Human–Computer Interaction: The Impact of Users' Cognitive Styles on Query Reformulation Behaviour During Web Searching

Khamsum Kinley, Dian Tjondronegoro, Helen Partridge and Sylvia Edwards
Information Systems School, Science and Engineering Faculty
Queensland University of Technology
2 George St Brisbane QLD 4000 Australia
{k.kinley, dian, h.partridge, s.edwards}@qut.edu.au

ABSTRACT
This paper discusses users' query reformulation behaviour while searching information on the Web. Query reformulations have emerged as an important component of Web search behaviour and human-computer interaction (HCI) because a user's success of information retrieval (IR) depends on how he or she formulates queries. There are various factors, such as cognitive styles, that influence users' query reformulation behaviour. Understanding how users with different cognitive styles formulate their queries while performing Web searches can help HCI researchers and information systems (IS) developers to provide assistance to the users. This paper aims to examine the effects of users' cognitive styles on their query reformation behaviour. To achieve the goal of the study, a user study was conducted in which a total of 3613 search terms and 872 search queries were submitted by 50 users who engaged in 150 scenario-based search tasks. Riding's (1991) Cognitive Style Analysis (CSA) test was used to assess users' cognitive style as *wholist* or *analytic*, and *verbaliser* or *imager*. The study findings show that users' query reformulation behaviour is affected by their cognitive styles. The results reveal that analytic users tended to prefer *Add* queries while all other users preferred *New* queries. A significant difference was found among wholists and analytics in the manner they performed *Remove* query reformulations. Future HCI researchers and IS developers can utilize the study results to develop interactive and user-cantered search model, and to provide context-based query suggestions for users.

Author Keywords
Human–Computer Interaction, Query Reformulation, Information Retrieval, Web Search Behaviour, Cognitive style

ACM Classification Keywords
H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION

OZCHI'12, November 26–30, 2012, Melbourne, Victoria, Australia.

Query reformulation has emerged as an important component of Web search behaviour and human-computer interaction (HCI) because a user's success of information retrieval (IR) depends on how he or she formulates queries. It also reflects the user's attempt to locate relevant information on the Web. Users constantly reformulate their queries to locate relevant information on the Web. Understanding how users generate their search queries and what factors influence their query reformulations can help towards developing adaptive Web search interface which in turn can help users find relevant information on the Web.

Several studies have focused on query reformulations to assist users to find required information (Fidel & Pejtersen, 2004; Liu, Gwizdka, Liu, Xu, & Belkin, 2010; Rieh & Xie, 2006; Tseng, Tjondronegoro, & Spink, 2009; Yoon & Chung, 2009). These studies show that effective query reformulations can help in locating relevant information. Studies also show that various factors, such as task complexity and users' cognitive styles, influence Web search behaviour in general (Gorrell, Eaglestone, Ford, Holdridge, & Madden, 2009; Gwizdka, 2009; Kim, 2000; Kinley & Tjondronegoro, 2010a, 2010b). Therefore, it is worth investigating how users' cognitive styles influence their query formulation behaviour.

This paper aims to investigate (1) the associations within query reformulation types, and (2) the effects of users' cognitive styles on their query reformulation behaviour. Investigation into users' cognitive styles and their query reformulation behaviour can HCI researchers and information systems (IS) developers to develop an interactive and user-cantered search model and search systems to provide effective query suggestions for users. Such functionalities and services can help the users to retrieve relevant search results.

RELATED STUDIES

Query Reformulations

Query reformulation is a process of modifying a given query in order to improve search results. Measuring the number of query reformulations can provide valuable information about the interactions between a user and the Web search engine. Jansen, Zhang and Spink (2007) reported that changing the query topic was the primary means to modify queries. Users submit queries to search engines to conduct their first search and frequently make changes to their queries in order to improve the result of

the search. Depending on the system feedback and searching results displayed, users reformulate their queries employing their prior knowledge and experience until the required information is retrieved.

Sutcliffe and Ennis (1998) proposed a process model of information searching activities and knowledge sources in which query formulation and reformulation were identified as one of the core components of the model. They proposed query formulation/reformulation as one of the main activities performed in an IR process. They found that the complexity of query formulation depends on the complexity of the IR system and the user's skill in generating queries. Complex queries can be formed if the user is skilled in query languages, such as Boolean and structured query language, where the information need is transformed into keywords and query syntax, which are employed by IR systems.

Jansen, Zhang & Spink (2007) found that Web searchers preferred to decrease the number of search results by means of three methods: AAN (noun after term), AAP (phrase after term), and CT (change to related terms). AAN and AAP relate to set restrictions to the previous queries in order to deduce the number of search results. For CT, Web user tends to modify the terms to a more specific one. Huang and Efthimiadis (2009) analysed query reformulations in the America Online (AOL) query logs. They reported that different query reformulation strategies had distinct characteristics and found certain query modifications, such as add/remove words, word substitution, and spelling correction most effective. They believe that users performed a query reformulation because they are dissatisfied with the previous results.

In a study of a large records of Dogpile.com log analysis, Jansen, Booth & Spink (2009) reported that 22.73 % of the search queries were query reformulations in which searchers modify their previous queries. They also reported that Web searchers move to narrow their query at the start of a session and then gradually move to reformulation in the mid and later portions of the session.

Categories of Query Reformulations

While searching information on the Web, users tend to use various types of query reformulations. Different information behaviour researchers tend to classify query reformulations differently. Rieh and Xie (2006) characterized Web query reformulation as having three facets: content, format and resource. *Content* modifications refer to those instances in which users made changes to the meaning of a query. *Format* modifications include the cases in which users made changes without altering the meaning of the query by means of using operators, etc. *Resource* modifications refer those instances in which users intended to make changes in types of information resources (e.g., news, images, and music files). They reported that content modifications accounted for 80.3% of the query reformulations, while 14.4% of the query reformulations were format modifications and 2.8% of the modifications were related to resource.

Recent researchers have used automatic detection methods to classify queries into several taxonomies of query reformulation classes. Jansen et al. (2009) classified query formulations into six groups while investigating Dogpile.com search logs:

New: The query is the first query from a unique User Identification–Cookie, or the query is on a new topic from this searcher

Assistance: This query is generated by the searcher's selection of an 'Are You Looking For?' feature.

Content Change: The user executed a query on another content collection. The available content collections were Web, Images, Audio, News, and Video.

Generalization: The current query is on the same topic as the searcher's previous query, but the searcher is now seeking more general information.

Reformulation: The current query is on the same topic as the searcher's previous query, and both queries contain common terms.

Specialization: The current query is on the same topic as the searcher's previous query, but the searcher is now seeking more specific information. Jansen et al determined a query reformulation to be specialization if the query contained more terms than the previous query by a particular user.

Jansen et al. (2009) reported that 63% of the total queries submitted corresponded to users initiating a new search. This indicates that *New* reformulation dominated amongst other forms of modifications.

Factors Affecting Web Query Reformulation Behaviour

A number of studies have shown contextual factors, such as task types having effects on users' Web query reformulation behaviour. Liu and his team (2010) examined how task types and task situation influenced users' query reformulation behaviour. The search task types were *simple*, where the information need is satisfied by a single, independent piece of information; *hierarchical*, where finding multiple characteristics of a single concept satisfies the information need; and *parallel*, where the information need is satisfied by finding multiple concepts that exist at the same level in a conceptual hierarchy.

Liu, et al. (2010) found that *specification* query reformulation (in which the succeeding query contained addition search terms that the preceding query) was most frequently used in *simple* tasks. The *word substitution* was most frequently used in parallel tasks. The frequency of the *generalization* query reformulations was the least, while *new* query was more likely to be used. In *hierarchical* tasks, the reformulated queries were sometimes dependent upon previous search results and users needed to find more specific information in their latter queries. In contrast, in *parallel* tasks, sometimes they only needed to change a single concept and did not need to change other query terms.

Comparing the distribution of the four modifications, i.e., generalization, specification, dynamic and constant, in multimedia searches, Tseng, Tjondronegoro, & Spink (2009) found that audio search users formulated more new search topics than image and video searches. Image and video users were found to have the same amount of modifications, while audio users showed slightly fewer modifications per session.

Cognitive Styles

Different authors refer to cognitive style with different terms, such as field-dependent/independent (Witkin, Moore, Goodenough, & Cox, 1977), holists-serialist (Pask, 1976), and wholist-analytic/verbal-imagery (Riding & Cheema, 1991). Riding and Cheema (1991) grouped the cognitive dimensions into two principal cognitive dimensions: the wholist-analytic and the verbal-imagery style dimensions.

The *wholist-analytic* (WA) dimension of cognitive style describes the habitual way in which people think about, view and structure information in wholes or parts. This affects the way they learn and organize information. The *verbal-imagery* (VI) dimension of cognitive style describes an individual's tendency to process information either in verbal or verbal mode of representation and thinking. It refers to ways in which an individual would represent knowledge and information in either words (verbal) or mental pictures (images).

Many researchers have developed their own instruments and tools to assess cognitive styles. Riding (1991) developed *Cognitive Styles Analysis* (CSA) test to measure both WA and VI dimensions of cognitive styles (Riding & Cheema, 1991). CSA is a computer presented test that classifies users as wholists or analytics on the WA dimension, and verbalisers or imagers on the VI dimension. This study utilized Riding's (1991) CSA test to assess participants' cognitive styles.

Several studies have shown users cognitive styles having effects on their Web search behaviour (example: Ford, Eaglestone, Madden, & Whittle, 2009; Kim, 2000). In a study investigating the impact of users' cognitive styles on Web search and navigation, Kim (2000) reported that the field-dependent users with little or no online experience had difficulties in retrieving information on the Web. The field-independent users tended to outperform the field-dependent users in the sense that they spent lesser time and visited a fewer nodes than the field-dependent users to complete a search task.

From a user study aimed at exploring the inter-relationships between Web users' searching behaviour and their cognitive style, Kinley, Tjondronegoro, & Partridge (2010) presented a conceptual model of Web searching and cognitive styles. The model based on the preliminary findings, illustrated relationships between different stages of Web searching and cognitive styles.

Several studies have attempted to investigate effects of users' cognitive styles on their Web search strategies and navigational styles. However, very limited studies have examined the effects of users cognitive styles on their query reformulation behaviour. Query reformulation is important because users' ability to transform information need into machine language that is understandable by IS dependents on their query reformulation skills.

Given the importance of cognitive style aspect of human-computer interaction, and query reformulation during Web searching, this study aims to explore (1) users' query reformulation, and (2) the effects of users' cognitive styles on their query reformulation behaviour.

RESEARCH AIMS AND QUESTIONS

The studies reviewed in the preceding sections show that a very limited body of work in HCI fields has explored effects of users' cognitive styles on their query reformulation behaviour. This study aims to investigate the relationships between users' cognitive styles and their query reformulation behaviour while performing Web searches.

The findings from this study will help search systems to provide effective and automatic query suggestions that can help users to improve their search results. This study attempts to address the following research questions:

RQ1: What are the types of query reformulations executed during Web searching and which types do users mostly follow during information searching on the Web?

RQ2: Are there any associations between different types of query reformulations?

RQ3: How does users' cognitive styles influence their query reformulation behaviour during Web searching?

RESEARCH DESIGN

Study Participants

A total of 50 volunteers from the Queensland University of Technology participated in the study. The research sample size was chosen based on the prevailing research practice in user studies. Many user study researchers tended to use a small group of participants, less than 70 participants (examples: Gwizdka, 2009; Liu, et al., 2010; Wang, Hawk, & Tenopir, 2000). Efforts were made to include equal number of males and females across different age group and occupations, such as student, academic or professional staff.

Search Tasks

Based on Borlund and Ingwersen's (1997) concept of "simulated work task situation" or scenarios, we developed three search tasks to ensure that these tasks are as close as possible to the real world situations. The simulated work task situation provides each searcher with the context, which ensures "a degree of freedom" to react in relation to his or her interpretation of the given situation. This approach has been used by several researchers in information seeking studies (examples include: Borlund, 2003; Kim, 2009).

Task 1: You, with your two friends, are planning a trek for one week in Solukhumbu in Nepal. The trekking will occur next month. You are told that tourists trekking in the place may get high-altitude illness. You decide that

you should know more about the place, and symptoms, seriousness and preventions of high-altitude sickness.

Task 2: You have recently moved to Austin, Texas, the US and would like to know the relevant laws passed by the Texas state government regarding child safety while travelling in vehicles. Identify three such rules.

Task 3: You recently heard about the Bermuda Triangle mystery, and you are curious and want to know more about it. So, you want to search any relevant information (articles, images and videos) about it and what effect it has on the travellers in the region.

Query Reformulation Taxonomy

Similar to the previous works in query reformulation type (Hoang, Nguyen, & Tjoa, 2008; Jansen, et al., 2009; Tseng, et al., 2009), we constructed five reformulation categories based on the common and difference in search terms use in two successive queries. Detailed definition of each of these query reformulation classifications with examples are given in Table 1.

Data Collection

Demographic information, including prior search experience, was collected by using a questionnaire. Riding's (1991) *Cognitive Style Analysis* (CSA) test was used to measure participants' wholist-analytic (WA) and verbal-imagery (VI) cognitive styles (Riding & Cheema, 1991).

The CSA test indicates the position of a user on the WA and VI cognitive style dimensions by means of a ratio. Based on the WA and VI ratios, participants were classified as *wholist* or *analytic* on the WA dimension and *verbaliser* or *imager* on the VI dimension of cognitive styles. In other words, participants were assessed on the two dimensions of cognitive style.

Participants scoring below 1.20 on the WA scale were classified as wholist, and those scoring 1.20 or above as analytic. Similarly, participants scoring below 1.03 on the VI scale were classified as verbaliser and those scoring 1.03 or above as imager. Thus, a participant can be wholist-verbaliser, wholist-imager, analytic-verbaliser, or analytic-imager.

Procedure

Prior to the Web searching experiment, each participant was briefed with the participation instruction and asked to fill up a questionnaire. The study participants were then assigned three search topics. Participants had the liberty to use any search engines and search options of their choice, and search at their own speed.

However, participants were recommended spending between 10 and 15 minutes on each task. The participants were reminded to save or bookmark the relevant information retrieved. More explanations were given to those participants who were not familiar with the search task topics. Each participant was provided with a laptop with Internet access. A monitoring program was used to capture Web search experiment.

Data Analysis

Participants' interactions with the search engines were logged using a monitoring program. The captured user-Web interactions for each participant were played several times to create Web search session logs. A standard search log file format with the following fields, similar to that of Jansen (2006), was adopted:

- User Identification: A unique number was used to identify a participant.
- Date: The date of the interaction.
- Time: The duration of the interaction.
- URL: The URL of the Web site visited.
- Search Terms: The query terms as entered by the user.

The data was then manually entered into the Web search session logs. We developed a program to automatically classify queries into five classifications: *New*, *Add*, *Remove*, *Replace* and *Repeat* queries (see Table 1 for definition). The query classification of the data were then manually checked and verified. Table 2 illustrates some examples of the Web search session logs with the automatically classified query reformulation types. While data in the first five

Query	Description	Query Examples
New	Qi and Qi+1 do not contain any common terms. All new session terms are assigned as a new query.	Q_i: "tour" Q_{i+1}: "Solukhumbu trek"
Add	Q_i is a subset of Q_{i+1}, that is, all the terms in Q_i are present in Q_{i+1} and Q_{i+1} contains more terms than Q_i.	Q_i: "Trekking Solukhumbu" Q_{i+1}: "Trekking Solukhumbu Nepal"
Replace	Q_i and Q_{i+1} contain at least one term in common and at least one different term.	Q_i: "Tour Nepal" Q_{i+1}: "Tour Solukhumbu"
Remove	Q_{i+1} is a super subset of Q_i, that is, all the terms in Q_{i+1} are present in Q_i and Q_i contains more terms than Q_{i+1}.	Q_i: "Solukhumbu tourist Nepal" Q_{i+1}: "tourist Nepal"
Repeat	Q_i and Q_{i+1} contain exactly the same terms; the order of these terms may be different.	Q_i: "trekking Solukhumbu Nepal" Q_{i+1}: "Nepal Solukhumbu trekking"

Note: Q_{i+1} is the succeeding query that follows the query Q_i in the same session

Table 1. Classifications of query reformulations with examples

User_ID	Date	Time	URL	Search Terms	QRT
41	03/02/10	14:00	www.google.com	hiking oxygen	New
41	03/02/10	14:00	www.google.com	portable oxygen tanks	Add
41	03/02/10	14:02	www.google.com	how to use hiking oxygen	Add
41	03/02/10	14:03	www.google.com	bermuda triangle	New
41	03/02/10	14:04	www.google.com	bermuda triangle	Repeat

Note: QRT: Query Reformulations Type

Table 2. Examples of Web search session logs with Query Reformulations Type

columns were added manually, the query reformulation types (QRT) in the last column of the table were inserted automatically by the query-classifying program.

A series of statistical analyses were conducted to examine the findings of the study. A basic frequency distribution analysis was conducted to find the occurrences of each of the query reformulation types. A parametric Pearson correlation analysis was performed to establish correlations between the various types of query reformulations. A series of one-way Analysis of Variance (ANOVA) were performed to find significant differences in query reformulation behaviour among the users of different cognitive style groups.

Path Analysis (PA) was used to reconfirm the associations established by the correlation analysis between query reformulation types, and then to construct a path model to graphically represent the relationships. Path analysis is an approach to modelling explanatory relationships between observed variables (Raykov & Marcoulides, 2006, p. 63). It is considered as a member of the Structural Equation Modelling (SEM) family (Kline, 2011). SEM analyses are based on correlation data to examine the relationships among two or more variables and constructs.

RESULTS

Demographic

A total of 50 users, comprising of students, and academic and professional staff from the Queensland University of Technology participated in this study. They regularly search the Web for information in the course of their academic, personal or administrative activities. 26 of them were males while, 24 were females. The participants' age varied from 20 years to 55 years. Their demographic information indicated that they had a minimum of 3 years Web search experience.

Although their demographic information might have contributed significantly to this study, participants were not differentiated by their demographic data, as it is not a controlled variable in this study. Only the participants' cognitive styles were considered in this study.

Cognitive Styles

Based on Riding's (1991) CSA test, participants were classified as *wholist* or *analytic* on the wholist-analytic (WA) dimension of cognitive styles, and *verbaliser* or *imager* on the verbal-imagery (VI) dimension of cognitive styles. In other words, a participant can be wholist-verbaliser, wholist-imager, analytic-verbaliser, or analytic-imager.

Table 3 illustrates the distribution of the study participants according to their cognitive style groups. On the WA dimension of cognitive styles, 23 participants out of 50 were classified as having a wholist cognitive style while, 27 participants were identified as having analytic cognitive style. On the VI scale, 24 participants were classified as verbal users while, 26 participants were imagers.

Query Reformulation Types (RQ1)

A total of 3613 unique search terms and 872 search queries were submitted by 50 users who engaged in 150 scenario-based search tasks. The average query length was 4.14. Based on the definition of the query reformulation types given in Table 1, we identified five types of query reformulation types: *New*, *Add*, *Remove, Replace*, and *Repeat*. We developed and used a program to automatically classify queries into the five classifications. The distributions of query reformulation types are illustrated in Table 4.

As reported in the prior works (Jansen, et al., 2009; Tseng, et al., 2009), the *New*, accounting for 28.90% of the total query reformulations, dominated amongst the

Cognitive Styles		Total
WA	Wholist	23
	Analytic	27
VI	Verbaliser	24
	Imager	26

Table 3. Distribution of participants according to their cognitive style group

Query type	Occurrence	Mean	SD	%
New	252	5.04	2.08	28.90%
Add	218	4.36	2.72	25.00%
Remove	102	2.04	1.58	11.70%
Replace	174	3.48	3.61	19.95%
Repeat	126	2.52	2.02	14.45%
Total	872			100%

Table 4. Distributions of query reformulations

query formulations while performing Web searching. Users are more likely to submit new queries when they change their topic of searches. The *Remove* query reformulation was the least of the query classifications with 11.70%, almost 20% lesser than the *New* query reformulations. The variation in the number of *Replace* queries (SD=3.61) by the users is relatively higher than the variation in any of the other query reformulations, the least being in *Remove* queries.

In general, Web users seemed to reformulate their queries by either adding completely new search terms, some new search terms or by replacing some terms. Among these queries, users tended to prefer *New* queries.

Associations within Query Reformulation Types (RQ2)

To find significant associations between different types of the query reformulations, a Pearson correlation was carried out. The test showed a significant positive correlation between *Add* and *Remove* (r =0.609, p<0.05), *Remove* and *Replace* (r =0.380, p<0.05), and *Add* and *Replace* (r =0.532, p<0.05) at significant level p<0.05. The Pearson correlation result indicated that when users perform *Remove* or *Replace* query reformulations, they also tended to add search terms to their queries. Similarly, when users replace search terms they also tended to remove keywords.

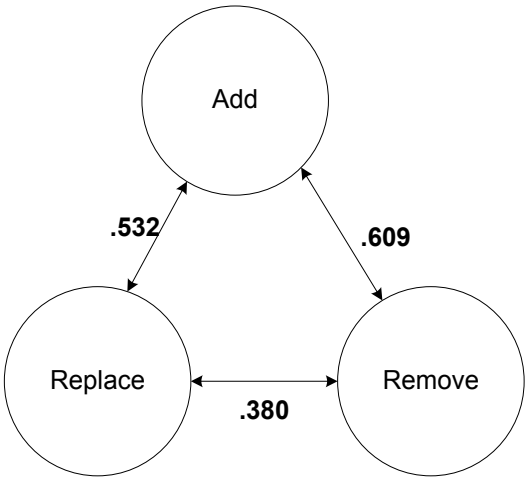
We constructed a path model to represent graphically the relationships between *Add*, *Remove* and *Replace* as indicated by the Pearson correlation test. Figure 1 illustrates the path model for these query reformulations. It summarizes the three-trio relationships between the three query reformulations. In the figure, the two-headed arrows represent the interrelationships among the *Add*, *Remove* and *Replace* query types.

The integer number between any two-query types (say 0.609 between *Add* and *Remove*) represents a Pearson correlation value between them at a significant level p<0.05. Higher the correlations value between two queries, stronger the relationship between them. As shown in the figure, the correlation between *Add* and *Remove* is stronger than between *Add* and *Replace*, or between *Replace* and *Remove*.

Effects of Cognitive Styles on Query Reformulation Behaviour (RQ3)

Table 5 illustrates the occurrence of query reformulations for each cognitive style group. The last column of the table illustrates aggregate number of queries submitted by a user to complete three search tasks.

The table indicated that wholists, verbalisers and imagers in general preferred *New* queries among the five types of queries, while analytic users preferred *Add* to other types of queries. *Remove* query type was the least preferred query for all users (in all cognitive style groups). This indicated that the least the users wanted to do during query reformulation was to remove search terms.



Note: All paths are statistically significant at P<0.05

Figure 1: Path Model for *Add*, *Remove* and *Replace* Query Reformulations.

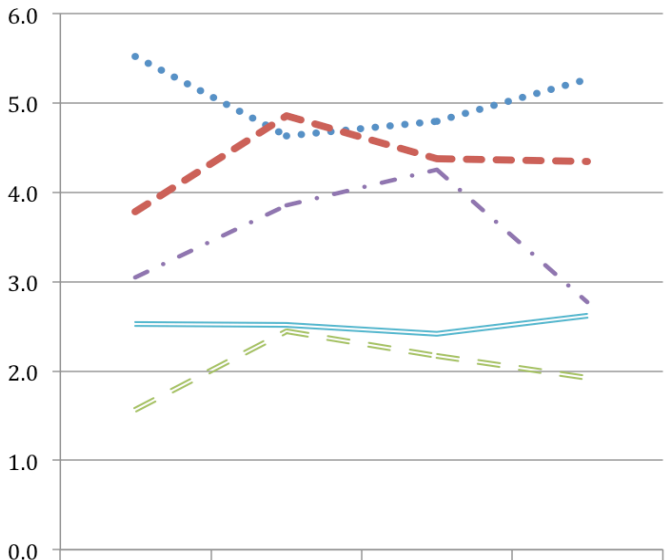
The table also shows that analytic users in general are more active in query reformulation, for they formulated relatively the highest number of queries on an average to complete three search tasks (Mean = 18.30). On the other hand, wholists executed the least number of queries to complete three search tasks (Mean = 16.43). It may be due to the fact that wholists are said to be less systematic and less logical compared to analytics (Wang, et al., 2000).

Figure 2 illustrates mean distribution of query reformulation types in each cognitive style group to complete three search tasks. When comparing the query reformulation distributions among the five types, on the WA dimension, as illustrated in the figure, wholist users submitted the highest number of *New* queries. They seemed to submit more *New* queries due to the fact that they are said to have problems separating ideas into discrete parts (Riding & Cheema, 1991), and they encounter more difficulties and confusion during Web searching (Wang, et al., 2000). *New* queries required lesser query reformulation skills compared to *Add*, *Remove* or *Replace* queries.

On the other hand, the analytic users submitted highest number of *Add*, *Remove* and *Replace* queries compared to wholists. They modified their existing queries by adding, removing or replacing some search terms. Analytics are said to perceive a situation as a collection of parts and focus on one or two aspects of the situation at a time (Riding & Cheema, 1991). Our findings are in line with what Ford and his team (2009) have found. They reported that analytics display consistency and more systematic logical

Cognitive Style	New	Add	Remove	Replace	Repeat	Total	Mean
Wholist (n =23)	127	87	36	70	58	378	16.43
Analytic (n =27)	125	131	66	104	68	494	18.30
Verbaliser (n = 24)	115	105	52	102	58	432	18.00
Imager (n = 26)	137	113	50	72	68	440	16.92

Table 5. Occurrence of query reformulations in each cognitive style group



	Wholist	Analytic	Verbaliser	Imager
New	5.5	4.6	4.8	5.3
Add	3.8	4.9	4.4	4.3
Remove	1.6	2.4	2.2	1.9
Replace	3.0	3.9	4.3	2.8
Repeat	2.5	2.5	2.4	2.6

Figure 2. Mean query reformulations for each cognitive style group

thought with their search transformations to a greater extent onto the more complex phrase-oriented expressions.

On the VI dimension of cognitive styles, verbalisers submitted relatively higher number of *Add*, *Remove* and *Replace* queries as compared to their imagery peers. These queries involve a greater use of combinations of keywords. Our findings confirm previous studies that analytics tend to be good at the use of language as opposed to imagers, with greater use of distinctive linguistic search transformations and good combination of keywords (Ford, et al., 2009).

On the other hand, imagers completed a higher number of *New* and *Repeat* query reformulations than verbalisers did. Imagers lack linguistic expression; therefore they either formulate their queries with completely new search terms or tended to search with the same query (on Google Web, Google images, and Google videos).

On a holistic view as illustrated in Figure 2, among the users, on aggregate wholists executed the maximum number of *New* queries (Mean = 5.5), analytic users formulated the highest number of *Add* (Mean = 4.9) and *Remove* queries (Mean = 2.4), verbaliser submitted the highest number of *Replace* queries (Mean = 4.3), while imagers formulated the highest number of *Repeat* queries (Mean = 2.6) to complete three search tasks.

To find significant differences among the users of different cognitive styles in their query reformulation behaviour (in terms of five types of queries), a series of One-way ANOVA were performed. However, the results showed a little difference among the searchers in their query reformulation behaviour. A significant difference was found among wholists and analytics in their *Remove* query reformulation behaviour, F (1, 48) = 4.103, $p< 0.05$, which indicated that the wholists and analytics performed *Remove* query reformulations differently.

DISCUSSIONS AND IMPLICATIONS

Study results indicated that users formulated a great amount of queries while searching information on the Web, which can be classified into five categories: *New*, *Add*, *Replace*, *Remove* and *Repeat*. Users seemed to execute more *New* queries to search information on the Web. They tended to use lesser *Remove* queries, which indicated that users seldom reduced their search terms while formulating and submitting queries to search engines in a search session.

A statistical correlation test revealed significant associations between *Add* and *Remove*, *Remove* and *Replace*, and *Replace* and *Add* query reformulation behaviour. This indicated that certain group of users tended to prefer certain query reformulations. The correlations between these pairs also varied from one pair to another. For instance, the correlation between *Add* and *Remove* is stronger than between *Remove* and *Replace*. We developed a path analysis model to demonstrate these inter-relationships (Figure 1).

The study findings also show relationships between users' cognitive styles and their query reformulation behaviour. Among the users of four cognitive style groups analytic users were found to be more active in query reformulation, for they submitted relatively more number of queries while searching information on the Web than users of other cognitive style groups.

On the WA dimension of cognitive style, wholists were found to utilize a higher number of *New* queries. This indicated that they seemed to lack query reformulation skills because new and repeated queries may not have retrieved more relevant information than otherwise would have done if they have added, removed or modified some search terms. On the other hand, analytic users submitted a higher number of *Add*, *Remove* and *Replace* queries compared to their wholist peers. This indicated that analytic users seemed to be better than their wholist peers in query reformulations.

Verbalisers are believed to think in terms of words and consider the information they read, see or listen in terms of words (Riding & Cheema, 1991), which added extra leverage to their query reformulation skills. Thus, they were found executing relatively a higher number of *Add*, *Remove* and *Replace* queries than their imagery peers. On the other hand, imagers were found searching information on the Web with same query (*Repeat*) on different search engines, such as Google Web, Google images and Google videos. It may be due to the fact that they think in terms of "mental pictures" (Riding & Cheema, 1991), therefore they lack linguistic expressions.

Our results showed that users' cognitive styles play an important role in their query reformulation behaviour while searching information on the Web. The study findings have theoretical implications for researchers and practical implications for IS developers. Understanding how users with different cognitive styles formulate and reformulate their queries while information searching can help HCI researchers and IS developers to develop an interactive and user-cantered search model, and search interface respectively.

CONCLUSION AND FUTURE WORK

The aim of this study was to examine the effects of users' cognitive styles on their query reformulation behaviour. A total of 3613 search terms and 872 search queries were submitted by 50 users who completed 150 scenario-based search tasks. We developed five types of query reformulation taxonomies and built a program to automatically classify users' queries into five query classifications. A series of statistical analyses were applied to determine (1) associations within five query reformulation types, and (2) effects of users' cognitive styles on their query reformulation behaviour. The study findings provided valuable insights into query reformulation behaviour amongst users with different cognitive styles.

Query reformulation transforms a user's information search into machine language that is understandable by information systems, and users' query reformulation behaviour is greatly influenced by their cognitive styles. Thus, understanding how users with different cognitive styles formulate queries during Web searching can help HCI researchers and IS designers to provide assistance to users to locate relevant information. Search engines can identify the type of information the user is looking for by capturing the trend of the query reformulations, and then provide effective query suggestions accordingly.

In future research, we aim to develop a model of query reformulation during Web searching that integrates users' cognitive styles. Modelling query reformulation with a greater understanding of users' cognitive styles can help to bridge the semantic gap between the information users and the IS. It can improve users' Web search experience by providing context based query suggestions. A model depicting relationships between users' Web search behaviour in general and query reformulation behaviour in particular and their cognitive styles will be reported in future papers.

We also aim to examine effects of users' demographic information, such as gender and age, and search tasks types on their query reformulation behaviour. These factors might have significant effects on users query reformulation behaviour, which were not considered in this study.

REFERENCES

Borlund, P. (2003). The IIR Evaluation Model: A Framework for Evaluation of Interactive Information Retrieval Systems. *Information Research, 8*(3), 8-3.

Borlund, P., & Ingwersen, P. (1997). The development of a method for the evaluation of interactive information retrieval systems. *Journal of Documentation, 53*(3), 225-250.

Fidel, R., & Pejtersen, A. (2004). From information behaviour research to the design of information systems: the Cognitive Work Analysis framework. *Information Research, 10*(1), 10-11.

Ford, N., Eaglestone, B., Madden, A., & Whittle, M. (2009). Web Searching by the "general public": An Individual Differences Perspective. *Journal of Documentation, 65*(4), 632-667.

Gorrell, G., Eaglestone, B., Ford, N., Holdridge, P., & Madden, A. (2009). Towards "metacognitively aware" IR Systems: An Initial User Study. *Journal of documentation, 65*(3), 446-469.

Gwizdka, J. (2009). What a difference a tag cloud makes: Effects of tasks and cognitive abilities on search results interface use. *Information Research, 14*(4).

Hoang, H., Nguyen, T., & Tjoa, A. (2008). A Semantic Web-Based Approach for Context-Aware User Query Formulation and Information Retrieval. *International Journal of Information Technology and Web Engineering, 3*(1), 1.

Huang, J., & Efthimiadis, E. (2009). *Analyzing and Evaluating Query Reformulation Strategies in Web Search Logs.* In Proc. the CIKM'09:The 18th ACM Conference on Information and Knowledge Management.

Jansen, B., Booth, D., & Spink, A. (2009). Patterns of query reformulation during Web searching. *Journal of the American Society for Information Science and Technology, 60*(7), 1358-1371.

Jansen, B., Zhang, M., & Spink, A. (2007). Patterns and transitions of query reformulation during Web searching. *International Journal of Web Information Systems, 3*(4), 328-340.

Jansen, B. J. (2006). Search Log Analysis: What it is, What's been done, How to do it. *Library and Information Science Research, 28*(3), 407-432.

Kim, J. (2009). Describing and predicting information-seeking behavior on the Web. *Journal of the American Society for Information Science and Technology, 60*(4), 679-693. doi: 10.1002/asi.21035

Kim, K. (2000). *Effects of Cognitive Style on Web Search and Navigation.* In Proc. the Conference on Educational Multimedia, Hypermedia & Telecommunications ED-MEDIA 2000: 496-501.

Kinley, K., & Tjondronegoro, D. (2010a). *The Impact of Users' Cognitive Style on Their Navigational Behaviors in Web Searching.* In Proc. the ADCS2010, Melbourne, Australia: 68-75.

Kinley, K., & Tjondronegoro, D. (2010b). *User-Web Interactions: How Wholistic/ Analytic Web Users Search the Web?* In Proc. the OZCHI 2010, Brisbane, Australia: 344-347.

Kinley, K., Tjondronegoro, D., & Partridge, H. (2010). *Web Searching Interaction Model based on User Cognitive Styles*. In Proc. the OZCHI 2010, Brisbane, Australia: 340-343.

Kline, R. B. (2011). *Principles and Practice of Structural Equation Modeling*. New York: The Guilford Press.

Liu, C., Gwizdka, J., Liu, J., Xu, T., & Belkin, N. (2010). *Analysis and Evaluation of Query Reformulations in Different Task Types*. In Proc. the American Society for Information Science and Technology.

Pask, G. (1976). Styles and Strategies of Learning. *British Journal of Educational Psychology, 46*(1), 128-148.

Raykov, T., & Marcoulides, G. A. (2006). *A first course in structural equation modeling* (2nd ed.). Mahwah, New Jersey London: Lawrence Erlbaum.

Riding, R. (1991). *Cognitive Styles Analysis*. Birmingham: Learning and Training Technology.

Riding, R., & Cheema, I. (1991). Cognitive Styles—An Overview and Integration. *Educational Psychology, 11*(3), 193-215.

Rieh, S., & Xie, H. (2006). Analysis of multiple query reformulations on the web: The interactive information retrieval context. *Information Processing and Management, 42*(3), 751-768.

Sutcliffe, A., & Ennis, M. (1998). Towards a cognitive theory of information retrieval. *Interacting with Computers, 10*(3), 321-351.

Tseng, L., Tjondronegoro, D., & Spink, A. (2009). *Analyzing web multimedia query reformulation behavior*. In Proc. the 14th Australasian Document Computing Symposium, University of New South Wales, Sydney, NSW: 118-125.

Wang, P., Hawk, W. B., & Tenopir, C. (2000). Users' Interaction with World Wide Web Resources: An Exploratory Study Using a Holistic Approach. *36*, 229-251.

Witkin, H., Moore, C., Goodenough, D., & Cox, P. (1977). Field-dependent and field-independent cognitive styles and their educational implications. *Review of educational research, 47*(1), 1-64.

Yoon, J., & Chung, E. (2009). Image query reformulation over different search stages. *Proceedings of the American Society for Information Science and Technology, 46*(1).