On Cognition, Emotion, and Interaction Aspects of Search Tasks with Different Search Intentions

Yashar Moshfeghi
School of Computing Science
University of Glasgow
Glasgow, UK
Yashar.Moshfeghi@glasgow.ac.uk

Joemon M. Jose
School of Computing Science
University of Glasgow
Glasgow, UK
Joemon.Jose@glasgow.ac.uk

ABSTRACT

The complex and dynamic nature of search processes surrounding information seeking have been exhaustively studied. Recent studies have highlighted search processes with different intentions, such as those for entertainment purposes or re-finding a visited information object, are fundamentally different in nature to typical information seeking intentions. Despite the popularity of such search processes on the Web, they have not yet been thoroughly explored. Using a video retrieval system as a use case, we study the characteristics of four different search task types: seeking information, re-finding a particular information object, and two different entertainment intentions (i.e. entertainment by adjusting arousal level, and entertainment by adjusting mood). In particular, we looked at the cognition, emotion and action aspects of these search tasks at different phases of a search process. This follows the common assumption in the information seeking and retrieval community that a complex search process can be broken down into a relatively small number of activity phases. Our experimental results show significant differences in the characteristics of studied search tasks. Furthermore, we investigate whether we can predict these search tasks given user's interaction with the system. Results show that we can learn a model that predicts the search task types with reasonable accuracy. Overall, these findings may help to steer search engines to better satisfy searchers' needs beyond typically assumed information seeking processes.

Categories and Subject Descriptors: H.3.3 Information Storage and Retrieval - *Information Search and Retrieval - Search Process*

General Terms: Experimentation, Human Factors

Keywords: Search Intents, Emotion, Cognition, Interaction, Prediction, Re-finding, Entertainment

1. INTRODUCTION

Recent advances in query log analysis for search engines such as Google, Yahoo! and Bing has shown that Web searchers' intentions[1] do not always fit into the typical taxonomy of informational, navigational or transactional intentions [37]. Although informational (i.e. to acquire information present on one or more sites), navigational (i.e. to reach a particular site) and transactional (i.e. to perform some Web-mediated activity) intentions are all commonly found in query logs [9], there is an increasing body of evidence showing that re-finding and re-retrieving visited information is a regular activity for searchers [37, 12]. In addition, there is also recent interest in the notion of entertainment, where the searchers do not have any particular information need [13] and the prime motivation behind their search is to satisfy an emotion need [28]. The aim of this paper is to study the characteristics of search processes motivated by different intentions, i.e. information seeking, re-finding information and/or entertainment.

Knowledge of the underlying intentions of searchers engaging in a search process on the Web can be very beneficial to better satisfy their needs. Web search engines constantly improve their retrieval effectiveness by mining information from query logs with regards to different aspects of the search processes, including prediction of the search (query) intents. The aim of this paper is to study and compare the seeking activities occurring during search processes with different intents. Through this, we can better quantify searchers' needs, and in turn improve the effectiveness of the recommendation, query suggestion, retrieval, presentation, etc.

As has traditionally been explored, the information seeking process is complex and dynamic [19, 6]. It is complex because it involves a searcher with an information need (IN) initiated by an anomaly in his or her current state of knowledge [7], who is seeking information to resolve the problem, and in turn satisfy the need. This need is usually transformed into a query and submitted to a retrieval system where a set of potentially relevant documents are retrieved and presented. However, the formulated query does not always provide an adequate description necessary to retrieve relevant documents [40] since it is only an approximation of the actual information need [36]. Likewise, it may not adequately define the characteristics of relevant documents, or indeed any relevant information, because of an *ill-defined* information need situation. Therefore, in the majority of cases, searchers are unsatisfied with the results obtained in response to their initial retrieval formulation [38], and must engage in further interaction with the system to resolve their

[1]When we say search intention we mean the underlying motivation of the search task. Intent and intention are used interchangeably.

WWW 2013, May 13–17, 2013, Rio de Janeiro, Brazil.
ACM 978-1-4503-2035-1/13/05.

needs. This introduces the dynamic aspect of the information seeking process which we will further examine.

The dynamic aspect of the information seeking process refers to the development and evolution of IN during a search session [35], from an initially vague state to a clear and well-defined one [20]. This change in IN can happen dramatically or gradually [30] as a result of new information a searcher is exposed to, i.e. the perusal of relevant, and even irrelevant documents [40]. This evolution of IN improves the searcher's query statement formulation[4], and changes what they consider relevant at both the early and late stages of the search [31]. However, two important questions emerge. First, whether such characteristics are generalizable to other search scenarios where the motivation is other than seeking information, e.g. re-finding or entertainment. Second, whether these search scenarios can be modelled via such characteristics.

In order to investigate our research questions, we follow a common approach for modelling and explaining information seeking behaviour. In general, it assumes that complex information seeking and retrieval process can be broken down into a relatively small number of activity stages (which we refer to as *phases*) [16] (p. 138) within which user behaviour is investigated. Kuhlthau's *information seeking process* (ISP) model investigated the affect, cognition and action of participants in different phases of an information seeking process [22]. Although Kuhlthau's ISP model was originally proposed for long term information seeking processes, she explains that the concepts of process, uncertainty and complexity emerging from her ISP model can be useful for designing user-centred IR systems [23]. Indeed many researchers follow this approach in information retrieval scenarios, e.g. White et al. [40]. Using an interactive IR framework, we investigate the characteristics of four search processes i.e. information seeking, re-finding, and two different entertainment intentions. In particular, we divide each search process into a small number of phases where the searchers' interactions within each phase is studied. Finally, we investigate the predictability of the search motivation given the interaction history at different phases of the search process.

This paper has three novel contributions. First, based on sociology literature, two simulated search tasks are designed for entertainment-based search processes. Second, the characteristics of search tasks with different intentions are studied using data gathered via questionnaires and interaction history. Finally, the predictability of the search intents is investigated given the interaction history at different points of a search process. The remainder of the paper is organised as follows: related works are presented in Section 2, the experiment methodology is described in Section 3. Results and discussion are presented in Section 4, and finally the paper is concluded in Section 5.

2. RELATED WORK

Seeking Information on the Web: Web search systems such as Google, Bing and Yahoo! are important tools for seeking and accessing information on the World Wide Web (the Web). As we discussed earlier (see Section 1), the information seeking process is complex and dynamic in nature. Many past theories and models have attempted to explain the complexity and dynamic nature of information seeking processes. In general they assume that complex information seeking and retrieval behaviour can be broken down into a relatively small number of activity stages [16] (p. 138). For example, Kuhlthau's *information seeking process* (ISP) model illustrates information seeking activities in six stages, each of which differentiated and determined the cognitive, affective and physical aspects of a searcher. Kuhlthau's model is one of the most popular models to investigate the cognitive, affective and physical aspects of a searcher in an information seeking process. She demonstrated that searchers adopt different information seeking strategies as they move from one stage of the ISP to another. Furthermore, she demonstrated that people's feelings, thoughts and actions interact within the ISP. The fundamental principle behind Kuhlthau's model is the *uncertainty principle* [22]. This refers to the existence of a cognitive state which causes feelings of anxiety and lack of confidence. Feelings of doubt, anxiety and frustration are associated with vague and unclear thoughts. The model shows that during a typical information seeking process, the thoughts of a searcher become clear; consequently, their confidence increases and their feeling of doubt, anxiety and frustration decrease. Following this approach, in this paper we investigate cognition, emotion and interaction aspects of search tasks with different intentions.

The cognitive aspect of the ISP has been an important theme of IR research since the earliest work in the field. Belkin et al. [7] explain that the information need motivating an ISP is derived from searchers' *anomalous state of knowledge* (ASK). The ASK refers to the gap between what searchers know and what they want to know. Taylor [36] explained that query formulation can be a cognitively demanding process, and the searchers' queries are approximate representations of the actual information needs. This query formulation problem can be magnified if searchers are facing a vague or ill-defined information need [35]. On the other hand, Saracevic [31] discussed how the concept of relevance can change for searchers during an ISP as a result of a change or development in their information needs throughout the search. These findings motivated a great deal of past research. However, user activity and expectations on the Web have now expanded well beyond traditionally studied information seeking process. This has led to an emergence of characteristically different search tasks, such as those of entertainment purposes [27]. As such, it is the right time to revisit and verify these findings. This paper is an attempt in this vein of research.

In recent years the emotion aspect of the information seeking process has gained much attention. For example, several works studied the emotional impact of search tasks within the ISP [29, 1, 25]. [29] examined how participants emotion responses are influenced by tasks of different nature. In particular, their results indicate that (i) artificial tasks have higher uncertainty and less sense of ownership than genuine search tasks, and (ii) more complex search tasks have lower positive emotions and more uncertainty before and after searching. Similar findings were earlier reported also by [1] in an information seeking activity. They concluded that users' emotions progressively transit from positive to negative valence as the degree of task difficulty increases. They also found that emotions both interweave with different physiological, psychological and cognitive processes during an information seeking process, and form distinctive patterns according to specific tasks [2]. However, [25] re-

ported no significant relationship between searchers' mood and search tasks due to the complexity involved in such studies. In another study, [18] investigated the relationship between the subjective (e.g. happiness levels, feeling lost during search, etc.) and objective (e.g. search outcomes and search task characteristics, etc.) factors in the ISP. Their results show that "higher happiness levels before the search and during the search correlate with better feelings after the search, but also correlate with worse search outcomes and lower satisfaction, suggesting that, perhaps, it pays off to feel some 'pain' during the search in order to 'gain' quality outcomes" [18]. Despite their valuable insights into various aspects of emotion in the ISP, they do not provide insights into how the trend of emotion varies for search sessions with different search intents. This paper investigates the answer to this research question.

The interaction aspect of the ISP is by far the most studied aspect in both research and practice. One of the important research topics is to understand the intent of search activity. Broder [9] studied Web query logs and categorised the submitted queries into three main search intents: informational, navigational, and transactional. The important outcome of this work was the realisation of the existence of search intents other than the informational intent, which was traditionally assumed to be dominant. In a more recent study, other search intents have been identified, such as re-finding [37] and entertainment [13]. Teevan et al. [37], by analysing Yahoo! Web query logs showed that up to 40% of the submitted queries were re-finding queries. Elsweiler et al. [12], in an empirical study, showed that it is possible to isolate re-finding behaviour in the logs through various qualitative and quantitative analyses. Other researchers investigated the possibility of predicting different search intents from the query log data. For example, Shen et al. [33] proposed a novel model for user intent, leveraging search sessions by learning intermediate hidden-dynamics between intent class labels and user behaviour variables. In general, query intent prediction can be categorised into two methods: context-aware and non-context-aware [33]. Non-context-aware methods learn a model by relying only on the current searchers' behaviours, e.g current submitted query and/or click-though data [24] whereas context-aware methods taken into account both current and past searchers' behaviour throughout the search session [10]. The underlying assumption behind the context-aware methods is that adjacent user behaviour is semantically related and follows the same search intent. It has been shown that context-aware methods outperform the non-context-aware ones [33]. In contrast to previous work, where the main task is to optimise the retrieval results with respect to the query intent, this paper studies search behaviour with different intents and tries to compare their characteristics across sessions.

Seeking Entertainment on the Web: Due to the ubiquity of the emotionally-rich content on the Web (e.g., news, music, movies, etc.), it is important to understand the search behaviour which has an entertainment aspect. It is not controversial to state that entertainment has an important role in human life. The consumable media on the Web, which is particularly easy to access nowadays, provides people with vast varieties of content, making it the most popular type of entertainment. It has therefore been of particular interest to understand its uses and effects on people. Zillmann [39] was instrumental in the development and establishment of a range of theories in this domain. His *mood management* theory explains why and how the audience seeks entertainment.

The underlying speculation of the mood management theory is that hedonistic motivation is a key factor in affecting the entertainment selection process [42]. Zillmann and his colleagues posit that entertainment choices are a reflection of a basic human need to enhance or retain positive states, and to lessen or steer clear of negative ones [28]. They have suggested two possible states in which there may be a need for regulation: physiological arousal and affect. In the case of physiological arousal, mood management theory suggests that users might be over-stimulated (i.e., stress) or under-stimulated (i.e., boredom). An individual experiencing such states will choose their entertainment content according to their expectations of what would lead them back to an optimal state [28]. In the case of affect states, mood management theory suggests that users might be in negative (i.e. dysphoric) or positive (i.e. upbeat) moods. An individual experiencing negative affects will choose entertainment content that helps them to alleviate or diminish the negative mood, and those experiencing positive affects will choose entertainment content that helps them to intensify or prolong their state [28].

Recently, there was an increased effort in the research community in developing IR systems that cope with search processes for entertainment intents. The "Entertain Me" [5] and "Search4FUN" workshops [14] began to explore this area and build resources, such as test collections and evaluation benchmarks. Besides the works published in these workshops, some research has been published that attempts to use emotion in information retrieval. We review this in the remaining part of this section. [13] attempted to understand the needs and motivation underlying leisure-based activities in the context of television viewing. They reported that the nature of the need and motivation are different between leisure-based and work-based situations where most of the needs reported for leisure-based situation are motivated by a desire to change mood, emotion, or arousal level. Another research related to leisure-based activity is the work by [32] which studies pleasure reading behaviour. Their results show that pleasure readers find information without having any prior information need. In addition, [41] and in a more in-depth version, [15], study the characteristics of casual-leisure search as an example of an exploratory search scenario. They reported that in casual search, the motivation is not to resolve an information need but is rather hedonistic, e.g. entertainment driven. Given their results, [15] define a casual-leisure information behaviour model highlighting the key differences between casual-leisure scenarios and typical information behaviour theory. Despite these attempts, our understanding of entertainment-based search processes is still in its infancy. Given the emergence of entertainment seeking intentions on the Web and the unexplored characteristics of them, in this paper we investigate two entertainment seeking tasks: entertainment by adjusting arousal level and entertainment by adjusting mood.

3. EXPERIMENTAL METHODOLOGY

3.1 Experiment Design

This study used a within subject design. The independent variable was the search intents (with four levels: "Informa-

tion Seeking" (INS), "Information Re-finding" (INF), "Entertainment by adjusting Arousal" (ENA), "Entertainment by adjusting Mood" (ENM)), which was controlled by the simulated search task given to the participants (see Section 3.2). We did not perform any control on the number of relevant and irrelevant results in order to simulate a real search scenario situation as much as possible. The dependent variables are the qualitative (gathered through questionnaire) and quantitative (gathered through system interaction logging) data.

3.2 Tasks

We prepared four search task scenarios, each simulating different search intent. The search tasks were presented using the structural framework of simulated need situations [8]. By doing so, we introduced short cover stories that helped us describe to our participants the source of their information need, the context of the situation and the problem to be solved. This facilitated a better understanding of the search objective and introduced a layer of realism, while preserving well-defined relevance criteria. In the following, each of the search tasks is explained in detail.

3.2.1 INS Task

This search task simulates the information seeking search scenario. Information seeking is the most studied topic in the field. This task is designed as a control group since the majority of cognition and emotion findings in the past were based on a similar search task intent. For this search task, we prepared a number of search topics that covered a variety of contexts in order to capture participants' interests as best as possible. The topics, presented in Table 1, were all checked manually, prior to the experiment, to ensure the availability of relevant documents. The simulated search scenario for INS task was as follows: *"Imagine you have graduated recently and are going to interview for a job in a local company. As part of the interview process, you are asked to explain and expand on the area you will be working on. You feel very enthusiastic about the interview; however, due to your lack of knowledge you would like to find out more about this particular topic before taking part in the interview."* Each participant was then asked to choose one of the topics that they were unfamiliar with but consider interesting. Using the video retrieval system, they had to find as many relevant videos as possible so that they could construct a good knowledge about their selected topic.

Table 1: Search topics for the information seeking task scenario.

Obtaining information regarding contraception methods.
Investigating new knowledge on global warming.
Formulating an opinion about existing social networking sites.

3.2.2 INF Task

This search task simulates the information re-finding search intent. There are two differences between this task and the previous one: (i) there is only one document that can satisfy the information need of the searcher, and (ii) the searcher has seen the relevant documents at some point before initiating the search process and is now attempting to re-find it. The similarity between this task and the previous one is that in both cases searchers have an information need.

For this search task, we prepared a number of videos that covered a variety of contexts in order to capture participants' interests as best as possible. The videos were intentionally selected as they would likely be very hard to find because they lacked textual description. The motivation was to simulate a challenging information re-finding process where the recalled terms are very ambiguous and do not lead directly to the relevant item. This would better represent many realistic re-finding tasks, when the user cannot recall the exact description of the item they are looking for. Of course this only considers one spectrum of re-finding tasks, and more exhaustive studies of such tasks needs to be done in the future. The simulated search scenario for INF task was as follows: *"Imagine you are discussing a video which you have seen few days ago with your friends. They are interested in seeing the video and have asked you to send them a link to it. You can remember the content of the video but you cannot remember its title or any textual information which can help you in retrieving it."* Each participant was then asked to select and watch one of the three videos presented to them (an animal[2], martial arts[3], or a science video[4]) that they are unfamiliar with and consider interesting. Once participants watched the video, they were asked to find it as fast as possible.

3.2.3 ENA Task

This search task simulates the entertainment-based search intent where searchers adjust their arousal level. The main difference between this task and the two tasks before is that the primary need is hedonistic rather than informational. Therefore, to accurately simulate such search processes, we avoid introducing any explicit information need. Thus, for this search task, we did not provide any pre-prepared search topic. This decision is motivated by the literature in sociology [28] and information seeking and retrieval domains on entertainment [13]. To the best of our knowledge, this is the first time that an entertainment-based search task is simulated in this way. The simulated search scenario for ENA task was as follows: *"Imagine you are working in a factory as a night-guard. You have just finished your routine checks and will be taking out more checks shortly. You are tired so you decide to watch some videos to wake yourself up and make yourself ready for your next round of checks."* Each participant was asked to find as many relevant videos as possible that make them feel excited.

3.2.4 ENM Task

This search task simulates the entertainment-based search intent where searchers adjust their mood. Similar to ENA task, searchers engage in such search processes with hedonistic need rather than informational. Therefore, to accurately simulate such search processes, we avoid introducing any explicit information need. Thus, similar to ENA task, we did not prepared any specific search topic. The simulated search scenario for ENM task was as follows: *"Imagine your boyfriend/girlfriend is travelling and communication access is very limited. It is now a few days since he/she has gone and you are missing him/her very much. You are feeling very sad and in order to change your mood, you have decided to watch some videos."* Each participant was asked to

[2] www.youtube.com/v/wx7rY11qb0k&showinfo=0
[3] www.youtube.com/v/pEUeP7hx8fs&showinfo=0
[4] www.youtube.com/v/P8Cs2t05w74&showinfo=0

find as many relevant videos as possible that make them feel happy.

3.3 Apparatus

For our experiment we used one desktop computer, equipped with a monitor, keyboard and mouse. The computer provided access to a custom-made search interface which allowed the participants to perform their search tasks. The interface was designed such that it logged participants' desktop actions, such as starting, finishing and elapsed times for interactions, mouse movement, and click-throughs using a common system time. Finally, we used entry, post and exit questionnaires in each session.

3.4 Questionnaires

At the beginning of the experiment, the participants were introduced to an *entry questionnaire*, which gathered background and demographic information, and inquired about previous experience with online videos, in particular, browsing and searching habits including their intentions. At the end of each task, the participants completed a *post-task questionnaire*, to elicit subject's viewpoint on certain aspects of the search process. The questions were divided into three sections that covered the encountered task and the cognition and emotion aspects of the search experience. All of the questions included in these questionnaires were a forced-choice type. Finally, an *exit questionnaire* was introduced at the end of the study. In this questionnaire we gathered information about the encountered system as well as the user study in general: which task they preferred and why, and their general comments about the user study.

3.5 Video Retrieval System

For the completion of the search tasks we used a custom-made search environment (named *VideoHunt*) that was designed to resemble the basic layout of existing search services, while retaining a minimum of graphical elements and distractions. VideoHunt works on top of Bing 2.0 API. For every submitted query it returned a list of fifty results (ten results on each page), stripped of their title, snippet and any other metadata. The title of a video was presented as tooltip by moving the mouse over a particular video. Even though this approach introduced our participants into artificial search situations which differ from real-life experiences, it was a necessary trade-off for capturing the browsing action of the user with the retrieved results.

3.5.1 Search Interface

VideoHunt applies a layered architecture approach, similar to that adopted in [3]. The first layer of the interface is dedicated to supporting any interaction that occurs during the early stages of the search process (such as query formulation and search execution). Any output generated during this phase is presented in the second layer. From there, the participants could easily select and preview any of the retrieved clips. The content of a clip is shown on a separate panel, in the foreground, which corresponds to the third layer of our system. The main reason behind this layered architecture was to isolate the viewed content from all possible actions allowing us to separate the time spent viewing video from other actions such as browsing the retrieved results or formulating queries. Upon viewing the clip, the participants had to explicitly indicate the relevance of the video.

3.5.2 User Tracking and Logging

All user actions were monitored and logged by the search interface including their queries and clicks, as well as the length of time spent watching a clip, expanding the result list and formulating a query. In addition, as it has been shown that mouse movements are correlated with user gaze [17], the system captured mouse events to determine the amount of time the user spent on each retrieved result item. This information was all captured in a log file allowing for easy association of the log information with the corresponding questionnaire responses.

3.6 Procedure

The user study was carried out in the following manner. The formal meeting with the participants took place in the laboratory setting. At the beginning of the session the participants were given an information sheet which explained the conditions of the experiment. They were then asked to sign a Consent Form and were notified about their right to withdraw at any point during the study, without affecting their legal rights or benefits. Then, they were given an Entry Questionnaire to fill in.

The session proceeded with a brief tutorial on the use of the search interface with a short training task. After completion of the training task, each participant had to complete four search tasks (explained in Section 3.2), one for each level of search process intentions (see Section 3.1). To negate the order and fatigue effects we counter-balanced the task distribution using a Graeco-Latin Square design. The subjects were asked every time to provide judgment for any video that they watch, and were given 10 minutes to complete their task, during which they were left unattended to work. At the end of each task, the subjects were asked to complete a post-task questionnaire. Questions in the post-task questionnaire were randomised to avoid the effect of fatigue. Between each task, a cooling-off period was applied to avoid the carry-over effect. An exit questionnaire was administered at the end of the session. Finally, the participants were asked to sign a payment form, prior to receiving the payment of £12.

Each study took approximately 120 minutes to complete; this is from the time they accepted the conditions until they signed the payment receipt. Users could only participate once in the study. The total cost of the evaluation was £348, including the cost of the pilot studies. A user study with the procedure explained above was conduced over a period of 10 days from 16th to 26th of July 2012. The results of these studies are presented in Section 4.

3.6.1 Participants

Participants consisted of 24 healthy participants with equal gender distribution (12 female and 12 male) all under the age of 41, with the largest group between the ages of 18-23 (45.8%) followed by the group between ages of 24-29 (36%). Participants tended to have a high school diploma or equivalent (4.16%), some college degree (4.16%), bachelor (41.66%) or graduate degree (50%). They were primarily students (62.5%), though there were a number of self-employed (16.6%), not employed (4.16%) and employed by a company or organisation (16.6%).

3.6.2 Pilot Studies

Prior to running the actual user study, a pilot study was performed using 5 participants to confirm that the process worked correctly and smoothly. A number of changes were made to the system based on feedback from the pilot study. The changes consisted of modifications to the questionnaires to clarify questions, modifications to the system to improve logging capabilities and improvements to the tasks. After the final pilot, it was determined that the participants were able to complete the user study without problems and that the system was capturing all necessary data. One of the important outcomes of the pilot study was that participants preferred the questions where the search session was broken down into three phases (i.e. beginning, middle, and end), compared to questions where a finer granularity was applied (e.g. broken down into four phases). Therefore, the questionnaire was adapted accordingly. However, for better granularity, the interaction log was divided into four phases.

4. RESULTS AND DISCUSSION

In Section 4.1, we first discuss the task perception expressed in the questionnaire. Following this, we discuss our analysis on the characteristics of the four search tasks outlined in Section 3.2; focusing on the cognition, emotion, and interaction aspects. Finally, we discuss the predictability of search task types based on features derived from the user interaction with the system in Section 4.2.

4.1 Qualitative and Quantitative Analysis

4.1.1 Task Perception

Figures 1 shows the box plots for the qualitative analysis of users' perception of the four tasks (i.e. INS, INF, ENA, and ENM). Each box plot reports data aggregated from 24 participants, along with five key statistics: the minimum, first, second (median), third, and maximum quartiles.[5] We performed an ANOVA test between measures obtained at each phase, across four search tasks for each user to check the significance of the difference among them. The test is suitable for this data as we have four groups of data, therefore we need to compare four means and variances. We use (*) and (**) to denote the fact that a measure had results different across four search tasks with the confidence levels ($p < 0.05$) and ($p < 0.01$), respectively.

In the post-task questionnaire we measured participants' perception of their performed task in terms of the difficulty of the task, the familiarity of the participant with the task, the extent to which they found the task stressful, interesting and clear by asking the following question *"The task we asked you to perform was [easy/stressful/interesting/clear/familiar] (answer: 1: "Strongly Disagree", 2: "Disagree", 3: "Neutral", 4: "Agree", 5: "Strongly Agree")"*. The results shown in Figure 1 indicate that participants found the INF task difficult and stressful, followed by the INS task, whereas they found other two tasks easy and not-stressful (the differences were statistically significant). The difference in the answer provided by the participants for interesting, clear and familiar measures is not statistically significant. In the post-task questionnaire we also asked the opinion of the participants with respect to the following statement *"I had enough time to do an effective search.] (answer: 1: "Strongly Disagree", 2: "Disagree", 3: "Neutral", 4: "Agree", 5: "Strongly Agree")"*. The results show that they found the given time enough to do an effective search task: INS: M=4.0 SD=0.88; INF: M=3.8 SD=0.96; ENA: M=4.0 SD=0.97; ENM: M=4.3 SD=0.56 (the differences are however not statistically significant across the tasks).

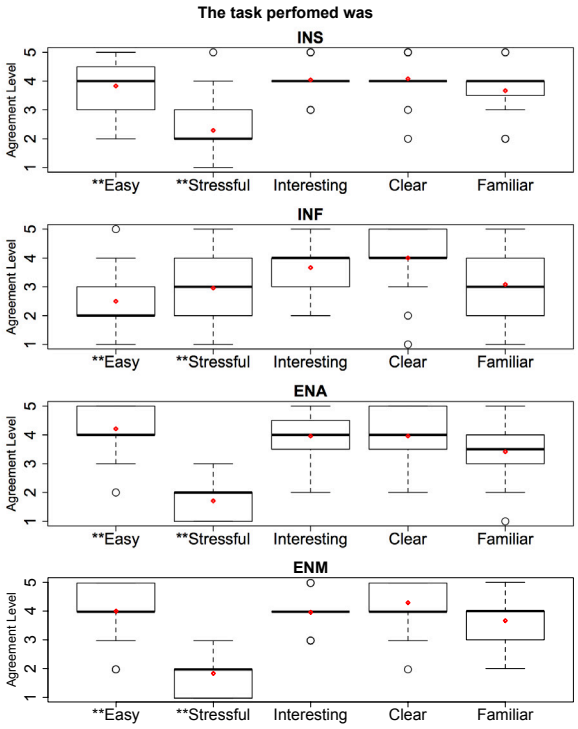

Figure 1: Box plot of the task perception based on the information gathered from 24 participants questionnaire. The diamond represents the mean value. (*) and (**) indicate confidence levels ($p < 0.05$) and ($p < 0.01$) respectively.

4.1.2 Main Results

Figures 2, 3 and 4 show the line plots for the qualitative analysis of users' cognition, emotion, and interaction aspects for the four search tasks (i.e. INS, INF, ENA, and ENM). Each line plot reports the data aggregated from the 24 participants, along with the mean and error bar for three phases of a search process: the beginning, middle and end. We performed the ANOVA test at each phase to check the significance of the difference among the tasks at each phase. We use (*) and (**) to denote the fact that a measure had results different across four search tasks with the confidence levels ($p < 0.05$) and ($p < 0.01$), respectively.

The key finding which emerged from the questionnaire is that search tasks with different intentions have varying characteristics in terms of cognition, emotion, and interaction aspects. From a cognition point of view, the complexity and dynamic concepts associated with an information seeking process do not hold for all four search intents, e.g. not all the search processes have an information need which starts from an ASK. From an emotion point of view, the uncertainty principle of Kuhlthau [22] does not hold (at least for the initiation phase) e.g. not all the search processes start

[5] Further information can be found in [26].

from negative emotions, such as anxiety associated with the realisation of lack of knowledge. However the end of the search process tends to agree with Kuhlthau ISP model [21] since, in the case of a successful search process, participants express satisfaction and in other case anxiety and/or anger. Though Kuhlthau's ISP model is originally proposed for information seeking tasks, it is assumed to hold in IR scenarios [23, 40]. The main result which emerged from interaction analysis is that participants interact differently with the search system throughout the search session when they have different intents. The interaction differences are in terms of the characteristic of the submitted queries and interaction behaviour with the retrieved results. In the remainder of this section we discuss each of these aspects in more details.

4.1.3 Cognition Aspect

To investigate the cognitive aspect of the our search tasks, we study the difficulty of query formulation (associated with complexity of search process), certainty of what to search for and what is relevant (associated with the dynamic concept of search process). Figure 2 presents the results for three questions posed in the post-task questionnaire: *(1) "At the [beginning/Middle/End] of the search session, I was certain about what I was going to search for.", (2) "At the [beginning/Middle/End] of the search session, I had difficulty formulating queries.", and (3) "At the [beginning/Middle/End] of the search session, I didn't have a clear idea what videos would be relevant." (answer: 1: Strongly Disagree, 2: Disagree, 3: Neutral, 4: Agree, 5: Strongly Agree).*

The findings show that for the INS task, at the early stage of the search session participants experience uncertainty of what to search for and have only a vague idea about what videos would be relevant. As they progress through the search session, their uncertainty decreases and their concept of relevant videos becomes more clear. Formulating queries also becomes less difficult as they progress through the search session. The cognitive process reported for the INS task follows the literature in the information seeking and retrieval community such as Belkin et al. [7] and Kuhlthau [22].

In contrast to INS, for the INF task, the findings show that participants at the beginning of the search session experience a high level of certainty, having little problem formulating queries and knowing what video would be relevant. As they progress through the search session, the uncertainty increases, formulating queries became more difficult and what a relevant video would be becomes more vague. Although the cognitive state at the later stage of the INF task is a by-product of the fact that they were unsuccessful in their search task, the cognitive state at the early phase of the search process is contrary to general assumptions in the IR community.

Finally, for the ENM and ENA tasks, the results show a consistent level of certainty throughout the search session. Although the ENA task is similar to INS in terms of answers given to the "difficulty of formulating queries" and "vagueness of the concept of relevance" questions, the ENM task shows consistent behaviour for all three questions across different phases, which indicates yet another cognitive behaviour pattern.

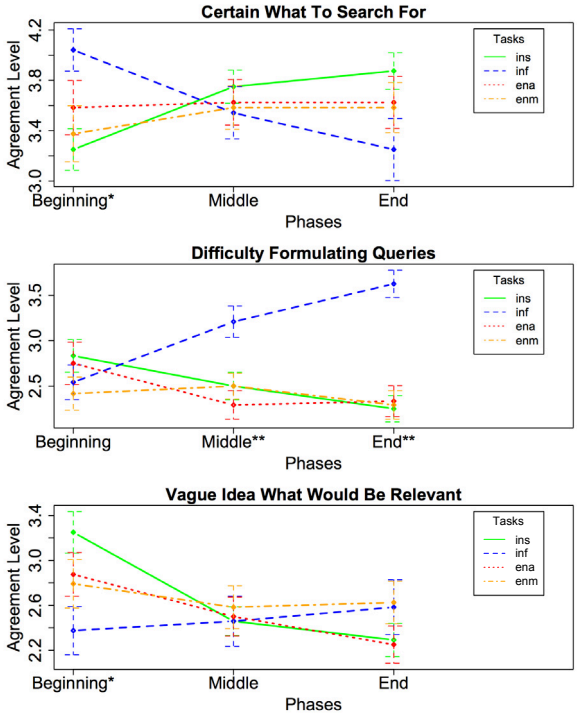

Figure 2: Line plot of the cognitive aspect based on the information gathered from 24 participant questionnaires. (*) and (**) indicate confidence levels ($p < 0.05$) and ($p < 0.01$), respectively.

4.1.4 Emotion Aspect

To investigate the emotion aspect of our search tasks, we study participants' emotion at three stages of the search session using six Ekman emotions [11] plus neutral, anxiety and satisfaction. Figure 3 presents the results for the posed questions in the post-task questionnaire: *(1) "At the [beginning/Middle/End] of the search session I experienced [Sadness/Happiness/Anger/Surprise/ Fear/Disgust/Anxiety/Satisfaction/Neutral]." (answer: 1: Yes, 2: No).*

The results for the INS task show that participants' satisfaction level increases as they progress through the search session. Although this is inline with previous literature (i.e. the increase of the feeling of satisfaction as a result of alleviating ASK), the results for anxiety is not inline with previous work, such as uncertainty principle of Kuhlthau [22]. This is because our findings show that the anxiety level for the INS task at the beginning of the search session is low and remains low throughout the session. Surprisingly, the results show that for this task, the majority of participants are neutral (experiencing no emotion) at the beginning of the search session and as they progress through, they experience emotion (reporting less neutral). In addition, we observe a diverse range of emotion reported at the middle phase of the INS search process (e.g. surprise, fear, disgust, anger) which we assume is a result of the topics and/or watched videos, rather than the search process itself.

For the INF task, the findings show a similar trend for a neutral level, as with what we reported for the INS task. However, we observe that for the INF task, anger constantly increases as the search session progresses, whereas for the

INS task, satisfaction increases instead. Experiencing anger can be a result of the cognitive process participants go through in their search session for this task. As we discussed earlier, the participants did not succeed in completing this task, hence the experience of anger is likely a result of this. In addition, the anxiety level increases as participants progress from the beginning to the middle phase of the search session, but surprisingly, it decreases from the middle to the end of the search session. This can be as a result of transforming anxiety to other negative emotions, e.g. sadness or disgust.

For the ENA and ENM tasks, the results show that fewer participants are neutral at the beginning of the search session compared to INS and INF, and this number further decreases as participants progress through the search session. The findings show that participants' experience of happiness for these two tasks are significantly higher than the INS and INF tasks, indicating the hedonistic meditation underlying such processes. These findings confirm Zillmann theory of mood management [42]. Similar to INS, for both ENA and ENM tasks, participants satisfaction level increases as they progress through the search session, indicating that the underlying need is getting satisfied. The results show that the feeling of satisfaction can also be experienced if the search task does not have a clear IN, which in these tasks relates to a hedonistic need. Another interesting observation is the variety of the emotions experienced at the beginning of the search session for the ENM and ENA tasks. This shows that it is possible that participants begin their search session experiencing emotions other than anxiety or any other negative emotions, as is the case for ENM (e.g. the feeling of sadness) and ENA (e.g. the feeling of surprise). This particular result may be an artefact of our experimental settings and needs to be explored further.

4.1.5 Interaction Aspect

To investigate the interaction aspect of our search tasks, we study a number of features extracted from interaction log data. As the distribution of the search task completion time can be very different, we aggregate the interaction data for each quartile. For each phase, we first calculated general statistics such as the average number of clicked, liked, and disliked videos, as well as the average number of times "Next Page" and "Previous Page" buttons were clicked, the average number of documents the mouse hovered over, and the average number of formulated queries. Then, we calculated the average query formulation time, video viewing time, and mouse hovering duration over the retrieved results. Finally, we calculated the similarity of the submitted queries as well as the similarity of the titles of the viewed videos. The similarity value was calculated using cosine similarity. Figure 4 presents the results for the analysis of the interaction with respect to the features explained above.

The analysis of interaction for the INS task shows that participants at each phase formulate fewer queries in comparison to other tasks, and the number of issued queries and the query formulation time decreases as they progress through the search session. The query formulation time can indicate the cognitive effort required. Therefore, the results obtained for the query formulation time is inline with what we observed from questionnaire results with respect to the cognitive aspect of the INS task where the difficulty of formulating queries decreases as the participants progress through the search session. Another interesting aspect is how they interact with the retrieved results. Participants mainly visited videos present on the first page, and checked a relatively lower number of videos at each phase compared to other tasks. The average video viewing time steadily increases as well as the number of videos judged as relevant.

On the other hand, for the INF task we observe a different interaction behaviour pattern. The number of formulated queries and the similarity between the formulated queries are the highest among the four tasks across all four phases, while the average query formulation time is the lowest. In contrast to the INS task, participants not only visited the retrieved results in the main page but also visited other results pages. They have also checked a larger number of videos at each phase. The time spent viewing selected videos was close to zero, showing that they were certain of what they were looking for and simply wanted to verify the selected video was the correct one. This results holds for difficult re-finding task scenario (see Section 3.2) and may not hold for other spectrum of such a task.

For the ENA and ENM tasks we observe that the number of formulated queries and the average query formulation time are between that of the INS and INF tasks across four phases. The greatest number of submitted queries is in the first phase, however from the second phase onwards the number of formulated queries remains almost static. What distinguishes the formulated queries between the ENA and ENM tasks is the similarity of the formulated queries at each phase. For the ENA task the similarity of formulated queries decreases as the search session progresses, whereas for the ENM task the similarity of formulated queries and the number of formulated queries follows a similar trend across all phases. Another aspect of interaction history which differentiates between the ENA and ENM tasks is the way participants interact with the retrieved results. For the ENA task, similar to the INS task, the visited documents were mainly from the first page across all phases, whereas for the ENM task, participants visited results from other pages as well, in particular during the third phase.

The next question this work addresses is whether we can learn a model from such interactions so that we can accurately predict the search task intentions. Furthermore, to what extent can we accurately predict search task intentions through the search process?

4.2 Prediction of Search Task Intentions

As discussed in Section 1, identification of the intentions behind the search processes can help search engines to better satisfy users' needs. Therefore, in this section, we investigate our second research question on whether the search intents can be modelled to predict given searchers' interaction data with the system. For this purpose, we used the set of features extracted from interaction log data explained in Sections 4.1.5.

For our four search intents, we have a multinomial classification problem where the classes are "INS" (indicating the participant had information seeking), "INF" (indicating the participant had information finding), "ENA" (indicating the participant had emotion need by adjusting arousal level), and "ENM" (indicating the participant had emotion need by adjusting mood). We used SMO, an implementation of SVM in Weka,[6] to discriminate between the four classes

[6] http://www.cs.waikato.ac.nz/ml/weka/

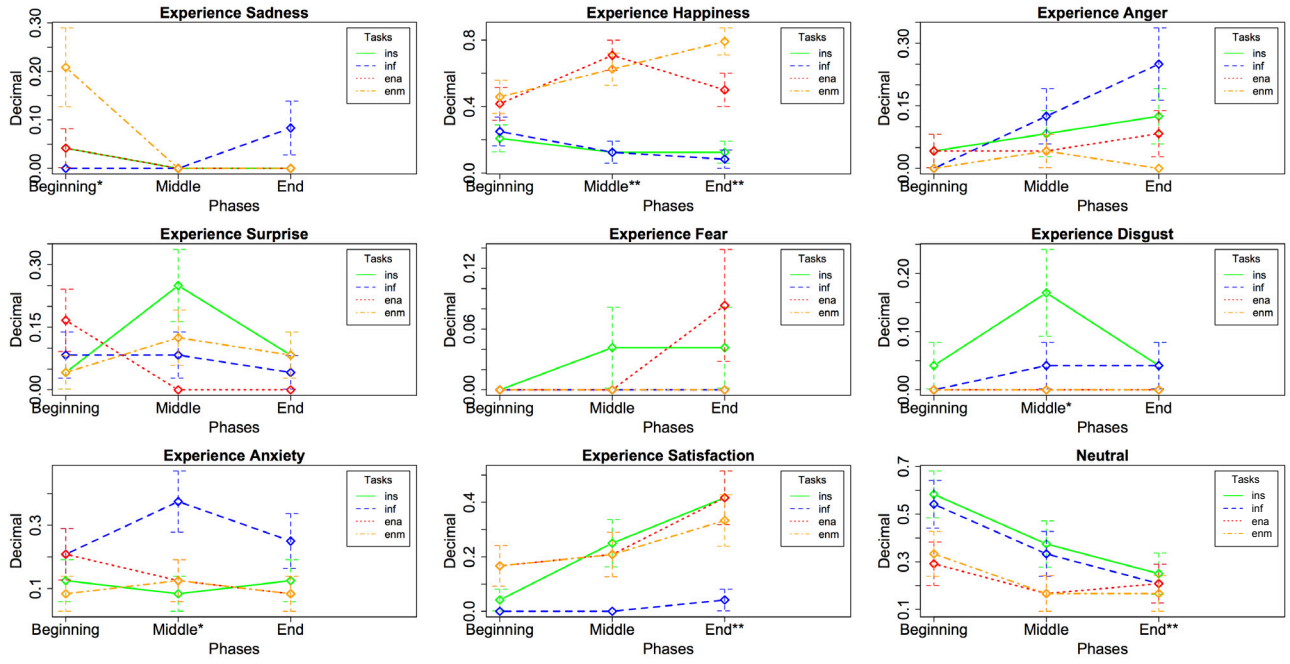

Figure 3: Line plot of the emotion aspect based on the information gathered from 24 participants questionnaire. (*) and (**) indicat confidence levels ($p < 0.05$) and ($p < 0.01$), respectively.

explained above. We trained our models using a normalised polynomial kernel which in the majority of cases outperformed other SVM kernels (e.g. polynomial and radial-basis) based on our analysis, not presented due to the space limits.

Table 2 shows the classification performance averaged over the 24 participants of the study at different phases of the search process. We measured the accuracy of the model (i.e. fraction of items in the test set for which the models' predictions were correct) using 10-fold cross-validation. We obtained more than 100% improvement over a model based on the discrete uniform distribution. The results indicate that at each phase of the search process we are able to successfully predict the search task intentions given the features extracted from interaction logs. The accuracy of the prediction increases as the searcher progresses through the search session. Our findings indicate that it is possible to build effective intention-aware search technologies in which the retrieved results and provided functionality adapt to the needs of the user.

Table 2: The accuracy of the search intent prediction at each phase. Phases are presented as columns. The best performing feature set for each dimension is highlighted in bold.

	Search Process Phases			
	Phase1	Phase2	Phase3	Phase4
Baseline	25%	25%	25%	25%
SVM	**51.04%**	**53.12%**	**56.25%**	**57.29%**

5. CONCLUSIONS

In this paper we investigated the cognitive, emotion, and interaction aspects of four search tasks namely information seeking (INS), re-finding (INF), entertainment adjusting arousal level (ENA) and entertainment adjusting mood (ENM). In order to do so, we devised four search tasks, each simulating one of these intents. Using a video retrieval system as a use case, we conducted a user study with 24 participants. We analysed the characteristics of each aspect across different phases of the search process. Our findings show differences in cognition, emotion and interaction aspects of search processes for different search tasks. In particular, from a cognition point of view, the complexity and dynamic concepts associated with an information seeking process do not hold for all four search tasks types, e.g. not all the search processes have an information need which starts from an ASK. The cognitive process reported for the INS task shows a parallel with prior findings in the information seeking and retrieval community such as Belkin et al. [7] and Kuhlthau [22]. In contrast, for the INF task, the findings show that participants at the beginning of the search session experience a high level of certainty, having little problem formulating queries and knowing what video would be relevant. Finally, for the ENM and ENA tasks, the results show a consistent level of certainty throughout the search session.

From an emotion point of view, the uncertainty principle of Kuhlthau [22] does not hold (at least for the initiation phase) e.g. not all the search processes start from negative emotions, such as anxiety associated with the realisation of lack of knowledge. However the end of the search process tends to agree with Kuhlthau ISP model [21] since, in the case of a successful search process, participants express satisfaction and in other case anxiety and/or anger. Surprisingly, the results show that for INS task, the majority of participants are neutral (experiencing no emotion) at the beginning of the search session and as they progress through, they experience emotion (reporting less neutral). For the INF task, the findings show a similar trend for a neutral

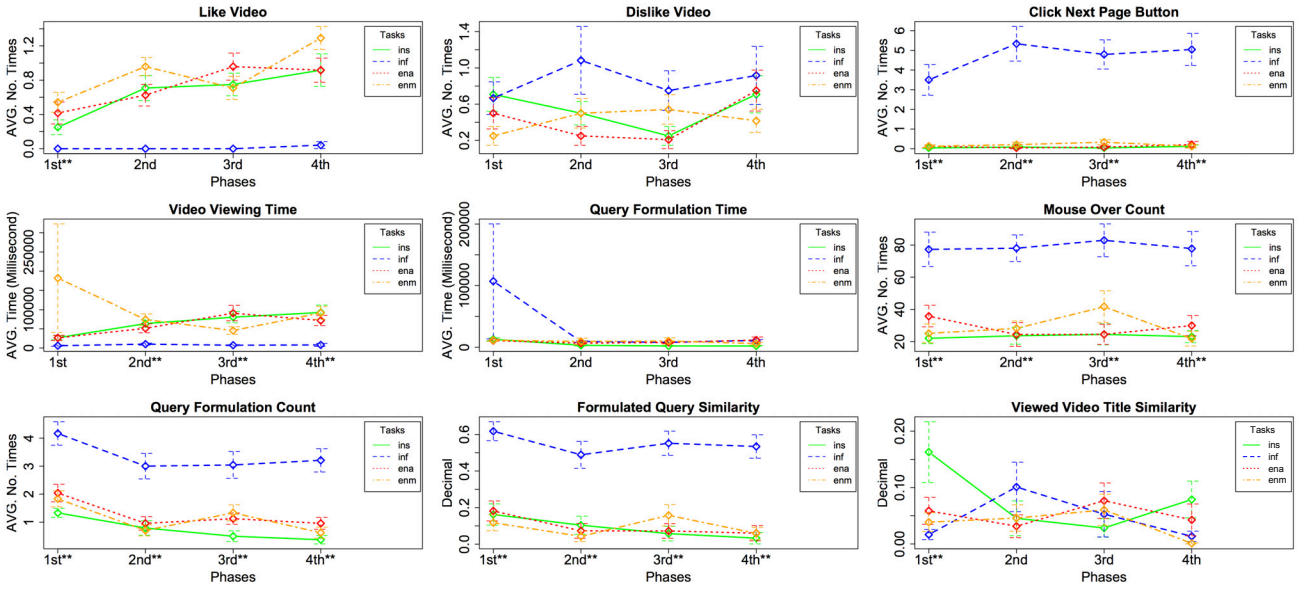

Figure 4: Line plot of the interaction aspect based on the information gathered from 24 participants interaction log data. (*) and (**) indicat confidence levels ($p < 0.05$) and ($p < 0.01$), respectively.

level, as with what we reported for the INS task. For the ENA and ENM tasks, the results show that fewer participants are neutral at the beginning of the search session compared to INS and INF, and this number further decreases as participants progress through the search session.

From an interaction point of view, the main result which emerged from analysis is that participants interact differently with the search system when they have different search tasks. The interaction differences are in terms of the characteristic of the submitted queries and interaction behaviour with the retrieved results. For the INS task, the findings show that participants at each phase formulate fewer queries in comparison to other tasks. Participants also mainly visited videos present on the first page. In contrast to the INS task, for the INF task, the number of formulated queries and the similarity between the formulated queries are the highest among the four tasks across all four phases. In addition, participants not only visited the retrieved results in the main page but also visited other results pages. What distinguishes the formulated queries between these two tasks is the similarity of the formulated queries at each phase.

We further study the possibility of predicting search intents at different phases of a search process, given features extracted from interaction log data. We have achieved more than 100% improvement over our baseline (discrete uniform distribution) on our search intents prediction task, given searchers' interaction data. Discrete uniform distribution is a naïve baseline, however there are no previous studies for comparison. Our prediction accuracy sets the baseline for future studies. The features used for our prediction are computationally inexpensive and easy to calculate. A more exhaustive exploration of features will be studied in future work.

The major implication of our work is in developing effective intent-aware search technologies. Given the fundamental differences in cognition, emotion and interaction aspects of the studied search intents, the next step can be to develop techniques facilitating effective seeking processes and in turn satisfying searchers. In addition, search behaviour characteristics at various phases can be easily detected and exploited for predicting user intents. This may open up avenues for revisiting current intent detection approaches.

Finally, the findings of this paper are limited to the video retrieval domain and we do not generalise it to other domains. Even though significant differences in cognition, emotion and interaction were detected across different search task types, we acknowledge the limitation of a lab-based study. However, our findings motivate the exploration of similar hypotheses in other domains. In future work, we plan to evaluate additional features to improve our prediction accuracy. Further, we want to continue to study how understanding search sessions can be leveraged to improve user satisfaction measures, and possibly build search engines that adapt based on user search process intentions.

6. ACKNOWLEDGEMENT

This work was supported by the EU FP7 LiMoSINe project (288024).

7. REFERENCES

[1] I. Arapakis, J. M. Jose, and P. D. Gray. Affective Feedback: An Investigation into the Role of Emotions in the Information Seeking Process. In *SIGIR*, 395–402, 2008.

[2] I. Arapakis, I. Konstas, and J. M. Jose. Using Facial Expressions and Peripheral Physiological Signals as Implicit Indicators of Topical Relevance. In *MM*, pages 461–470, 2009.

[3] I. Arapakis, Y. Moshfeghi, H. Joho, R. Ren, D. Hannah, and J. M. Jose. Enriching User Profiling with Affective Features for the Improvement of a Multimodal Recommender System. *CIVR*, 2009.

[4] N. Belkin. Helping People Find What They Don't Know. *Communications of the ACM*, 43(8):58–61, 2000.

[5] N. Belkin, C. Clarke, N. Gao, J. Kamps, and J. Karlgren. Report on the SIGIR Workshop on Entertain Me: Supporting Complex Search Tasks. In *ACM SIGIR Forum*, volume 45, 51–59. 2012.

[6] N. Belkin, C. Cool, A. Stein, and U. Thiel. Cases, Scripts, and Information-Seeking Strategies: On the Design of Interactive Information Retrieval Systems. *Expert systems with applications*, 9(3):379–395, 1995.

[7] N. Belkin, R. Oddy, and H. Brooks. Ask for Information Retrieval: Part i. Background and Theory. *Journal of Documentation*, 38(2):61–71, 1982.

[8] P. Borlund and P. Ingwersen. The Development of a Method for the Evaluation of Interactive Information Retrieval Systems. *Journal of Documentation*, 53(3):225–250, 1997.

[9] A. Broder. A Taxonomy of Web Search. In *ACM SIGIR forum*, volume 36, 3–10. 2002.

[10] Z. Cheng, B. Gao, and T.-Y. Liu. Actively Predicting Diverse Search Intent from User Browsing Behaviors. In *WWW*, 221–230. 2010.

[11] P. Ekman and R. J. Davidson. *The Nature of Emotion: Fundamental Questions*. Oxford University Press, 1994.

[12] D. Elsweiler, M. Harvey, and M. Hacker. Understanding Re-finding behavior in Naturalistic Email Interaction Logs. In *SIGIR*, 35–44. 2011.

[13] D. Elsweiler, S. Mandl, and B. Kirkegaard Lunn. Understanding Casual-Leisure Information Needs: a Diary Study in the Context of Television Viewing. In *IIiX*, 25–34, 2010.

[14] D. Elsweiler, M. Wilson, and M. Harvey. Searching4fun. 2012.

[15] D. Elsweiler, M. Wilson, and B. Lunn. Understanding Casual-Leisure Information Behaviour. *Library and Information Science*, 211:241, 2011.

[16] K. E. Fisher, S. Erdelez, and L. McKechnie. *Theories of Information Behavior*. Information Today Inc, 2005.

[17] Q. Guo and E. Agichtein. Towards Predicting Web Searcher Gaze Position from Mouse Movements. In *CHI Extended Abstracts*, 3601–3606, 2010.

[18] J. Gwizdka and I. Lopatovska. The Role of Subjective Factors in the Information Search Process. *JASIST*, 60(12):2452–2464, 2009.

[19] P. Ingwersen. *Information Retrieval Interaction*. Number s 246. Taylor Graham London, 1992.

[20] P. Ingwersen. Polyrepresentation of Information Needs and Semantic Entities: Elements of a Cognitive Theory for Information Retrieval Interaction. In *SIGIR*, 101–110, 1994.

[21] C. Kuhlthau, J. Heinström, and R. Todd. The 'information Search Process' Revisited: Is the Model Still Useful. *Information Research*, 13(4):13–4, 2008.

[22] C. C. Kuhlthau. A Principle of Uncertainty for Information Seeking. *Journal of Documentation*, 49(4):339–355, 1993.

[23] C. C. Kuhlthau. Accommodating the User's Information Search Process: Challenges for Information Retrieval System Designers. *Bulletin of ASIS&T*, 25(3):12–16, 2005.

[24] X. Li, Y. Wang, and A. Acero. Learning Query Intent from Regularized Click Graphs. In *SIGIR*, volume 339, 346. 2008.

[25] I. Lopatovska. Searching for good mood: Examining relationships between search task and mood. *ASIST*, 46(1):1–13, 2009.

[26] R. McGill, J. W. Tukey, and W. A. Larsen. Variations of Box Plots. *American Statistician*, 32(1):12–16, 1978.

[27] Y. Moshfeghi and J. M. Jose. Role of Emotion in Information Retrieval for Entertainment (Position Paper). *Searching4FUN Workshop in ECIR*, 2012.

[28] M. B. Oliver. Mood management and Selective Exposure. *Communication and Emotion: Essays in Honor of Dolf Zillmann*, 85–106, 2003.

[29] A. Poddar and I. Ruthven. The Emotional Impact of Search Tasks. In *IIiX*, 35–44. 2010.

[30] D. Robins. Shifts of Focus in Information Retrieval Interaction. In *ASIST*, volume 34, 123–134, 1997.

[31] T. Saracevic. Relevance: A Review of and a Framework for the Thinking on the Notion in Information Science. *JASIST*, 26(6):321–343, 1975.

[32] C. Sheldrick Ross. Finding without Seeking: The Information Encounter in the Context of Reading for Pleasure. *Information Processing & Management*, 35(6):783–799, 1999.

[33] Y. Shen, J. Yan, S. Yan, L. Ji, N. Liu, and Z. Chen. Sparse Hidden-Dynamics Conditional Random Fields for User Intent Understanding. In *WWW*, 7–16. 2011.

[34] J. A. Singer and P. Salovey. Mood and Memory: Evaluating the Network Theory of Affect. *Clinical Psychology Review*, 1988.

[35] A. Spink, H. Greisdorf, and J. Bateman. From Highly Relevant to Not Relevant: Examining Different Regions of Relevance. *Information Processing & Management*, 34(5):599–621, 1998.

[36] R. Taylor. Question-Negotiation an Information-seeking in Libraries. Technical report, DTIC Document, 1967.

[37] J. Teevan, E. Adar, R. Jones, and M. Potts. Information Re-Retrieval: Repeat Queries in Yahoo's Logs. In *SIGIR*, 151–158. 2007.

[38] C. van Rijsbergen. (Invited Paper) A New Theoretical Framework for Information retrieval. In *SIGIR*, 194–200. 1986.

[39] P. Vorderer. Entertainment theory. *Communication and Emotion: Essays in Honor of Dolf Zillmann*, 131–153, 2003.

[40] R. W. White, J. M. Jose, and I. Ruthven. An Implicit Feedback Approach for Interactive Information Retrieval. *Information processing & management*, 42(1):166–190, 2006.

[41] M. Wilson and D. Elsweiler. Casual-leisure Searching: the Exploratory Search Scenarios that Break our Current Models. In *HCIR*, 28–31, 2010.

[42] D. Zillmann. Mood Management: Using Entertainment to Full Advantage. *In L. Donohew, H.E. Sypher, & E.T. Higgins (Eds.), Communication, social cognition, and affect*, 147–171, 1988.