

Factors Affecting Users' Measure of Interest: A Study of the Effect of Task, Document Difficulty and Document Familiarity

Stephen Akuma

Department of Mathematics and Computer Science, Benue State University, Makurdi, Nigeria
E-mail: akumas@uni.coventry.ac.uk

Chrisina Jayne

School of Computing Science and Digital Media, Robert Gordon University, Aberdeen, UK
E-mail: chrisina.draganova@gmail.com

Received: 31 January 2019; Accepted: 17 February 2019; Published: 08 May 2019

Abstract—Data on the web is constantly growing which may affect users' ability to find relevant information within a reasonable time limit. Some of the factors previously studied that affect users searching behaviour are task difficulty and topic familiarity. In this paper, we consider a set of implicit feedback parameters to investigate how document difficulty and document familiarity affects users searching behaviour in a task-specific context. An experiment was conducted and data was collected from 77 undergraduate students of Computer science. Users' implicit features and explicit ratings of document difficulty and familiarity were captured and logged through a plugin in Firefox browser. Implicit feedback parameters were correlated with user ratings for document difficulty and familiarity. The result showed no correlation between implicit feedback parameters and the rating for document familiarity. There was, however, a negative correlation between user mouse activities and document difficulty ratings.

Also, the dataset of all the participants in the experiment was grouped according to task type and analysed. The result showed that their behaviour varies according to task type. Our findings provide more insight into studying the moderating factors that affect user searching behaviour.

Index Terms—Document difficulty, Document familiarity, implicit indicators, Information Retrieval, Task Type.

I. INTRODUCTION

The present challenge in the field of Information Retrieval (IR) is to improve the recommendation of relevant resources to users [1, 2]. The user query is essential in retrieving documents, but most users find it difficult to formulate queries that can adequately represent their intentions. Information systems should be designed to take into consideration users characteristics

towards obtaining their information need. To effectively do this, the personalisation of information is necessary. Personalisation is a way of solving the problem of being 'lost in hyperspace', caused by document overload on the web. Personalisation considers either users' subjective perspective or common behaviour exhibited by a group of users through the use of contextual tools [3]. When such tools are used to link users to their web experience, it is referred to as web personalisation [4]. Personalisation in information retrieval involves recommending retrieval results to users based on their behaviour, domain knowledge, context, task type and other moderating factors [5]. Systems that are designed to incorporate aspects of user behaviour helps to personalize search to users; understanding the factors that affect user behaviour and interest is important in choosing the best sets of features for personalisation of information retrieval [6]. Users' interest in a given document depends on how relevant it is to a given task [7]. Documents that are of interest to a user are said to be relevant to the user. Considering that users vary in their search behaviour [8], there is a need to investigate the behaviours that affect user search activities in order to personalise information retrieval.

Previous studies on the moderating factors that affect user behaviour focused on topic familiarity and task difficulty [9-11]. Studies [11] have shown that topic familiarity affects the way users seek information. Similar study specific to reading time shows that topic familiarity has an effect on user reading time [12, 13]. Other studies of IR moderating factors focused on the use of search words [14], search effectiveness and length of the query [15]. The research by Liu, Liu and Belkin [6] suggests that topic familiarity affects user behaviour. Task type and task difficulty are two factors that also affect user behaviour [16]. Although not all task types affect user behaviour, user task still remains an important contextual factor affecting user searching behaviour [10]. This research builds on existing research but it narrows down from topic familiarity and task difficulty to

document familiarity and document difficulty. In this study, we investigate the effect of document familiarity and document difficulty on user searching behaviour in a task-specific context. We also studied the effect of task type on user behaviour. Document difficulty is viewed as how easy a user understands a current document; while document familiarity entails previous visit or knowledge of a current or related document. The following research questions were considered:

1. Do document familiarity and document difficulty affect users' implicit feedback parameters?
2. Does task type affect the result of the relationship between document familiarity/difficulty and users' implicit feedback parameters?

This paper is organized in the following order: Section 2 reviews the literature in this area. Section 3 describes the user study. The analysis of the results is presented in Section 4. Section 5 is the discussion, while Section 6 states the conclusion which summarizes the research and the limitation of the work. It also highlights the future work.

II. RELATED WORK

The main aim of measuring users' characteristics in a search process is to evaluate the extent of user individual differences and how the knowledge of users' intentions, interest and context can improve recommendation [17]. In IR, contextual measures describe the state in which information seeking activities occurs. Although the present search engine has provided some support to users, like the ranking of documents according to a user query, spelling and query suggestion, the interpretation of the suggestions remains the duty of the user. Contextual retrieval approach combines different search technologies and user context to meet users' information needs [18]. Multi-purpose personalisation approaches have been developed over the years but this has been limited to laboratory use. Research [19, 20] has shown that such generic personalisation approach does not adequately capture users' interest in a real-life situation. To create an efficient and robust system that will personalise user interaction with the web, contextual approaches need to be employed to capture users' information seeking behaviour and incorporate it into the system. Although contextual retrieval is important, it has not been fully implemented because of the difficulty in developing the right instrument to capture user's knowledge, intelligence, cognitive style, personality, memory, document familiarity, task or domain of interest. Attempts have been made to use pre-task and post-task questionnaires to predict user satisfaction [21], but it is yet to effectively capture user context; for instance, users might be instructed to use a 6-point rating scale to state their familiarity with a given web document but it might not be possible to obtain information about how much understanding the user has to a current topic. Although Mao et al [22] claimed that with a clear search context,

external assessors can annotate usefulness, they, however, did not state how to adequately capture search context. Domain knowledge can be measured based on experience or previous study but such classification may not lead to a concise interpretation of the results of a study. However, careful task-based experimentation can be used in place of questionnaires to capture these contextual factors while users browse the web. A new measurement approach for capturing user interest was proposed by Jiang et al [23]. They called it ephemeral state relevance (ESR), which they defined as the amount of useful information a user obtain when he/she click a result page and assessed it immediately during an interactive session. They evaluated the approach and inferred that it effectively captures users' current state of mind and it measures the usefulness of the information a user obtains per result. The next two subsections discuss Task and Domain knowledge in relation to user behaviour.

A. Task

Research in information retrieval in the last decade has focused on user interest and how context can be used to improve users' search experience. When an information system makes use of user context, it improves the quality of task outcomes [17]. Tasks are activities that people make an attempt to accomplish to meet a particular goal. These have been viewed from a different perspective by researchers [24]. They influence humans' social and psychological search behaviour [25]. Information seeking task refers to the general problems of a user which can be met by searching any related resources like documents, people or information system.

Previous researchers have investigated the effect of task type on users' information seeking behaviour [26-28] and they inferred that the needs of information seekers keep changing as they progress from one task to another task. This is viewed under different streams; identifying an accurate task stream is difficult because the demarcation of boundaries is not clear enough. However, common streams have classified different tasks along features like information gathering against fact-finding [29] and specifically defined instructions and non-specifically defined instructions [11]. Studies have also shown that such differences in task streams affect users' implicit behaviour [6, 28]. Kellar et al [29] say that information gathering tasks are more complex among the task types and require longer completion time and page views. Another study on the effect of the task as a contextual factor on reading time conducted by White and Kelly [30] found that under a task-specific context, dwell time can be used to measure document usefulness and improve the performance of an implicit feedback system. Liu and Wu [31] also reported similar findings; they examined whether task type is a good contextual factor that can be used for document prediction. Their result shows that task type assists in inferring document usefulness at the initial task stage and also influences the total time spent while performing the task. Kelly and Belkin [32] studied 7 users' behaviour in a naturalistic setting for 14 weeks and they reported that reading time

is best used for measuring document usefulness in a task-specific context. Lee and Pang [11] investigated the effect of task and topic knowledge in evaluating websites. They grouped user tasks into 'specifically' and 'non-specifically' types and their result indicates that users who followed specifically defined instructions in conducting a search task showed stricter credibility evaluations than those who did not follow specific instructions - confirming previous research that task types affect user searching behaviour.

Researchers [11, 16, 25, 33] have also investigated the effect of task difficulty on user behaviour. Cole et al [34] used eye movement patterns to examine the relationship between user interaction behaviour and task difficulty. They inferred that their technique can accurately use task difficulty patterns to differentiate between tasks types. Kelly et al [35] reported that the more cognitively complex the task, the more the searcher activity in terms of dwell time, mouse clicks and queries. Similar contextual research was carried out by Järvelin and Ingwersen [24]. The study by Li and Belkin [16] investigated the effect of task difficulty on user behaviour (reading time and 'hits' per query). They asked users to perform 6 tasks and rate the tasks according to difficulty. They infer that reading time is a good measure of user behaviour but it cannot predict task difficulty. Lee and Pang [11] investigated the effect of task and topical knowledge in evaluating websites and they found that the level of difficulty of a subject to a user affects the user's credibility evaluation of a website. A study conducted by C. Liu, Zhang and Huang [33] on whether task difficulty and domain knowledge affect user query formulation was carried out. Their findings indicate that task difficulty affects the searcher's performance. This work is similar to the previous research on factors affecting user behaviour. It, however, presents a novel approach which examines multiple indicators that affect user behaviour.

B. Domain Knowledge

One of the limiting factors of actualizing the personalisation approach is the inconsistency in user behaviour [28, 36] and the diverse subject areas [16]. Previous studies have used features like query input, search techniques and dwell time to examine how domain knowledge affects searchers' behaviour. In some cases, only a few indicators (reading time and search efficacy) were examined in relation to topic difficulty, task type and topic familiarity [9]. The research by Kelly and Cool [9] investigated the effect of topic familiarity on user behaviour. They inferred that users search efficiency increases and their dwell time on documents decrease when they are familiar with the search topic. Lee and Pang [11] research used 160 university students to examine the effect of topic knowledge on user web behaviour. They found evidence linking topic knowledge and credibility perception/evaluation. Their result showed that users who have less or no subject knowledge tend not to be thorough in appraising webpages. Capra, Arguello and Zhang [37] studied how search task affects user behaviour in terms of pre-task perceptions and post-task

outcomes. They found that when searchers know the task outcome, information requirement and process involved, it affects only the searchers' pre-task perceptions, but not the behaviour of the searcher and search outcomes. Bhavnani [12, 13] conducted a study to examine novice and experts search behaviour. They recruited 5 healthcare experts and 5 shopping experts for the study. Their findings suggest that while novice users begin their search with the general search engines, domain experts go straight to websites that will provide them with their needed information. A similar study was conducted by Hembrooke et al [38] to investigate the effect of domain knowledge on how searchers enter and reformulate queries. Their result shows that domain experts entered the complex and longer queries than domain novices. Zhang et al [39] used engineering domain to study the relationship between user search success and domain knowledge. They gave graduate and undergraduate engineering students 200 engineering terms to state whether they are familiar with the terms. Their findings show that the domain experts rated that they are familiar with more terms than the novices. Unlike previous small-scale studies, a large scale log analysis of searchers' behaviour on four domains was carried out by White et al [40]. They developed a model that can predict domain experts based on how they search for information. Gómez Ferragud, Solaz-Portolés and Sanjosé [41] research also found a difference between expert and novice problem-solving ability. Liu, Liu and Belkin [6] affirm that users with different level of topic familiarity differ in their search behaviour. A recent comparative study of novice and expert search behaviour in relationship with topic familiarity was conducted by Liu, Zhang and Huang [33]. Their finding was, however, not consistent with previous studies. They didn't find a significant difference between domain expert and novice in terms of query similarity. This was further investigated in this research with a focus on document familiarity.

As briefly described above, dwell time, queries and search techniques have been used by previous research to investigate whether domain knowledge affects user behaviour. In this work, an experimental approach is used to unfold the effect of domain knowledge with a focus on document familiarity with user behaviour.

III. USER STUDY DESCRIPTION

Part of the data generated from this study has been analysed in a previous publication (Akuma et al. 2016). In the study, two task types were employed, and 77 undergraduate students of the Faculty of Engineering and Computing participated. The study was conducted for 45 minutes for each of the task group and in the study. The participants were all above 18 years of age and they had a high proficiency in the use of computers. They were allowed to perform the study in a naturalistic setting by using their laptops at any place of their choice or in a controlled environment in selected computer laboratories of the University. This intertwined approach (naturalistic and controlled) was to make up for the disadvantage of

collecting data from just one of the ways. Most of the participants, however, performed the study in a controlled setting in the laboratory. A brief instruction and tutorial about the tasks were given to the participants, after which they completed and signed a consent form. A zipped Mozilla Firefox Portable injected with JavaScript plugin to capture user behaviour and explicit ratings were installed on the users' computers. They were asked to use the browser in searching for answers to the given tasks. The users were prompted by the plugin to enter their unique User Id and the URL of each of the documents visited by the users were captured along with their unique User ID and other features like their dwell time, mouse activity, key activity and their explicit ratings. In order to engage the users to search as many documents as possible, they were asked to visit and read a minimum of 7 web documents relating to the given tasks. Some students, however, did not visit up to 7 documents. For every web document they visited, they were to read and explicitly rate it according to familiarity and difficulty. The ratings were done immediately after the user closes the current document. The users were then asked to write a 200 words report of the solution to the tasks.

A. Behavioural features

The following behavioural measures were captured as users visit web documents through the time they spent on the web and their mouse and key activity. These "low-key" indicators have been proven to capture users' behaviour on the web to a reasonable degree of accuracy [42]. We grouped the events captured into Interactive and site structural data.

Interactive Events Captured

- *Dwell Time (DT)*: The dwell time (active time on document) is the actual period at which a web document is in focus. It is the total time (in seconds) that a user spends while reading a web document in one session.
- *Mouse Distance (MD)*: The mouse distance is calculated by the cursor movements along the x and y-axes of the monitor screen for every 100ms. It is the Euclidean distance of the mouse as shown in (1):

$$MD = \left(\sum_{i=1}^n \sqrt{(x-x_i)^2 + (y-y_i)^2} \right) \quad (1)$$

Where x , x_i , y_i and y are the locations of the mouse cursor on the monitor screen, n is the total number of location points moved by the mouse.

- *Mouse Movement (MM)*: The mouse movement is calculated as the mouse hovers along the x and y-axes of the monitor screen. The count for the movement along the x and y-axes are incremented by the change of its current value at each movement.
- *Mouse Duration Count (MDC)*: This is the total number of 100ms intervals that occurred while the mouse moved on the monitor screen.
- *Mean Mouse Velocity (MMV)*: The mean mouse

velocity is the average speed covered by the movement of the mouse on the screen. The formula for computing the mean mouse speed is given in (2):

$$MMV = \left(\left(\sum_{i=1}^n \sqrt{(x-x_i)^2 + (y-y_i)^2} \right) / t \right) / MDC \quad (2)$$

Where x , x_i , y_i and y are the locations of the mouse cursor on the monitor screen, n is the total number of location points moved by the mouse; t is the time covered (100ms). MDC is as explained above.

- *Number of Mouse Clicks (NMC)*: This is the total amount of mouse clicks on the current web document. The count is incremented for each page every time the mouse is clicked by a user.
- *Amount of Scroll (AS)*: The vertical length of most web pages is longer than the monitor height. Users normally scroll the web page by either clicking or dragging the scrollbar. The count increments anytime the scrollbar is dragged or clicked.
- *Number of Keystrokes (NK)*: This is the total number of keystrokes on a web document. The count for each page is incremented when a user strikes a key.
- *Amount of Copy (AC)*: This is the number of times the text is copied to the clipboard from a web document. Anytime a text from a particular document is copied, the count for the document is incremented by one.
- *Document Familiarity*: This is the users' statement of familiarity with the current document. The rating is done on a two-point scale.
- *Document Difficulty*: This is the users' statement of whether the current document is difficult to understand. The rating is done on a two-point scale.

Site Structural Data

- *Time Stamp*: This is the exact time and date in GMT when a document is loaded (open timestamp) and when a document is closed (close timestamp).
- *Page Height*: This is the vertical length of the document measured in pixels.
- *IP Address (IP)*: This is the unique string of numbers that represents the location of each computer. It is the internet protocol address of the user machine.
- *URL*: This is an acronym for Uniform Resource Locator. Each web document has a unique web address. It is the HTTP address of each web document visited by a user.

B. Tasks Assigned

We used simulated work situations and also took into consideration the participants' background [43]. They were designed to engage the participants and enable them to easily relate to the tasks and find the topics interesting. We also used Li & Belkin's [16] classification scheme as modified by Liu et al. [6] to group our tasks components

into a single scheme. Table 1 shows the different work task components. The tasks domain for the study was in the area of Computer Science. Two search tasks were given to two groups of students as shown below:

Simulated Work Task Situation 1

GIG Software Development company employed you as a consultant to provide a solution to the Company's pressing problem of developing a customised software within a minimal time frame. Some professional software developers achieved this by using the Rational Unified Process while others used the waterfall model.

Indicative request 1

Which of the approach would you consider for a small project of a few lines of code (LOC) and what stage of the software lifecycle do you consider to be the most important? State the reason for your answer in your report.

Simulated Work Task Situation 2

Google is looking for young and ambitious students of Computer science for an internship to work with the Company's Service Management Department. Consider that you are shortlisted for an interview among 2000 applicants and you are asked to search the internet and find answers to questions related to Information Technology Infrastructure Library (ITIL):

Indicative request 2

- a) What are the five stages of the ITIL lifecycle?
- b) What are the five stages of the ITIL lifecycle?
- c) What are the differences between ITIL v1, v2 and v3 (2007)?
- d) What are the ITIL processes?
- e) What are ITIL functions?
- f) Who should use ITIL?
- g) When should ITIL be used?
- h) What are the differences between ITIL and ISO/IEC

Table 1. Task components by Li & Belkin (2008) and extended Liu et al (2013)

Component	Values	Operational Definitions/Rules
Product	Physical	A task which produces a physical product
	Intellectual	A task which produces new ideas or findings
	Decision (Solution)	A task which makes a decision or solves a problem
	Factual Information	A task locating facts, data, or other similar items in information systems
	Image	A task locating image(s) in information systems
	Mixed Product	A task of locating different types of items in information systems
Goal (Quality)	Specific goal	A task with a goal that is explicit and measurable
	Amorphous goal	A task with a goal that cannot be measurable
	Combined goal	A task with both concrete and amorphous goals
Objective Task complexity	High complexity	A work task involving, at least, five activities during engaging in the task; a search task involving searching at least three types of information sources
	Moderate	A work task involving three or four activities during engaging in the task; a search task involving searching two types of information sources
	Low complexity	A work task involving one or two activities during engaging in the task; a search task involving searching one type of information source
Level	Document	A task for which a document as a whole is judged
	Segment	A task for which a part or parts of a document are judged

C. Tasks Classification

The grouping of the search task given to the participants is shown in Table 2 and it follows Liu, Liu and Belkin [6] task classification.

Task 1 is considered a mixed product (Decision and Intellectual task) because it involves making a decision to solve a problem with the most efficient method (RUP or Waterfall Model). It also focuses on 'how' a problem can be solved. It asked for the most important stage of the lifecycle, making it also an intellectual task. The goal of

the task is specific because participants have to find which approach is better for a few lines of code and it is of high complexity because a minimum of 7 documents was to be sourced.

Task 2 is considered a Factual product because facts are to be located and it focuses on gathering information about a thing or subject; it is specific because participants were to find specific information which was explicitly measurable. The complexity of this task is high because a minimum of 7 web documents was to be consulted.

Table 2. Component grouping of the two tasks

Task	Product	Goal (Quality)	Objective complexity	Level
Task 1	Mixed	Specific goal	High	Document
Task 2	Factual	Specific goal	High	Document

IV. RESULTS

The result of the data analysis is presented in this section. Among the 77 students that participated in the survey, 343 web documents were visited. Although users were asked to visit and read a minimum of 7 documents, some of them visited and rated less than 7 documents. Documents whose active time was above 600 seconds (10 minutes) were reduced and fixed at 600seconds. This was done on the assumption that in a 45 minutes' task, users will spend maximum time of 10 minutes on a document if they are not distracted by something else. A manual random check of the documents also showed that the dwell time on most of them is less than 10 minutes. Only documents rated by users according to difficulty and familiarity were analysed.

Statistical Concepts

This section describes the statistical concepts used for analysing the results of the experiments. The concepts explained include the Pearson correlation, Independent T-test, and significance test. The statistical rules for conducting each of the testings were followed.

Pearson Correlation

The Pearson correlation (denoted as r) generates a coefficient which measures the direction and the strength of the linear relationship between two variables. The value ranges from -1 to +1 depending on the strength of the relationship. A positive coefficient means that there is a positive linear correlation and a direct relationship between two parameters. A negative coefficient means that there is a negative linear correlation and an inverse relationship between two parameters. When the coefficient is closer to 1 or -1, the linear relationship is stronger. When the value of the coefficient is zero (0), it means that there is no correlation. The hypothesis is given as:

$H_0: \rho = 0$; the sample coefficient is equal to zero

Else

$H_A: \rho \neq 0$; the sample correlation is not equal to zero

Independent T-Test

The independent T-Test statistically determines whether there is a significant difference between the means of two independent groups. The independent T-test is carried out on large data set regardless of the test for normality since Independent T-Test and ANOVA are not very sensitive to a large distribution of data that slightly deviate from normality [44-46].

A. The effect of document familiarity on user implicit behaviour

Only documents rated for familiarity were analysed. Among the 324 rated by the participants, 203 were rated as not familiar while 121 were rated as familiar. Pearson correlation was run on the dataset to examine if there is a correlation between the implicit indicators and the user ratings for familiarity. The result, as shown in Table 3, indicates that there is no statistically significant correlation between the implicit indicators and the user ratings for familiarity. The non-correlation between the implicit indicators and familiarity might be due to the time given for the experiment; users read the documents regardless of whether they were familiar or not while performing the tasks.

The Independent T-Test conducted also did not reject the null hypothesis. Table 3 also shows that although the mean ratings for familiar and non-familiar documents appeared different, it was not statistically significant. There is no significant difference between the two groups of familiarity in terms of the implicit features. It suggests that users' familiarity with a document does not have a significant effect on their searching behaviour.

B. The effect of document difficulty on user behaviour

Unlike the test for familiarity where there was no correlation between the implicit indicators and document familiarity, in the case of document difficulty, there is a correlation between the mouse activities and the rating for document difficulty on the 317 documents rated by the users. The Mouse movement along the X and Y axes and the mouse distance correlated negatively with the ratings for document difficulty with the correlation of coefficient as -0.111, -0.115 and -0.119 respectively. This indicates that the greater the difficulty of the document, the less movement of mouse activity. Table 4 shows the Pearson correlation between the implicit indicators and difficulty ratings.

Independent T-Test was used to compare the means of the implicit indicators for documents rated as difficult to understand and those rated as non-difficult to understand. Although the results showed a difference in mean for all the indicators measured, the mean difference for the copy indicator, mouse movement along X and Y axes of the screen, the mouse distance, the mouse duration count and the mean mouse speed were, however, significant ($p < 0.05$) with their p -values as 0.008, 0.004, 0.003, 0.002 and 0.043 respectively. Users copied and performed more mouse activities on documents they considered not difficult than on documents they considered difficult to understand, as highlighted in Table 4. This indicates that when users find a document that is difficult to understand, they perform fewer activities on the document and move to documents they can easily comprehend.

Table 3. Pearson correlation between the implicit indicators and familiarity ratings. Comparison of implicit indicators based on document familiarity groups

Implicit Indicators	Pearson Correlation(r) with Document Familiarity rating	Significant coefficient level (p)	Mean Not Familiar (N=203)	Mean Familiar (N=121)	T-TEST (p)
Clicks	0.008	0.882	1.70	1.64	0.882
Height	0.085	0.128	4194.36	4742.67	0.128
Copy	-0.009	0.878	0.53	0.51	0.878
Scroll	0.079	0.154	173.29	217.02	0.196
Mouse movement X	0.027	0.634	3882.77	4175.37	0.634
Mouse movement Y	0.031	0.583	4787.45	5118.16	0.583
Dwell time	0.98	0.08	124.665	158.901	0.08
Mouse distance	0.035	0.534	7381.16	7994.09	0.534
Mouse duration count	0.072	0.196	114.51	139.63	0.196
Mean mouse speed	0.001	0.980	870.06	871.67	0.980
Keystroke	0.057	0.305	1.62	2.86	0.305

Table 4. Pearson correlation between the implicit indicators and difficulty ratings. Comparison of implicit indicators based on document difficulty groups

Implicit Indicators	Pearson Correlation(r) with Document difficulty rating	Significant coefficient level (p)	Mean Not Difficult (N=267)	Mean Difficult (N=50)	T-TEST (p)
Clicks	-0.057	0.310	1.79	1.24	0.31
Height	0.111	0.49	4251.83	5203.32	0.116
Copy	-0.096	0.088	0.57	0.28	0.008
Scroll	-0.052	0.359	198.03	159.86	0.359
Mouse movement X	-0.111	0.049	4235.16	2621.54	0.004
Mouse movement Y	-0.115	0.04	5191.97	3527.16	0.003
Dwell time	-0.079	0.158	144.610	107.320	0.158
Mouse distance	-0.119	0.034	8063.30	5258.18	0.002
Mouse duration count	-0.072	0.202	129.10	95.76	0.202
Mean mouse speed	-0.086	0.125	892.30	761.90	0.043
Keystroke	-0.007	0.897	2.15	1.94	0.897

C. The Effect of Task-Type on User Behaviour

This section investigates the effect of task type on user searching behaviour. Two task types with a high number of participants were used as case studies in evaluating task effect on user behaviour. The analysis of the tasks

(Task 1 and Task 2) performed by the participants is presented below:

Mixed Task VS Factual Task: *Task Specific grouping of the relationship between implicit indicators and document familiarity ratings*

For the Mixed task, among the 197 documents rated for familiarity, 106 were rated not familiar and 91 were rated familiar. There was no significant correlation between the rating for document familiarity and user-generated implicit indicators. The Independent T-Test also did not produce any significant difference in mean between documents rated familiar and those rated as non-familiar with respect to the implicit indicators.

For the Factual task, among the 92 documents, 63 were rated as not familiar while 29 were rated as familiar. Like the mixed task, no significant correlation between the implicit indicators and the ratings for familiarity was found. However, the T-test as shown in Table 5 produced a significant difference in mean for the copy parameter between documents rated familiar to those rated as unfamiliar. It showed that texts were copied from documents considered to be unfamiliar than those considered to be familiar.

Mixed Task VS Factual Task: *Task Specific grouping of the relationship between implicit indicators and document difficulty ratings*

For the Mixed task, among the 192 documents rated, 34 were rated as difficult to understand and 158 were rated as not difficult to understand. There was no significant correlation between the Implicit indicators and ratings for document difficulty, but the Independent T-test shows significant difference between the two groups

of difficulty for the following features: Mouse movement along Y-axis ($p = 0.010$), total distance ($p = 0.009$) and Mouse movement along X-axis ($p = 0.021$), as can be seen in Table 6. The difference in mean, as highlighted in the two groups in Table 6, indicates that users of Task 1 (Mixed task) moved the mouse more on documents they considered not difficult to understand than documents they considered to be difficult.

In terms of the Factual task, 92 documents were rated

by the users. Among these, 12 were rated as difficult to understand and 77 were rated not difficult to understand. There was no correlation between the implicit indicators and the ratings for difficulty. Only the copy parameter was significantly different for the difficult and non-difficult groups as shown in Table 6. This suggests that in a Factual task, users' copied text on documents that were rated not difficult more than documents that were rated difficult to understand.

Table 5. Task Specific grouping of the relationship between implicit indicators and familiarity ratings

Implicit Indicators	Mixed task				Factual Task			
	Mean Not Familiar (N=106)	Mean Familiar (N=91)	Pearson Correlation(r)	T-TEST (p)	Mean Not Familiar (N=63)	Mean Familiar (N=29)	Pearson Correlation(r)	T-TEST (p)
Clicks	1.75	1.44	-0.46, $p = 0.519$	0.519	2.17	1.52	-0.074, $p = 0.483$	0.483
Height	4509.84	5062.33	0.094, $p = 0.187$	0.187	3627.48	3910.55	0.040, $p = 0.708$	0.708
Copy	0.42	0.53	0.074, $p = 0.302$	0.302	0.86	0.28	-0.160, $p = 0.128$	0.039
Scroll	183.76	213.74	0.057, $p = 0.427$	0.427	157.63	126.07	-0.075, $p = 0.477$	0.477
Mouse movement X	3934.44	4343.33	0.034, $p = 0.633$	0.633	4799.49	3644.55	-0.114, $p = 0.279$	0.279
Mouse movement Y	5117.63	5243.71	0.011, $p = 0.874$	0.874	4775.65	3840.79	-0.098, $p = 0.353$	0.353
Dwell time	116.274	151.242	0.104, $p = 0.145$	0.145	176.143	144.103	-0.081, $p = 0.444$	0.444
Mouse distance	7761.75	8276.45	0.028, $p = 0.698$	0.698	8043.49	6221.14	-0.113, $p = 0.282$	0.282
Mouse duration count	123.91	143.78	0.053, $p = 0.458$	0.458	137.40	106.76	-0.099, $p = 0.347$	0.347
Mean mouse speed	825.88	857.53	0.029, $p = 0.690$	0.690	863.86	990.76	0.097, $p = 0.358$	0.358
Keystroke	1.39	3.55	0.088, $p = 0.218$	0.236	1.19	3.90	0.128, $p = 0.226$	0.404

Table 6. Task Specific grouping of the relationship between implicit indicators and difficulty ratings

Implicit Indicators	Mixed task				Factual Task			
	Mean Not Difficult (N=158)	Mean Difficult (N=34)	Pearson Correlation(r)	T-TEST (p)	Mean Not Difficult (N=77)	Mean Difficult (N=12)	Pearson Correlation(r)	T-TEST (p)
Clicks	1.72	1.21	-0.059, $p = 0.415$	0.415	2.0	1.50	-0.047, $p = 0.660$	0.660
Height	4697.97	5289.65	0.077, $p = 0.288$	0.368	3632.45	4464.83	0.085, $p = 0.429$	0.429
Copy	0.50	0.32	-0.089, $p = 0.219$	0.219	0.77	0.08	-0.137, $p = 0.202$	0.003
Scroll	213.99	137.91	-0.109, $p = 0.131$	0.131	153.04	126.08	-0.047, $p = 0.665$	0.665
Mouse movement X	4401.06	2678.26	-0.111, $p = 0.127$	0.021	4605.95	2479.75	-0.159, $p = 0.136$	0.136
Mouse movement Y	5523.84	3593.12	-0.132, $p = 0.068$	0.010	4727.90	2335.08	-0.184, $p = 0.084$	0.084
Dwell time	142.051	94.147	-0.108, $p = 0.135$	0.135	178.052	99.833	-0.143, $p = 0.181$	0.181
Mouse distance	8546.61	5401.82	-0.130, $p = 0.073$	0.009	7837.13	3991.25	-0.178, $p = 0.095$	0.095
Mouse duration count	139.13	101.74	-0.076, $p = 0.292$	0.292	130.19	83.75	-0.113, $p = 0.294$	0.294
Mean mouse speed	864.03	729.09	-0.093, $p = 0.199$	0.072	954.87	655.83	-0.166, $p = 0.120$	0.120
Keystroke	2.89	0.35	-0.078, $p = 0.281$	0.281	2.42	0.08	-0.080, $p = 0.459$	0.459

V. DISCUSSION

The research investigated the effect of document difficulty and document familiarity on user behaviour, and the effect of task type on user behaviour. Two task types were given to a different set of students of the same grade. The mixed tasks focus on 'how' a problem can be solved while the factual tasks focus on finding facts. The

findings in the result sections show that there is no significant difference in user behaviour in terms of documents they consider to be familiar with those they consider as unfamiliar. The result show a significant difference between documents considered as difficult to understand and those considered not difficult to understand in terms of mouse activities and the amount of copy. Among the behavioural features (Mouse Duration Count, Distance of Mouse Movement, Total Mouse

Movement, Mean Mouse Velocity, Number of Mouse Clicks, Amount of Scroll, Number of Keystrokes, Amount of Copy and Active Time Spent on the Document) examined, the results suggest that users perform fewer mouse activities and copy less text from documents they consider difficult to understand. The results also show that there is no significant difference in user implicit behaviour in terms of documents they consider to be familiar from those they consider as unfamiliar. We found a significant difference between documents considered difficult and those considered not difficult to understand in terms of user search behaviour. The results of this research are consistent with previous research on the effect of topic difficult on user behaviour [11, 33]. Although this study focuses on the effect of document difficulty and document familiarity on user behaviour, it produces a similar result with previous studies that examined the effect of task difficulty and topic familiarity on user behaviour.

The analysis of the different task types shows that in the mixed task, users moved the mouse more and spent more time on documents considered as not difficult to understand. Also, users of the factual task moved the mouse quicker on pages considered to be easy than pages considered to be difficult. This indicates that users probably abandoned documents they consider difficult to understand, and moved swiftly to documents they understood better. It may also be that users were mindful of the time given for the experiment and they paid more attention to documents that were not difficult to understand. The result of this research also suggests that users who are looking for facts and specific information moved the mouse quicker on 'easy to understand' documents. Our result demonstrates that there is a relationship between user search behaviour and user level of perceived understanding of a document. When reading documents that are easy to understand, users will likely move the mouse more on the documents. When they encounter a perceived difficult document, there is a tendency that they may leave the document out of frustration and move to the next available documents. Since the researchers found that task type has an influence on user behaviour, additional task components and type will be investigated further and a comparative analysis between the result of this research and previous results will be evaluated.

VI. CONCLUSION

This paper has presented research on the effect of document familiarity and difficulty of user's search behaviour. Our study was both controlled and naturalistic with 77 participants. Two different task components were used to conduct the study. The contribution of this paper within the context explored is as follows: 1) It showed that there is no significant difference between a familiar document and a non-familiar document in terms of user search behaviour. 2) It showed that document difficulty affects users' behaviour in the aspect of mouse activity. 3) Also, the research shows that user understanding of a

document differs according to task type. This work is important in determining the moderating factors to be considered when designing a personalised learning system based on implicit features.

There were limitations to this study. The user study was conducted in a Computer Science domain, so the dataset and the result obtained from it may not be easily generalizable to other domains. An investigation needs to be carried out for other domains. The number of participants and the documents obtained from the experiment was not large enough. Although a lot of task types have been listed, only two task types were employed for this research.

In the future, more data will be collected and more task components will be explored to determine the significant behavioural features that can be used to distinguish documents based on difficulty and familiarity. Also, predictive models of document familiarity and difficulty will be developed based on user behaviour.

REFERENCES

- [1] G. Jawaheer, P. Weller and P. Kostkova, "Modeling User Preferences in Recommender Systems: A Classification Framework for Explicit and Implicit User Feedback," *ACM Transactions on Interactive Intelligent Systems*, vol. 4, pp. 1-26.
- [2] X. Zhu, J. Huang, B. Zhou, A. Li and Y. Jia, "Real-time personalized twitter search based on semantic expansion and quality model," *Neurocomputing*, vol. 254, pp. 1339-1351.
- [3] S. Akuma, R. Iqbal, C. Jayne and F. Doctor, "Comparative analysis of relevance feedback methods based on two user studies," *Comput.Hum.Behav.*, vol. 60, 7, pp. 138-146.
- [4] B. Mobasher, R. Cooley and J. Srivastava, "Automatic Personalization Based on Web Usage Mining," *Commun ACM*, vol. 43, no. 8, pp. 142-151.
- [5] A. Alhindi, U. Kruschwitz, C. Fox and M. Albakour, "Profile-Based Summarisation for Web Site Navigation," *ACM Transactions on Information Systems*, vol. 33, no. 1, pp. 1-40.
- [6] J. Liu, C. Liu and N. Belkin, "Examining the effects of task topic familiarity on searchers' behaviors in different task types," *Proceedings of the ASIST Annual Meeting*, vol. 50, no. 1.
- [7] N.J. Belkin, D. Hienert, P. Mayr and C. Shah, "Data requirements for evaluation of personalization of information retrieval - A position paper," *CEUR Workshop Proc.*, vol. 1866.
- [8] S. Akuma, C. Jayne, R. Iqbal and F. Doctor, "Implicit predictive indicators: Mouse activity and dwell time," *IFIP Advances in Information and Communication Technology*, vol. 436, September 19-21, pp. 162-171.
- [9] D. Kelly and C. Cool, "The effects of topic familiarity on information search behavior," *Proceedings of the ACM International Conference on Digital Libraries*, pp. 74-75.
- [10] J. Liu, C. Liu, J. Gwizdka and N.J. Belkin, "Can search systems detect users' task difficulty? Some behavioral signals," *SIGIR 2010 Proceedings - 33rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 845-846.
- [11] H. Lee and N. Pang, "Understanding the effects of task and topical knowledge in the evaluation of websites as information patch," *J.Doc.*, vol. 74, no. 1, pp. 162-186.

- [12] S.K. Bhavnani, "Domain-specific search strategies for the effective retrieval of healthcare and shopping information," *Conference on Human Factors in Computing Systems - Proceedings*, pp. 610-611.
- [13] S.K. Bhavnani, "Important cognitive components of domain-specific search knowledge," pp. 571-578.
- [14] B. Allen, "Topic knowledge and online catalog search formulation," *Library Quarterly*, vol. 61, no. 2, pp. 188-213.
- [15] X. Zhang, H.G.B. Angheliescu and X. Yuan, "Domain knowledge, search behaviour, and search effectiveness of engineering and science students: An exploratory study," *Information Research*, vol. 10, no. 2.
- [16] Y. Li and N.J. Belkin, "A faceted approach to conceptualizing tasks in information seeking," *Information Processing and Management*, vol. 44, no. 6, pp. 1822-1837.
- [17] P. Bennett, K. Collins-Thompson, D. Kelly, R. White and Y. Zhang, "Overview of the Special Issue on Contextual Search and Recommendation," *ACM Transactions on Information Systems*, vol. 33, no. 1, pp. 1-7.
- [18] D.K. Limbu, A.M. Connor, R. Pears and S.G. MacDonell, "Improving web search using contextual retrieval," *ITNG 2009 - 6th International Conference on Information Technology: New Generations*, pp. 1329-1334.
- [19] M. Busby, "Learn Google," 2003.
- [20] R. Iqbal, A. Grzywaczewski, J. Halloran, F. Doctor and K. Iqbal, "Design implications for task-specific search utilities for retrieval and reengineering of code," *Enterprise Information Systems*, pp. 1751-1757.
- [21] A. Crescenzi, R. Capra and J. Arguello, "Time pressure, user satisfaction and task difficulty," *Proceedings of the ASIST Annual Meeting*, vol. 50, no. 1.
- [22] J. Mao, Y. Liu, K. Zhou, J. Nie, M. Zhang, S. Ma, J. Sun and H. Luo, "When does Relevance Mean Usefulness and User Satisfaction in Web Search?" *SIGIR '16 Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval*, July 17 -21, 2016, pp. 463-472.
- [23] J. Jiang, D. He, D. Kelly and J. Allan, "Understanding ephemeral state of relevance," March 07-11, 2017, pp. 137-146.
- [24] K. Järvelin and P. Ingwersen, "Information seeking research needs extension toward tasks and technology," *Information Research*, vol. 10, no. 1.
- [25] R.A. Hamid, J.A. Thom and D.N.F.A. Iskandar, "Effects of relevance criteria and subjective factors on web image searching behaviour," *J.Inf.Sci.*, vol. 43, no. 6, pp. 786-800.
- [26] C.C. Kuhlthau, "A principle of uncertainty for information seeking," *Journal of Documentation*, vol. 49, no. 4, pp. 339-355.
- [27] T.D. Wilson, "Models in information behaviour research," *Journal of Documentation*, vol. 55, no. 3, pp. 249-270.
- [28] R. Mehrotra and E. Yilmaz, "Terms, topics & tasks: Enhanced user modelling for better personalization," *ICTIR - Proc. ACM SIGIR Int. Conf. Theory Inf. Retr.*, pp. 131-140.
- [29] M. Kellar, C. Watters and M. Shepherd, "A field study characterizing Web-based information-seeking tasks," *J.Am.Soc.Inf.Sci.Technol.*, vol. 58, no. 7, pp. 999-1018.
- [30] R.W. White and D. Kelly, "A study on the effects of personalization and task information on implicit feedback performance," *International Conference on Information and Knowledge Management, Proceedings*, pp. 297-306.
- [31] D.-. Liu and I.-. Wu, "Collaborative relevance assessment for task-based knowledge support," *Decis.Support Syst.*, vol. 44, no. 2, pp. 524-543.
- [32] D. Kelly and N.J. Belkin, "Display time as implicit feedback: Understanding task effects," *Proceedings of Sheffield SIGIR - Twenty-Seventh Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 377-384.
- [33] C. Liu, X. Zhang and W. Huang, "The exploration of objective task difficulty and domain knowledge effects on users' query formulation," *Proceedings of the Association for Information Science and Technology*, vol. 53, no. 1, pp. 1-9.
- [34] M.J. Cole, X. Zhang, C. Liu, N.J. Belkin and J. Gwizdka, "Knowledge effects on document selection in search results pages," *SIGIR'11 - Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 1219-1220.
- [35] D. Kelly, J. Arguello, A. Edwards and W. Wu, "Development and Evaluation of Search Tasks for IIR Experiments using a Cognitive Complexity Framework," September 27-30, 2015.
- [36] R. Iqbal, A. Grzywaczewski, A. James, F. Doctor and J. Halloran, "Investigating the value of retention actions as a source of relevance information in the software development environment," *Proceedings of the 2012 IEEE 16th International Conference on Computer Supported Cooperative Work in Design, CSCWD 2012*, pp. 121-127.
- [37] R. Capra, J. Arguello and Y. Zhang, "The Effects of Search Task Determinability on Search Behavior," April, 2017, pp. 108-121.
- [38] H.A. Hembrooke, L.A. Granka, G.K. Gay and E.D. Liddy, "The effects of expertise and feedback on search term selection and subsequent learning," *J.Am.Soc.Inf.Sci.Technol.*, vol. 56, no. 8, pp. 861-871.
- [39] X. Zhang, H.G.B. Angheliescu and X. Yuan, "Domain knowledge, search behaviour, and search effectiveness of engineering and science students: An exploratory study," *Information Research*, vol. 10, no. 2.
- [40] R.W. White, S.T. Dumais and J. Teevan, "Characterizing the influence of domain expertise on web search behavior," *Proceedings of the 2nd ACM International Conference on Web Search and Data Mining, WSDM'09*, pp. 132-141.
- [41] C.B. Gómez Ferragud, J.J. Solaz-Portolés and V. Sanjosé, "Effects of topic familiarity on analogical transfer in problem-solving: A think-aloud study of two singular cases," *Eurasia J.Math.Sci.Technol.Educ.*, vol. 11, no. 4, pp. 875-887.
- [42] G. Pasi, "Implicit feedback through user-system interactions for defining user models in personalized search," *Procedia Comput. Sci.*, vol. 39, no. C, pp. 8-11.
- [43] P. Borlund, "The IIR evaluation model: A framework for evaluation of interactive information retrieval systems," *Information Research*, vol. 8, no. 3.
- [44] G.V. Glass, P.D. Peckham and J.R. Sanders, "Consequences of failure to meet assumptions underlying fixed effects analyses of variance and covariance," *Rev. Educ. Res.*, vol. 42, pp. 237-288.
- [45] M.R. Harwell, E.N. Rubinstein, W.S. Hayes and C.C. Olds, "Summarizing Monte Carlo results in methodological research: the one- and two-factor fixed effects ANOVA cases," *J. Educ. Stat.*, vol. 17, pp. 315-339.
- [46] L.M. Lix, J.C. Keselman and H.J. Keselman, "Consequences of assumption violations revisited: A quantitative review of alternatives to the one-way analysis of variance F test," *Rev. Educ. Res.*, vol. 66, pp. 579-619.

Authors' Profiles



Dr. Stephen Akuma is currently a Lecturer of Computer Science in the Department of Mathematics/Computer Science at Benue State University. He holds a PhD in Computing and a Masters degree in Software Development (Distinction) from Coventry University, United Kingdom. Stephen also holds a

Bachelor's degree in Computer Science (2.1) from Benue State University. His research area is informational retrieval and personalisation. He has 10 years of Teaching Experience and has published papers in International Conferences and Journals.



Prof. Chrisina Jayne received her Ph.D. degree in applied mathematics from Sofia University, Sofia, Bulgaria in 1998. She is currently Head of the School of Computing and Digital Media at Robert Gordon University, Scotland. Her research includes developing and applying neural network methods to a number of

applications. Professor Jayne is a Fellow of the British Computer Society and UK Higher Education Academy. She is a coordinator of the International Neural Network Society's Special Interest Group on Engineering Applications of Neural Networks. She chaired the Engineering Applications of Neural Network Conference in 2012 and 2014 and she has been involved in numerous program committees of neural network related conferences.

How to cite this paper: Stephen Akuma, Chrisina Jayne, "Factors Affecting Users' Measure of Interest: A Study of the Effect of Task, Document Difficulty and Document Familiarity", *International Journal of Information Technology and Computer Science(IJITCS)*, Vol.11, No.5, pp.47-57, 2019. DOI: 10.5815/ijitcs.2019.05.06