distant products on the shelf slightly more often than in goal-oriented search tasks (AVGDISTPERCENT).
We find that NUMBERDIFFPROD is a strong predictor of the search process, particularly in the first seconds of the search. However, NUMBERDIFFPROD loses predictive power later on in the process. This could be caused by the unequal dropout rate in the two situations: Participants in the exploratory search task finished the process earlier than in the goal-directed search task. Thus, the value for NUMBERDIFFPROD increases for the goal-directed group, whereas it stays more or less constant for the exploratory group as soon as most of the participants in the exploratory group have dropped out. This explains why both curves seem to converge. One way of addressing this problem would be to determine the number of different products fixated for a sliding window of, for example, only the previous 10 seconds. However, since our main goal was to classify tasks early on in the search process, this issue could be neglected because it is only important in the later stages of the process.

It should be noted that these first results are based on a rather small sample. From a statistical perspective, larger samples sizes would be useful to improve the robustness of our prediction model because the relatively small number of observations in our case might result in an overfitting of the model parameters to the observed sample.

Our model is able to classify the search situation with an accuracy between 70% and 80% already early on in the search process. This result is very promising, because (1) the logistic regression is a rather simple model; (2) the model is based on data from a real-world environment where participants are disturbed by other objects in the environment; (3) it relies only on eye-tracking data; (4) it does not rely on any semantic information, for example, information about product features. By using only attentional processes as predictors, it is unlikely to achieve an accuracy of 100%.

Furthermore, we have seen that the information acquisition behavior changes during the task, which supports the theory that different decision stages exist (proposed by Russo and Leclerc 1994). Consumers consider more detailed information later on in the process, which is in line with the idea of a switch from an orientation stage to an evaluation stage. The average distance between the products considered is higher at the beginning than at the end in both search tasks. In our particular study, the focus on products...
located nearby on the shelf occurred quite early on in the process and stabilized after just 30 seconds. It is possible that acquiring information of proximate products helps in the acquisition and memorization of information. Indeed, it possibly requires less effort to compare products that are located nearby because head movements and body movements can be avoided. Overall we have learned that in goal-oriented search situations, consumers need more time to study the detailed product information. On average, consumers look at the detailed information of a product for about 4 seconds. Moreover, our results suggest that there is vast potential for supporting consumers at the point-of-sale. In the goal-oriented condition, for example, only 42.42% of respondents found the product that fulfilled all search criteria.

Conclusions

The key contribution of this paper is that it demonstrates how eye-tracking information can be used for the classification of two search processes: goal-oriented search and exploratory search. We think that our findings are new and essential to the information systems field: First, our paper is the first paper that investigates search processes in a real-world in-store setting. Second, our paper is the first that distinguishes goal-oriented and exploratory search based only on eye-movement data (instead of using the clickstream data of consumers like in the paper by Moe (2003)). Third, our paper demonstrates how the goal of classification of information acquisition behavior can be achieved.

We think that our findings are especially relevant for building recommendation agents to be used at the point-of-sale. These recommendation agents could provide the consumer with relevant product information by identifying the consumer’s information acquisition processes recorded during the purchase decision and by giving product recommendations automatically when they are needed. Several researchers have emphasized the importance of developing such recommendation agents for in-store use (Lee and Benbasat (2010), van der Heijden (2006)). The great advantage of developing such systems using mobile eye-tracking data, instead of relying uniquely on the sensors currently built into mobile devices like smartphones, is the reduced need for explicit user input. Reduced user input results not only from the automatic detection of human information behavior but also from a more fine-grained localization, not only of the user, but also of the user’s target of visual attention.

In the field of marketing research, the investigation of attention at the point-of-purchase is a relatively new endeavor. In most empirical studies, researchers have conducted eye-tracking experiments in lab-like situations. However, mobile eye tracking equipment could be used to more rigorously test the influence of marketing variables such as price and brand on attention and in-store purchase probabilities. For example, researchers could investigate when and why consumers pay attention to prices in purchase situations (Dickson and Sawyer 1990), and which in-store factors trigger attention to prices.

At the core of this research paper is the idea of combining information about eye movements, such as the number of fixations or timing, and information about the targeted objects in order to identify user behavior in real-time. This information provides grounds for attentive user interfaces that can adapt on-the-fly to current user behavior without the need for explicit communication between the user and the system. While our own research targets complex mobile scenarios, the method can be applied more easily to common desktop environments with a static remote eye-tracking system. Potential areas of application include support in logistics and fulfillment, training (e.g., for medical doctor trainees to learn how to assess medical images based on patterns learned from experts), the adaptation of search engine results for web applications, or monitoring supervisory personnel (flight controllers, tunnel surveillance, production monitoring) to provide feedback when their alertness decreases.

A major challenge for the development of future applications is the real-time analysis of eye-movement data. We were able to record the eye-movements for 80 search situations (20 respondents times 4 decisions situations). The annotation of these videos required about 80 hours of manual coding. These numbers illustrate that as long as manual coding is necessary, the determination of the information acquisition measures cannot be done and will constitute a great amount of work. The development of software that is able to automatically annotate fixations in real-time is therefore a prerequisite for building recommendation agents. For this reason, we are currently developing methods that can be used for real-time annotation of fixations to products (Harmening and Pfeiffer 2013). We think that much more research needs to be done to reach the goal of building a mobile recommendation agent based on eye-tracking data.
References


