Contents | **Author index** | **Subject index** | **Search** | **Home**

What a difference a tag cloud makes: effects of tasks and cognitive abilities on search results interface use

Jacek Gwizdka
Department of Library and Information Science, School of Communication and Information, Rutgers University, New Brunswick, NJ 08901, USA

Abstract

Introduction. The goal of this study is to expand our understanding of the relationships between selected tasks, cognitive abilities and search result interfaces. The underlying objective is to understand how to select search results presentation for tasks and user contexts
Method. Twenty three participants conducted four search tasks of two types and used two interfaces (List and Overview) to refine and examine search results. Clickthrough data were recorded. This controlled study employed a mixed model design with two within-subject factors (task and interface) and two between-subject factors (two cognitive abilities: memory span and verbal closure).
Analysis. Quantitative analyses were carried out by means of the statistical package SPSS. Specifically, multivariate analysis of variance with repeated measures and non-parametric tests were performed on the collected data.
Results. The overview of search results appeared to have benefited searchers in several ways. It made them faster; it facilitated formulation of more effective queries and helped to assess search results. Searchers with higher cognitive abilities were faster in the Overview interface and in less demanding situations (on simple tasks), while at the same time they issued about the same number of queries as lower-ability searchers. In more demanding situations (on complex tasks and in the List interface), the higher ability searchers expended more search effort, although they were not significantly slower than the lower ability people in these situations. The higher search effort, however, did not result in a measurable improvement of task outcomes for high-ability searchers.
Conclusions. These findings have implications for the design of search interfaces. They suggest benefits of providing result overviews. They also suggest the importance of considering cognitive abilities in the design of search results' presentation and interaction.

CHANGE FONT

Introduction

Web search is becoming a dominant mode of obtaining information in our society. Almost half of the Internet users use search engines daily (Fallows 2008). This makes Web search, after e-mail, the second most popular activity in which people engage on the Internet. With the widespread use comes an increasing diversity of people who search for information on the Web. It has been argued (e.g., Belkin 2008, Ingwersen & Järvelin, 2005) that the *one size fits all* approach to information search systems is not likely to produce many further significant improvements and that to

achieve future improvements the search systems will need to take into account other (traditionally not considered) factors that affect the search process. These other factors include user characteristics, user situation and context (Belkin, 2008, Ingwersen & Järvelin, 2005). In particular, human perception and cognition are engaged at multiple levels during information search. First, information search itself is cognitive in nature (Ingwersen 1996). Second, interaction with computers (that mediate the search) relies on human perception and cognition (Card *et al.* 1983). An ever-wider diversity of users who use the Web increases the range of cognitive characteristics that need to be considered by system designers. Clearly, the role of cognitive characteristics of users in the search process should be taken into account. The study presented in this paper aimed to expand our understanding of the relationships between selected cognitive abilities, search result interfaces and user tasks.

Related work

This section first introduces the concepts of cognitive abilities and styles. It then presents very selectively research in human-computer interaction and information science that considered the effects of differences in cognitive characteristics. This short overview aims to show a sampling of obtained results. It then focuses on research that examined search results presentation and particularly on studies that examined relationships between result presentations and cognitive factors. The section closes with a short discussion of tagging and tag clouds.

Cognitive abilities and styles

With the recognition of information search as a cognitive process (Belkin 1980, Ingwersen 1996) the need to understand the role of cognitive factors in the search process has been acknowledged since the 1980s. Cognitive characteristics of people were studied in relation to general human-computer interaction, but also in more specific information task contexts. Researchers have studied mainly two kinds of cognitive characteristics of users: cognitive abilities and cognitive styles.

The notion of *cognitive abilities* comes from the tradition of psychometric intelligence testing in psychology (Kline 2000). Cognitive ability refers to human ability to perform cognitive tasks, that is, tasks '*in which correct and appropriate processing of mental information is critical to successful performance*' (Carroll 1993: 10). One of the best known systems of cognitive abilities is Carroll's three-stratum theory. The following cognitive abilities (from Carroll's first stratum) are of particular interest to research in information science:

- *Memory span*: ability immediately to recall digits, letters, or other items (one of the measures of working memory performance);
- *Working memory*: ability to temporarily store and perform a set of cognitive operations on information that requires divided attention and the management of the limited capacity resources of short-term memory;
- *Spatial ability* (spatial relations and spatial scanning): ability to visualize figures in different orientations (perceive, manipulate, visually explore);
- *Verbal closure*: ability to identify visually presented words when some letters are missing, scrambled, or embedded among other letters.
  (These definitions are based on McGrew (1997) and Kline (2000).)

Cognitive styles represent a person's typical modes of perceiving, remembering, thinking and problem solving (Messick 1970). Riding & Rayner (1998) described cognitive styles as person's preferred and habitual approach to organizing and representing information. Harrison & Rainer (1992) see the styles as personality dimensions that influence how an individual collects, analyses, evaluates and interprets information. It is thus not surprising that cognitive styles have received attention in the information science community and that several styles were studied as factors influencing person's interaction with computing systems on information tasks. The most commonly studied styles include field dependence versus field independence (Messick 1970) and visualizer versus verbalizer (Witkin & Goodenough 1981).

Consideration of cognitive factors clearly relates to individual differences between people, which are systematically studied in psychology. In information science and human-computer interaction the differences among users are not the main focus, hence their study was not always seen as relevant. Ford (1986) was probably the first to draw attention to role of individual differences in information behaviour, followed by Borgman (1989) and Saracevic (1991). In human-computer interaction, one of the first was Egan (1988). Following these early papers, the last two

decades have seen numerous studies concerned with the effects of cognitive characteristics of users. Below, we provide a short overview of these studies.

Egan (1988) discussed the differences in user performance on several common computing tasks (including typing and form filling) and showed that the differences can be on the order of 20:1. Egan suggested that these results could be ascribed to individual differences (such as computing experience, technical aptitudes and cognitive abilities), that they could be predicted and that they could be modified (minimized) through appropriate design. Following Egan's work, other researchers demonstrated reduction of the performance gap between different population groups through appropriate design modifications. Sein *et al.* (1993) conducted a study examining the effects of visual ability on the users' ability to learn three software applications. Use of a direct manipulation interface led to a reduced difference between high- and low-visual-ability users in their study. Zhang & Salvendy (2001) investigated the effects of users' visualization ability and Website structure display design on Web browsing performance. They found that structure preview reduced the differences in performance between high and low visualization ability users.

These projects studied the effects of cognitive abilities from the perspective of improving interface usability to accommodate the general public. Other researchers focused more specifically on the effects of cognitive abilities on user interaction with computer-based systems. Two abilities, working memory and spatial ability, received considerable attention. The effects of spatial ability on performance were studied in virtual environment navigation as well as in hypertext (Modjeska and Chignell 2003), in textual information retrieval (Westerman 1995) and in visual information retrieval (Westerman and Cribbin 2000). Limited capacity of working memory is a well known bottleneck in human information processing (Baddley 1986, Miller 1956). The role of individual differences in the capacity of working memory in graphical information processing was shown, for example, by Lohse (1997).

In information science, most research on the effects of individual differences in cognitive factors focused on cognitive styles. In particular, field dependence versus field independence was shown to reflect how well an individual is able to restructure information based on the use of salient cues and field arrangement (Weller *et al.* 1994). The effects of field dependence versus field independence on user performance of information tasks were demonstrated in several studies by Allen, Kim and their colleagues (Allen 2000, Kim 2001, Kim and Allen 2002, Palmquist and Kim 2000). More recently, Ford *et al.* (2001, 2005a and 2005b) showed relationships between the searchers' individual differences (not only cognitive styles, but also including experience and age), search task performance and search strategies. Field dependence versus field independence was also shown to have a significant effect on users' information seeking behaviour (Chen and Ford 1998, Ford and Chen 2000). In this line of research, Chen and colleagues (Chen *et al.* 2005, Chen *et al.* 2004) created a flexible Web directory that accommodated both users who were field independent and those who were field dependent.

Search results presentation

Representation of search results is one of the important aspects of information retrieval systems. At the fundamental level, representation is important because it affects the users' accuracy, effectiveness and efficiency in their assessment of relevance. This assessment can apply to individual results, to result sets or to whole databases. The representation should inform users about the relationship between their information need (expressed in a query or in a selection of clicked links) and the obtained documents. Relevance assessment is facilitated by the amount and the type of information available about each returned result. A typical Web search engine's result listing shows for each returned result a page title, URL, text snippet and the page size. Text continues to be the dominant representation of search results, although a number of graphical elements were proposed to complement or replace textual results. Web languages (e.g., HTML, CSS) allow for an expression of a wide range of visual attributes, such as page layout and colour, which could be used to facilitate relevance assessment. For example, page thumbnails were proposed by many researchers (Dziadosz and Chandrasekar 2002, Joho and Jose 2008, Woodruff *et al.* 2001, Woodruff *et al* 2002). Thumbnails can be optionally added to Google results. TileBars interface (Hearst 1995) used visual surrogates to show the relationships between the words in a query and the returned results.

Our discussion thus far has been mainly concerned with individual result surrogates. Making explicit the structure and relationships between search results could be useful for searchers. A few existing and proposed search systems added metadata (understood very broadly) to the search interface or to the search results display. The metadata were added next to the search results or the search results were embedded in the metadata. A number of different types of metadata were used in the experimental search systems. These included categories (Buntine *et al.* 2005, Dumais *et al.* 2001); classification codes (Shiri and Revie, 2003) and facets (English *et al.* 2002, Kules *et al.* 2009). The results of

evaluations showed, in general, the usefulness of metadata. For example, Kules *et al.* (2009) showed the usefulness of facets in the exploratory search process. Dumais *et al.* (2001) showed that embedding individual results in an explicit representation of a collection structure (categories) improved user efficiency.

Providing overviews that represent collections of returned search results was suggested by Greene and colleagues (Greene *et al.* 2000). Recently, search results overview (summary) in a form of a cloud was examined for PubMed searches by Kuo *et al.* (2007). They found that the cloud used to summarize Web search results is good for presenting descriptive information and reducing user frustration, but that cloud display is less effective for discovering relations between concepts. Simpler, descriptive questions were answered more correctly by using a cloud, while more complex, relational question were answered more correctly in the standard PubMed interface. Descriptive questions were answered faster using the cloud, while relational questions were answered two times faster in the standard interface. Although users were slower in the tag cloud, they rated their satisfaction higher.

The design space of search result surrogates is quite large. We highlighted just a few selected approaches from the large body of work that explored many dimensions of this space. However, in spite of the many techniques that were proposed, the systematic study of effectiveness of search results display is relatively limited (Joho and Jose 2008: 227). Joho and Jose evaluated effectiveness of adding top ranking sentences and thumbnails to search result display and found that less experienced searchers are more likely to find additional representations useful in relevance assessments. However, their systematic study of document surrogates is one of few. A systematic study should include examination of interactions among user characteristics, tasks and search result presentations. Previous research suggested that different presentations of search results might be appropriate for different tasks and contexts (Joho and Jose 2008, Woodruff *et al.* 2002). Results from prior work also indicated that users' search experience could influence perception of search interface features and their use (Joho and Jose 2008, White *et al.* 2003). Interaction with a search interface is also affected by its overall visual complexity. More complex visual representations of search results carry with them a cost to the user. Visual clutter (Rosenholtz *et al.* 2007) can contribute to increasing cognitive load that is imposed by the search system on its users (Harper *et al.* 2009). Given that human cognitive resources are limited (Baddley 1986, Sweller *et al.* 1998), the searchers trade mental effort needed for completion of their information search tasks for possibly unnecessary effort required by the system. The goal of designing search interfaces should be to minimize cognitive load imposed by the interface on their users.

Based on the earlier discussion of cognitive abilities and styles it should be clear that one could expect their influence on user interaction with different search result presentations. At the basic level, one can expect that different people will express preferences for specific information presentations. Indeed, Krishnan and Jones (2005) found that some people preferred to access files through folders shown in a spatial representation, while others preferred textual keyword-based search to access their files. One could also expect interaction among individual cognitive differences and search result presentations. However, there is little work that deals with these aspects. One of few notable exceptions is the work by Chen *et al.* (2004, 2005). They investigated the relationships between cognitive styles and information presentations in Web directories and demonstrated that cognitive style is correlated with user preferences for specific presentation elements in these directories. The researchers found that field dependent and field independent users demonstrate different preferences for the organization of subject categories, the presentation of results and the screen layout. In particular, field independent users prefer alphabetical sorting of results, while field dependent users prefer sorting by relevance and extra help in assessing relevance. We note that result presentation in Web directories bears many similarities to result presentation in search engines, hence similar findings could be expected in the context of search engine results.

Collaborative tagging and tag clouds

Many Web 2.0 Websites allow their users to associate words (known as tags) with the users' own and with other users', content. The three sets of entities (users, Web resources and tags) become associated with each other in the process of collaborative tagging (Moulaison 2008). The result is an additional layer that provides descriptions of Web resources. Its creation is motivated by the need to communicate and organize information (for self, for friends and for the public) (Nov *et al.* 2008). The created associations between tag sets and Web resources are used to facilitate Web resource re-finding, to promote understanding of resources, to identify resource qualities, to enable social navigation, to create communities of interest, to attract attention, to identify ownership and to express opinions (Golder and Huberman 2006, Marlow *et al.* 2006), to name just a few. From the perspective of information search, tags can be viewed as serving two main purposes: 1) promoting understanding of the associated set of Web resources and 2) facilitating navigation of the collection of Web resources. At each step of the navigation process, tags describe the set

of associated resources, but it also serve as starting points for further navigation.

Tags are frequently presented visually in a form of a list or a cloud (Rivadeneira *et al.* 2007, Sinclair and Cardew-Hall 2008) (Figure 1). The cloud representation was popularized by social tagging Websites such as Delicious, Digg, Citeulike).

Figure 1. A sample tag cloud. The tag cloud represents topics covered in an HCI course taught by the author.

The design space of clouds has a number of features. It is still not clear how to design the *best* tag cloud and how to select the *best* cloud features. The features include tag layout (e.g., box layout, radial layout), tag font size (e.g., no variation, variation according to tag frequency, where the font size is often proportional to the log of its frequency (Sinclair and Cardew-Hall 2008)), tag font colour and order of tags (e.g., based on tag frequency, or alphabetical). A number of studies evaluated selected tag cloud features (Rivadeneira *et al.* 2007, Halvey and Keane 2007, Bateman *et al.* 2008). We learned that different visual features seem to support different cloud functions. For example, alphabetically ordered clouds were shown to aid information finding (Halvey and Keane 2007), larger tag font size was shown to better support recall of words (Rivadeneira *et al.* 2007) and the tag font size was found to affect the perception of tag importance in a cloud (Bateman *et al.* 2008). We used these results as the basis for the design of the cloud used in the study presented here (see the section Search interfaces and data set).

Summary

Related work presented in this section aimed to improve interactive information retrieval on the Web by considering a wider range of user factors (such as cognitive abilities) and by incorporating these factors into the design of search results presentation. However, we note that improvements to Web search are not limited to changes at the user

interface. Other work aims to de-couple information retrieval from direct human interaction. An example of one such approach are mobile software agents, which, once they are given a task by a user, asynchronously find, collect and return the collected search results back to that user (Lieberman 1997, Parry 2008, Lam *et al.* 2009)

Our interest in cognitive characteristics of searchers is from the perspective of understanding their effects on searchers' performance and preference for information presentations and, in particular, for search results presentation and browsing. In this section, we have highlighted the role of cognitive characteristics in human-computer interaction and information science and briefly discussed work on displaying search results. The remainder of the paper is structured as follows. The next section describes the study method. This is followed by sections presenting the results and the discussion. The paper ends with conclusions and possible directions for future research.

Method

Twenty three undergraduate students in an Information Technology and Informatics programme (five females and eighteen males) participated in the study conducted in a controlled experimental setting. The study was approved by Rutgers' Institutional Review Board. Students received a partial course credit for their participation.

User tasks

Each participant was given four search tasks. Travel, sightseeing and shopping were selected as everyday search topics familiar to the general public. The tasks were constructed in the spirit of *situated work task situations* (Borlund 2003). Scenarios were used to present realistic situations and provide participants with the search context and the basis for relevance judgments. The tasks were designed to differ in complexity. Simpler tasks involved finding (one) fact that satisfied specified criteria (e.g., name of a hotel located close to an airport). The more complex tasks involved information gathering about several items of interest and selecting those that satisfied several criteria (e.g., finding three museums that collectively carried collections of three different kinds) (Toms *et al.* 2008). All tasks are listed in the Appendix.

The search tasks were performed using two different search interfaces. The interface was switched after the second tasks. Before the first use of each interface, participants performed a training task using that interface. The order of tasks was balanced with respect to task complexity and the search results' interface. Four combinations of two task complexities (Table 1) and two interfaces yielded a total of eight different task and interface rotations. The interfaces are described in the next subsection.

Fact finding	Information gathering	Fact finding	Information gathering
Fact finding	Information gathering	Information gathering	Fact finding
Information gathering	Fact finding	Fact finding	Information gathering
Information gathering	Fact finding	Information gathering	Fact finding

Table 1: Search task order, by complexity

Given our interest in understanding the relationships among search results presentation, search tasks and cognitive abilities, we focused on the participants' interactions with the search results and not their ability to formulate the initial search query. Accordingly, the initial search results page was provided to participants, who continued the process by searching within them and by examining the individual results. Each initial search was performed by using two keywords, one referring to place and one to a topic of interest (e.g., London museums). Since London was used in all scenarios, the term London appeared in all initial queries.

In addition to search tasks, participants performed cognitive tasks. We assessed two cognitive abilities of the study participants: memory span (Francis and Neath 2003) and verbal closure (McGraw *et al.* 1997) (Table 3). These particular cognitive factors were selected as likely to affect the searchers' performance (Gwizdka and Chignell 2004).

In our data analysis, the values of cognitive task performance were split at the median and participants were divided into higher and lower ability groups.

Search interfaces and data set

As described in the Related work section, previous research has explored numerous variations of how to present search results and what document surrogates to use (e.g., Dumais *et al.* 2001, Greene *et al.* 2000, Joho and Jose 2008, Woodruff *et al.* 2002). The focus of our work is not to investigate new presentations, but to explore relationships among cognitive abilities and the presentation of results for different tasks. Hence, we decided to use one interface similar to a traditional textual result list (the *list interface*, see Figure 2) and to use a second interface that differs only in one feature: an added overview of the returned results (the *overview interface* or *list and tag cloud*, see Figure 3). The interfaces were created specifically for this project.

Both interfaces display a textual list of results that shows for each result its surrogate composed of title, URL and a list of descriptive words (Figure 4). Each search result in our system is a Website. The descriptive words are shown in a form of a *document tag cloud*. The overview interface adds an overview of results in a form of an *overview tag cloud*. The *overview tag cloud* contains descriptive words from all returned results. Thus, each interface displays previews of individual results, while the second interface adds an overview of all results (as suggested by Greene *et al.* 2000). The function of a tag cloud as a visual summary was described by Sinclair and Cardew-Hall (2008).
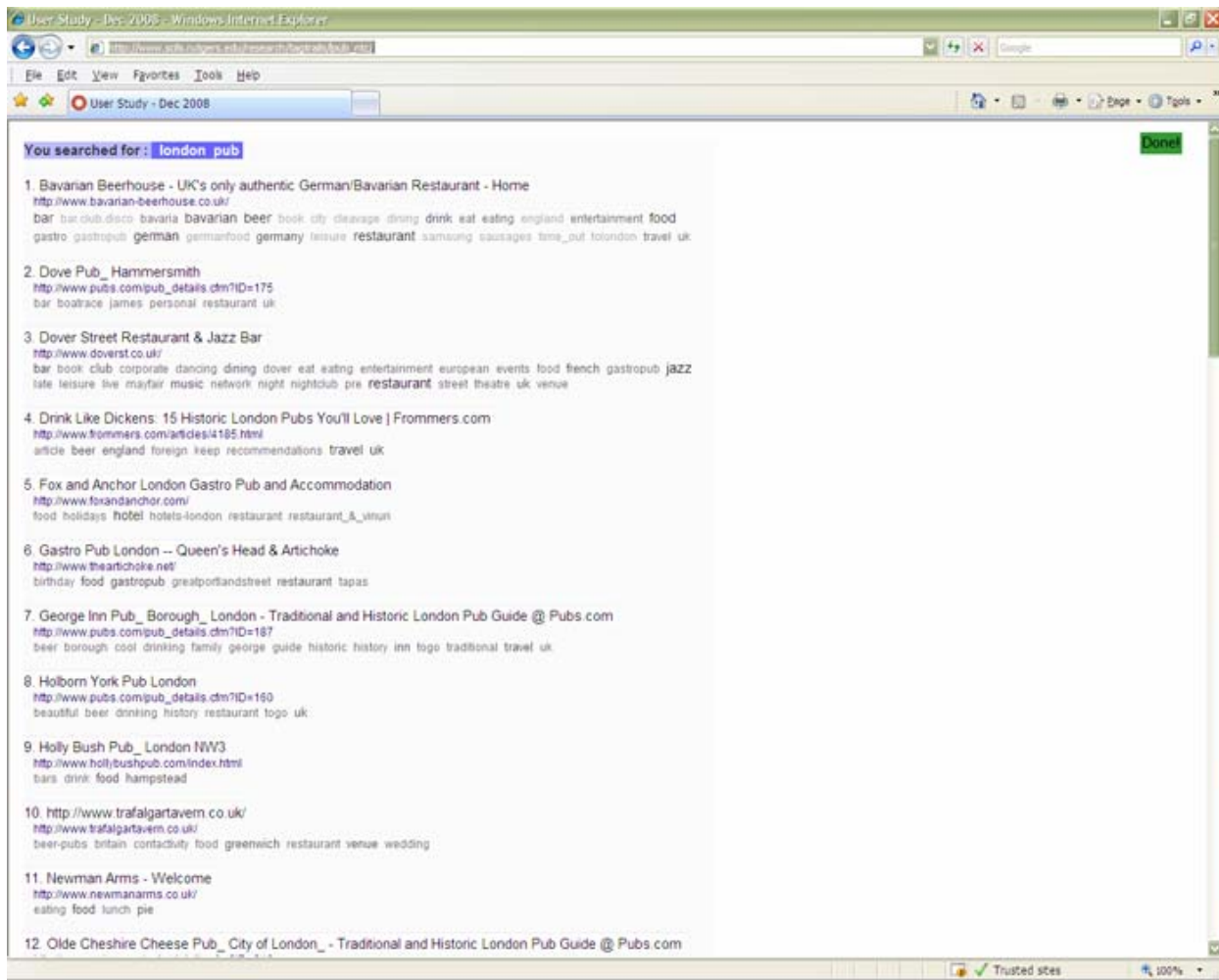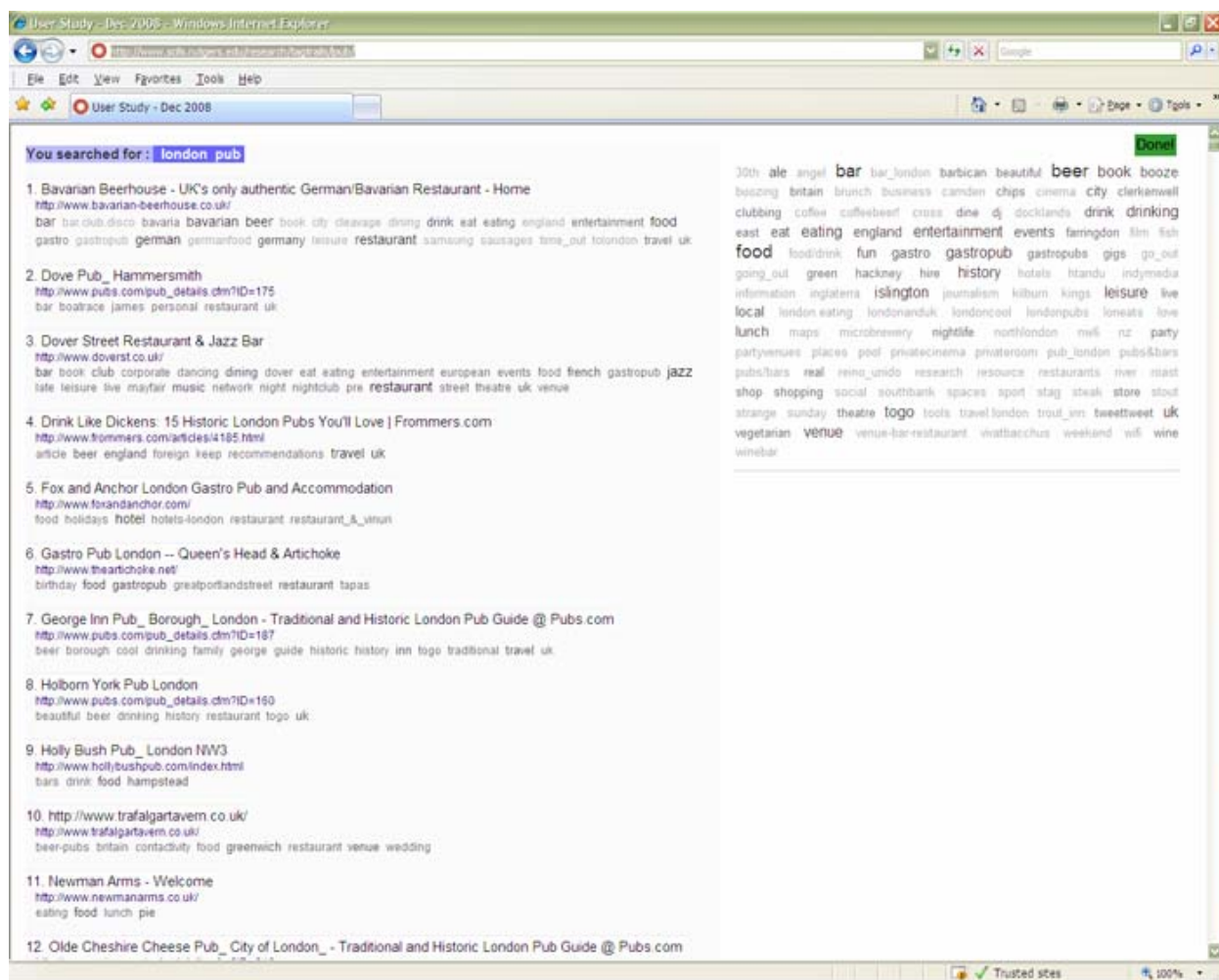

Figure 2: The list interface

Figure 3: The overview interface (list and overview tag cloud)

Henceforth, we will refer to the descriptive words as tags. Tags were displayed in different font sizes. In the overview tag cloud, tag font size reflects the term frequency; in the document tag cloud, tag font size reflects document frequency. The overview tag cloud used six distinct font sizes, while the document tag cloud used three distinct font sizes. In both cases, tags were ordered alphabetically. Tags in all clouds served as links that could be followed to obtain a new view of the associations among the Web resources and tags.
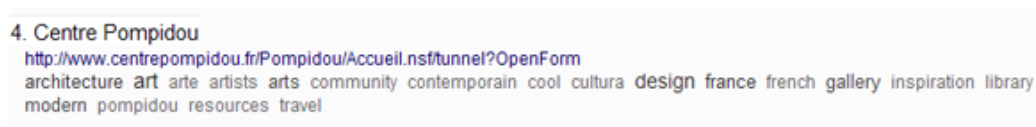

Figure 4. Surrogate for one search result: clickable title, URL and a list of descriptive terms (document tag cloud).

The current query was shown at the top of the interface (Figure 2 and Figure 3). The two search interfaces supported search within the initial set of results. Narrowing down the search was performed by adding terms to the query. A term was added by clicking on a tag (in either document or overview tag cloud). The results were progressively narrowed down as more terms were added to the query. Terms could be removed from a query in any order by clicking on a tag-remove-link, represented by '-' (a minus symbol) shown in the top-right corner of each added term (Figure 5). One could also directly open individual search results (content pages) by clicking on their URLs. When searchers completed each task, they clicked on a *Done* button in the upper right corner of browser window (Figure 2 and Figure 3). The interfaces were written in PHP and optimized for display on screen set to 1280x1024 pixels. The interfaces can be accessed on-line.

You searched for : paris museum culture   contemporary

Figure 5. Interface detail showing query terms.

The data set used in the experiment was obtained by crawling the Delicious social bookmarking site. In November 2008, we collected approximately 18,000 unique bookmarks along with associated tags (tagged by 600,000 users with a total of 380,000 tags). The data were collected on topics related to travel, sightseeing and shopping in several European cities. We used the study task topics to further select approximately 100,000 tagging instances (combinations of unique URL-tag pairs) that were applied to 1,700 bookmarks. Before using the data in the study, they were first cleaned to remove bookmarks to non-existing Web pages and *noisy* tags. Two types of noisy tags were removed: 1) words in non-Latin alphabets and 2) technical tags referring to the source from which bookmarks were imported to Delicious. The latter category includes Web browsers' names (e.g., Firefox, Safari), explicit descriptors of bookmark import activity (e.g., import, export), designations referring to a personal library as the source of bookmarks (e.g., mylibrary, mylinks). The study tasks used data related to London (UK), while the training tasks used data related to Paris. Since the search was performed on a relatively small set of words (i.e., it was not full-text search), we used MySQL as the search engine. To improve system responsiveness, the data were stored in a separate table for each search task.

Our main focus is on the visual presentation of terms in relation to the factors of interest and the source of terms is secondary to our concern. The processes that generated the terms are important and we assume that the words associated by the users of Delicious with the bookmarked Web pages sufficiently describe their content. Our assumption is supported, at least in part, by the work of Golder and Huberman (2006), who analysed user activities on social tagging Websites and described the various reasons that motivate people to enter tags associated with Web pages. Among these reasons are resource re-finding, identification of resource qualities and understanding the content of Web resources.

User actions and search moves

Our search system supported a well-defined set of user actions. Users started from the initial list of search results and their task was focused on finding information within these results. User states and transitions between them are described in Table 2, while Figure 6 shows their state diagram. User actions can be conceptualized as cognitive and physical search moves (Shiri and Revie 2003).

Clicking on tags was conceptually equivalent to adding terms to the search query. Deleting tags was conceptually equivalent to removing terms from the search query. A plausible difference between these two ways of entering query terms is the amount of cognitive load required in these two cases: more, when user types in a query; less, when user scans a tag cloud and clicks on a tag (Sinclair and Cardew-Hall 2008). The number of tags added and deleted by a user corresponded to the number of issued query refinements and, hence, to the number of examined result lists (all results in each list were shown on one Web page). We observed that the participants used tag deletion as both a cognitive move (i.e., when the searcher formulates a new query) and as a physical move (i.e., the searcher goes back to a previous result list). In another study, we observed that searchers used the Backspace key in a similar way. Backspace was used to navigate to a previous Web page (a physical move) and to revise a query by deleting one or more terms (a cognitive move) (Gwizdka 2008, Gwizdka in press). However, based on the data recorded in the current study, we cannot distinguish the cognitive actions from physical ones. Therefore, we use a simple heuristic assuming that a half of the delete tag actions are cognitive and a half are physical. In the Results section, we refer to the number of query re-formulations that was calculated as the number of new tags plus 50% of the deleted tags.

No.	Name	Description of cognitive action	Physical expression
1	*User state*: view result list	Users could use information contained in each individual result surrogate to assess the result's relevance. In the *Overview* interface users could also use the overview	Users are in this state between narrowing down or widening the results and examining the single result Web pages

|			
		cloud to assess all results and to find terms for further query refinement	
2	**User state**: examine one search result	If relevance judgement could not be made based on the result's surrogate, a user could visit the result (Kuo *et al.* 2007).	Users enter this state by clicking the result's URL; users leave this state by clicking the browser's Back button (or using the right-mouse menu to go back); users remain in this state as long as they follow links between external sites
3	**Transition**: narrow down search results	Add a new tag; this action was equivalent to adding new terms to the search query.	Click on a new tag, either in the overview cloud or in of the document clouds
4	**Transition**: widen search results *or* go back to a previous page	Remove a previously added tag; this action was equivalent to removing terms from the search query. **Note**, this action could express either a cognitive or a physical action.	Click on a tag-remove-link.
5	**Transition**: decide when done	Decide that information need has been satisfied and that the answer was found	Click on the *Done* button

Table 2: User states and transitions
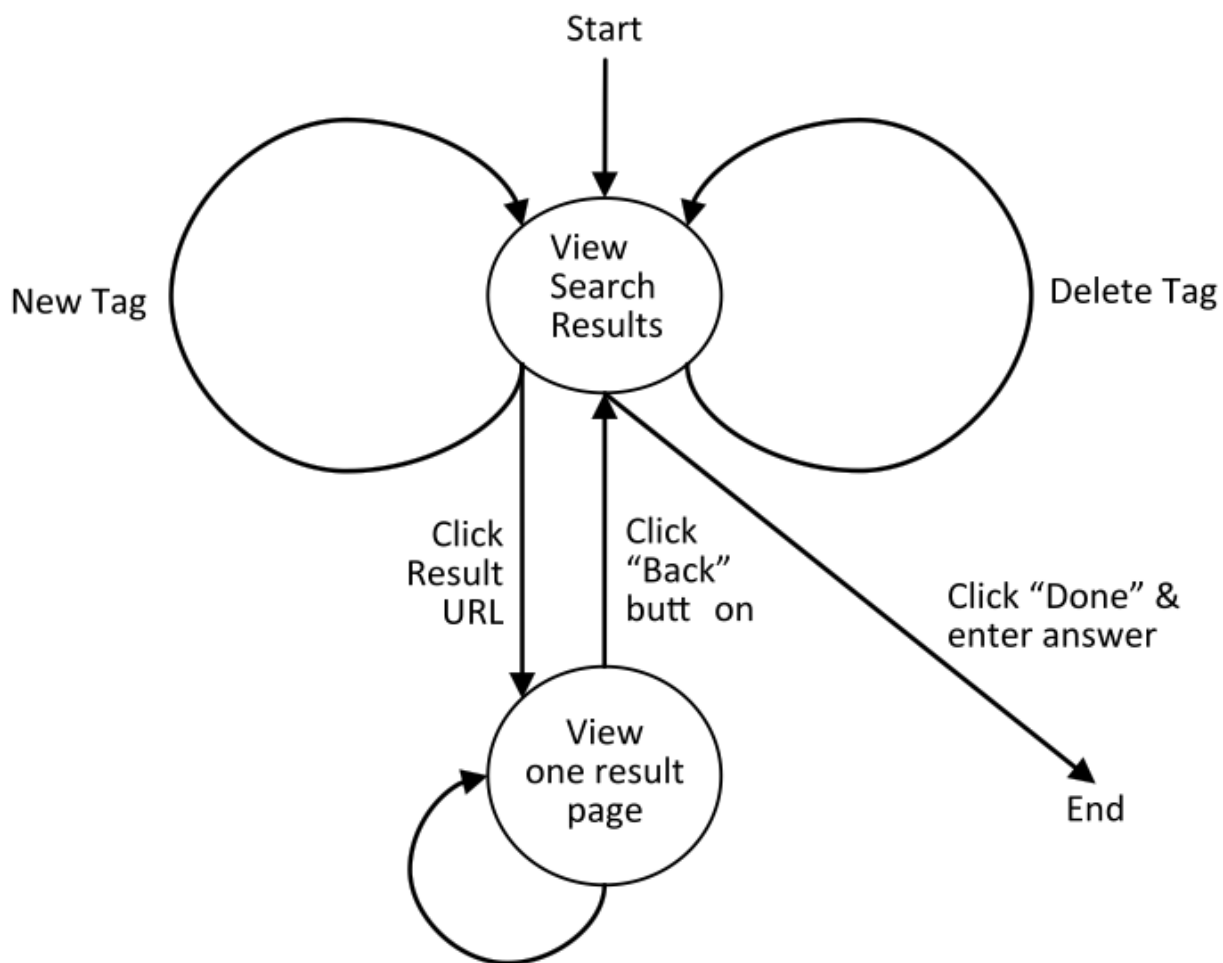
Figure 6: State diagram of user states and transitions

Measures and experiment design

Independent variables

Task complexity and interface type were the controlled, independent variables. Two additional independent variables were the levels of the two selected cognitive abilities (Table 3). Each participant performed tasks at both complexity levels using both interfaces; these were the *within-subject* factors. Cognitive abilities that characterize each individual person were the *between-subject* factors. Thus, the study had a mixed-factorial design.

Cognitive factor	Variable	Test name	Short description	Reference
Memory span (Short-term memory)	Memory span. Higher score = higher ability	CogLab on CD (Wadsworth)	The ability immediately to recall digits, letters, or other items.	(Francis and Neath 2003)
Verbal closure	CV: A measure of verbal closure. Higher values = higher ability	Closure, Verbal. CV-2 (ETS)	The ability to identify visually presented words when some letters are missing, scrambled, or embedded among other letters.	(Ekstrom *et al.* 1976)

Table 3: Cognitive abilities measured in the study

Dependent variables

Task performance was assessed by using search efficiency, search effort and search task outcomes (Table 4). Search efficiency was measured by the task completion time, while search effort was measured by counting the earlier described user actions. Search effort was expressed in multiple variables: the number of tags added and removed, the number of times different types of clouds were clicked (document vs. overview) and the number of times single search results were visited. These measures were obtained from the local Web server logs. As we noted above, adding and deleting tags was equivalent to query reformulation. Search task outcome was measured by assessing the relevance (on a scale from 1-not relevant to 5-relevant) and completeness (0.0 No items found; 0.33: 1 out of 3 found; 0.67: 2 out 3 found; 1: all requested items found) of the user-saved results by a human expert. The outcome was calculated as relevance multiplied by completeness (thus the search task outcome measure can range from 0-minumum to 5-maximum).

Variable type	Variable name	Description
Search efficiency	Time	Task completion time
	URLTagsRatio	The ratio between the visited content pages (URLs) and the clicked tags (new tags) = (URLs+1)/new tags
Search effort	New tags	The number of tags that were added (clicked)
	Deleted tags	The number of tags that were deleted (removed)
	Queries	New tags + 0.5 * deleted tags
	Document clouds	The number of tags clicked in one of the document clouds
	Overview cloud	The number of tags clicked in the overview cloud
	otherUrlCount	The number of times single search results were visited
Search outcome	taskOutcome	Relevance * completeness

Table 4: Dependent variables

An additional derived variable was constructed to capture the ratio between the visited content pages (URLs) and the clicked tags (new tags). Since in 98% of task instances (90 out of 92 cases), participants started their searches by clicking on a tag, we decided to give a small advantage to URLs over tags and we calculated the ratio of URLs to tags in the following way:

$$URLTagRatio = (URLs+1)/newTags$$

Results

The experiment had a mixed model design, with two within-subject factors (task and interface) and two between-subject factors (the two cognitive abilities: verbal closure and memory span). The underlying mixed model, also called univariate repeated measures model, assumes that the dependent variables are responses to levels of the within-subject factors. We performed a multivariate analysis of variance with repeated measures (as implemented by GLM procedure in SPSS 16). In addition, we ran separate t-tests and non-parametric tests for selected variables.

Interface and task effects

There was a significant main effect of user interface ($F_{(6,15)}=50.63$, $p<0.001$). Within-subject contrasts reveal that

interface had a significant effect on time and tag cloud use. The effect of tag cloud use was expected, as it is a product of the interface design. Hence, we will not be discussing it further. The Overview interface was faster than the List interface by 64 seconds (Table 5).

Dependent measures (mean values)	User interface (UI)		Statistics (within-subject contrasts)
	List	Overview (list + tags)	
Time (seconds)	239	175	$F(1,20)=7.42$, $p=0.013$
Document clouds	4.22	0.15	$F(1,20)=125.34$, $p<0.001$
Overview cloud	n/a (no overview cloud)	3.72	$F(1,20)=84.21$, $p<0.001$

Table 5: Effects of user interface

There was a significant main effect of task ($F(6,15)=3.5$, $p=0.022$). Within-subject contrasts show that task had a significant effect on time, number of new and deleted tags and overview tag cloud use. Generally, more complex tasks required more effort and more time. On the average, participants used three more queries in more complex tasks than in simple tasks (Table 6).

Dependent measures (mean values)	Tasks		Statistics (within-subject contrasts)
	Simple task (fact finding)	Complex task (information gathering)	
Time (seconds)	171	242	$F(1,20)=5.5$, $p=0.029$
New tags	3	5.1	$F(1,20)=9.6$, $p=0.002$
Deleted tags	1.5	3.5	$F(1,20)=23.1$, $p<0.001$
Queries	3.8	6.9	n/a

Table 6: Effects of task type

There was a significant interaction effect f user interface and task ($F(6,15)=6$, $p=0.002$). Within-subject contrasts show a significant effect of task * interface on the use of the overview tag cloud ($F(1,20)=23.1$, $p<0.001$). This effect is, in part, a result of differences in the interface (the List interface did not have the overview tag cloud) and, in part, a result of the task (the complex task required searchers to issue more queries and view more result lists).

We then examined the interfaces and tasks separately. Within each interface we considered the effects of tasks and within each task the effects of the interface. We note that searcher performance between subsequent tasks and between interfaces is essentially independent. Time on a task and the number of tags significantly differed between the two types of tasks only in the Overview interface. Simpler tasks were faster and involved less effort (i.e., fewer tags clicked, less reformulation of the query) (Table 7). The direction of these effects is the same as for the main task effect (Table 6). Since there were no such differences in the List interface, we infer that the main task effect is mainly due to how the Overview interface influences user performance.

Dependent measures (mean values)	Tasks		Statistics (between task conditions)
	Simple task	Complex Task (information	

	(fact finding)	gathering)	
Time (seconds)	124	225	t(42.7)=-2.67, p=0.011
New tags	2.5	5.3	t(41.9)=-4.1, p<0.001
Deleted tags	1.1	3.6	t(43.1)=-3.41, p=0.001
Queries	3	7	n/a

Table 7:
Effects of task type in the Overview interface

Examining performance within each task, we found that significant differences between the List and Overview interfaces were limited to time on simple tasks. The Overview interface (124s) was faster than the List interface (225s) and the gain in efficiency came from performance of simple tasks (Table 8).

User interface	Tasks				Statistics (between task conditions)
	Simple task (fact finding)		Complex task (information gathering)		
	Time (seconds)	Queries	Time	Queries	
List	218	4.4	259	6.7	not significant (n/s)
Overview (list +tags)	124	3	225	7	see Table 7
Statistics (between user interface conditions)	t(43.3)=2.13, p=0.039		n/s		

Table 8: Task by interface results
(Note: Emphasised cells show values that differed significantly.)

Cognitive ability effects

There were no main effects of cognitive abilities. The interactions between cognitive abilities and within-subject variables (task and interface) were also not significant. However, the effect sizes (as measured by partial Eta squared) of these interactions were relatively large (they corresponded to explaining 25% to 45% of variation). Furthermore, several of the within-subject univariate tests (the within-subject contrasts) were significant. Since our examination of cognitive ability effects had an exploratory character, we considered the univariate tests and we report significant as well as a few borderline significant results. The borderline significant result do not satisfy the p<0.05 criteria; their p values can be in the range of 0.05<0.095. To further examine the nature of relationships between cognitive abilities and task and between cognitive abilities and interface, we examined them separately.

In general, the findings indicate that the verbal closure ability tended to interact with user interface conditions (significant differences found in the List interface), while the memory span ability tended to interact with task types (significant differences found on complex tasks).

There were significant differences in performance between high and low memory span people on complex tasks, while there were no such differences on simple tasks. For complex tasks the number of new tags was borderline significant and the number of deleted tags was significant. People low on memory span issued in complex tasks on average 3.5 queries, while those with high on memory span issued approximately eight queries (Table 9).

	Memory span	

Dependent measures (mean values)	Low	High	Statistics (between memory span levels)
New tags	2.2	5.8	t(43.6)=-1.95, p=0.058 (borderline)
Deleted tags	2.6	4.3	t(42.4)=-2.2, p=0.037
Queries	3.5	8	n/a

Table 9: Differences between high and low memory span on complex tasks

When we split participants according to the two levels of memory span, we found significant differences in performance between simple and complex tasks only for people high on memory span. Time on task and the numbers of new and deleted tags were significantly different. On simple tasks, high memory span people issued on average 3.5 queries, while on complex tasks they issued approximately eight queries. Compared with low memory span people, high memory span people were more efficient on simple tasks and performed better on complex tasks, in a sense of exploring more result pages, while not being significantly slower (Table 10).

Task	Memory span			
	Low		High	
	Time (seconds)	Queries	Time	Queries
Simple task (fact finding)	208	4.1	143	3.5
Complex task (information gathering)	222	5.5	258	8
Statistics (between task conditions)	n/s	n/s	t(40.9)=-2.94, p=0.005	new tags: t(45.8)=-4.0, p<0.001 del. tags: t(46.3)=-3.56, p=0.001

Table 10: Summary of results: Task by memory span
(Note: emphasised cells show values that differed significantly.)

Examining the relationships between the two user interface conditions and cognitive abilities we found that people high on memory span were significantly faster in Overview than in List interface. There was no such difference for low memory span people (Table 11).

User interface	Memory span	
	Low	High
	Time (seconds)	Time (seconds)
List	235	243
Overview (list + tags)	195	159
Statistics (between user interface conditions)	n/s	t(41.8)=2.05, p=0.047

Table 11: Summary of results: user interface x memory span
(Note: emphasised cell shows value that differed significantly.)

We found significant differences in the List interface between people with high and low verbal closure. The number of new and deleted tags differed. In List interface low verbal closure people used four queries, while high verbal closure used seven (Table 12). This difference between low and high disappeared in the Overview interface. Both groups used about the same number of queries (about five on the average). The number of queries used by both groups in the Overview interface was similar to the number of queries used by low verbal closure people in the List

interface.

Dependent measures (mean values)	Verbal closure		Statistics (between verbal closure levels)
	Low	High	
New tags	3.2	5.1	t(43.5)=-2.16, p=0.037
Deleted tags	1.7	3.6	t(40.6)=-2.2, p=0.034
Queries	4	7	n/a

Table 12: Differences between high and low verbal closure in the List interface

We examined closer the differences in performance of high and low verbal closure people between the two user interface conditions. There were no differences for people low on verbal closure, however we found significant difference in time for people high on verbal closure. There were also borderline significant differences for the number of new tags and the number of deleted tags (Table 13). High verbal closure people issued about seven queries in List interface and 4.5 in Overview interface. Hence, they were more efficient in Overview interface and performed better in List interface (in a sense of exploring more result pages, while not being significantly slower).

User interface	Verbal closure			
	Low		High	
	Time (seconds)	Queries	Time (seconds)	Queries
List	238	4	240	7
Overview (list + tags)	206	5.5	146	4.5
Statistics (between UI conditions)	n/s	n/s	t(39)=2.22; p=0.032	new tags: t (42.1) = 1.96, p = 0.056; del. tags: t (42.6) = 1.74, p = 0.088 (both borderline)

Table 13: Summary of results: user interface x verbal closure
(Note: emphasised cells show values that differed significantly.)

Task outcome

Most participants succeeded in the sense that they found information that was requested by each task scenario. The average search task outcome was 4.5 (on a rating scale of 0-minumum to 5-maximum, see the Dependent variables section). Participants achieved maximum score (5.0) on 78% of tasks, while they achieved less than or equal to the neutral score (3.0) only on 12% of tasks. Hence the distribution of task outcome was highly skewed and we used non-parametric statistics to test it. There was a borderline significant effect of task type on outcome as shown by non-parametric Mann-Whitney test (U=885.0, Z=-1.9, p=0.061). Task outcome differed between task types in the expected direction; more difficult tasks were associated with worse outcomes, however the difference was only 0.5 points (on average participants achieved 4.8 for simple tasks and 4.3 for complex tasks). The difference between tasks was significant in the Overview interface (Mann-Whitney test: U=187.5, Z=-2.3, p=0.019; for simple tasks 4.2 vs. 4.9 for complex tasks), while there was no difference in the List interface.

It appears that performance on tasks of different complexity was more differentiated in the Overview interface. Although not statistically significant, the tendency observed in the interaction between interface and memory span is worth mentioning. Low memory span people tended to have worse outcome in List than in Overview (4.2 and 4.7 respectively), while the difference was smaller for high memory span people (4.7 and 4.5 respectively). The Overview interface helped people low on memory span to achieve levels of performance closer to the high memory

span group, while the List interface differentiated more between people characterized by both lower and higher levels of memory span.

Use of result lists vs. content pages

We sought to explain further search behaviour by comparing the number of viewed result lists (i.e., the number of clicked tags) to the number of visited content pages (i.e., the number of individual items from the result lists). As described in the Method section, we calculated the ratio between the visited content pages and the clicked tags (URLTagRatio). The number of clicked tags corresponded to the number of issued queries (for more detailed explanation please see the User actions and search moves section), which is the same as the number of examined result lists. As a shortcut, we will refer to the visited content pages as URLs.

Overall, searchers clicked about two times more often on tags than on URLs. There were no significant differences between the two interfaces. However, the ratio of clicked URLs to tags was slightly higher in Overview than in List interface. On complex tasks, searchers clicked about five times more frequently on tags than on URLs. On simple tasks, they clicked about the same number of tags and URLs. For simple tasks in List interface, the number of tags and URLs was about equal. However, for simple tasks the URL to tag ratio was higher in the Overview interface than in the List interface. Possibly, simple search tasks in the Overview interface did not require trying many different queries and viewing many result lists, but, instead, were satisfied by visiting more content pages. In the Overview interface, searchers clicked more tags on complex tasks than on simple tasks, which is consistent with the main effect of task on URLTagRatio.

For complex tasks, the effects of cognitive abilities were not significant. However, for simple tasks, the levels of cognitive abilities made a difference. For low levels of both verbal closure and memory span, searchers clicked about the same number of tags as they did URLs. In contrast, for high levels of cognitive abilities, the differences between user interface conditions were, in most cases, more pronounced. High verbal closure and high memory span people clicked about the same number of tags and URLs in the List interface, while they used more URLs than tags in Overview interface. Detailed results of this analysis are shown in Appendix 2.

Use of specific tags

We examined the use of specific tags for one of the simple tasks (Task id: FF_1 – see Appendix 1). This task was selected because we could expect that searchers would use one of few specific tags (e.g., vegetarian). We found that the use of the expected tags varied between the two interfaces. In the List interface half of participants (7/14) used the expected tags, while in the Overview interface almost all participants (8/9) used these tags. This difference seems to reflect the interface effect. It indicates, again, the advantage of the Overview interface over the List interface. It was easier to find target words in the overview cloud than to find these words in the individual document clouds. The data used in this analysis was very sparse; hence the power of conclusions that we could draw from it is limited.

Discussion

Our main focus is on the relationships between the users' cognitive abilities, the search interface and task. In this section, we present a summary of the results and discuss their implications.

In the Overview interface the overview tag cloud was the prevalent method of adding terms to the query, while the document tag clouds remained virtually unused in this condition. In the List interface document tag clouds were the only mechanism available for adding terms. Overall, the Overview interface was faster. It seems that the Overview helped users to be more efficient. The difference in speed between the interface conditions was significant for simpler tasks, while it was not significant for complex tasks. This finding bears some similarity to the results described by Kuo *et al.* (2007), who found that a tag cloud that provided summary of Web search results was better (faster and more accurate) than standard PubMed interface for finding descriptive information (simpler task) than for identifying relationships among multiple concepts (more complex task).

As we had expected, the simpler tasks were overall faster than the more complex tasks. More interestingly, the user's performance was more differentiated in the Overview than in the List interface. In List interface performance on both task types was about the same. In Overview interface the simpler tasks required less effort and time. Apparently, the

overview tag cloud made simpler task even easier. A similar kind of effect on performance time of the interaction between interface and task was observed by Dumais *et al.* ([2001](#)). They found that for difficult tasks (difficult queries) the result list interface was much slower than the category interface, while for easy tasks there was little effect attributable to the interface.

Cognitive abilities affected search behaviour in some interface and task conditions; they did not seem to influence behaviour independently from the interface nor task. The verbal closure ability tended to interact with the user interface, while the memory span ability tended to interact with the task type.

Cognitive abilities and user interface

The differences in behaviour and in performance of searchers characterized by the two levels of cognitive abilities (memory span and verbal closure) were larger in the List than in the Overview interface. The user interface mainly affected people with different levels of verbal closure, but it also affected people with different levels of memory span with respect to one dependent variable (time on task). People high on memory span were significantly faster in Overview than in List interface. For the searcher groups characterized by different levels of verbal closure, the nature of the difference depended on the interface. In the List interface, it was reflected in the number of added and removed tags (i.e., the number of query reformulations). Searchers high on verbal closure used more queries (2.5 more) than those low on verbal closure. There was no such difference in terms of queries between low and high verbal closure people in the Overview interface. On the other hand, high verbal closure people were significantly faster (146 seconds) in the Overview than they were in the List interface (240 seconds). Thus, the expected advantage of high verbal closure people manifested differently between the varied interface conditions. The List interface required more perceptual and cognitive processing than the Overview interface. More frequent use of tags by high verbal closure compared to low verbal closure people in this interface condition could possibly be explained by their better ability to locate perceptually and to extract words from the search results display. Given that there was no significant difference in performance time in List interface between these groups, we can say that high verbal closure people were more efficient than low verbal closure in terms of the number of queries issued per unit of time. However, since the overall time to complete a task and the task outcomes were about the same for both user groups, one could also consider the increased effort of high verbal closure people in the List interface as *wasted*. In contrast in the Overview interface, the better ability of high verbal closure people (and also of high memory span people) resulted in their faster performance at the same level of effort invested in the task as by the low verbal closure people. In this condition, higher cognitive ability gave a clear advantage to searchers.

Cognitive abilities and task types

The task type did not differentiate performance for the levels of verbal closure ability. In contrast, behaviour and performance of people characterized by the two levels of memory span differed on complex tasks. The number of queries issued was significantly different. High memory span people used on average eight queries while than those low on memory span only 3.5. Examining these two groups separately, we found a significant difference between tasks for high memory span people only. They used approximately 4.5 queries more on complex tasks than on simple tasks and, at the same time, they were slower on complex tasks (258 seconds) than on simpler tasks (143 seconds). These kinds of differences were not significant for low memory span people, who used about the same number of queries in both tasks (between 4.1 and 5.5) and spent on the average 215 seconds on a task. In summary, on simple tasks people low on memory span took more time and used about the same number of queries as high memory span people. On complex tasks, however, low memory span people used fewer queries than high memory span people, while spending about the same amount of time. We can conclude that high memory span people tended to be overall more efficient than low memory span people in a sense of the number of queries issued and results pages examined per unit of time.

The pattern of interactions between task effects and memory span ability is similar to that of interactions between user interface conditions and verbal closure abilities. The increased effort of higher ability people in more demanding situations (complex tasks and List interface) did not result in their better performance and thus could be considered as *wasted*. In less demanding situations (simple tasks and Overview interface), however, higher cognitive abilities resulted in faster performance at the same level of effort invested and thus gave a clear advantage to this group of searchers. The increased effort invested in search by the high cognitive ability people may reflect an extra cognitive capacity ([Wickens 2002](#)) that this group of people possesses and a lower effective cognitive load ([Sweller *et al.* 1998](#), [Gwizdka 2008](#), [Gwizdka in press](#)) imposed on these people by more demanding situations. Since it was not the case

that people low on memory span had significantly lower task outcomes, the interpretation of the lower number of queries issued per unit of time by this group of people remains open and we cannot say that they issued too few *necessary* queries. Based on the collected data, we cannot answer what the high memory span people gained on complex tasks or what the high verbal closure people gained in the List interface by using more queries and by examining more search result lists.

The next aspect of the searchers' behaviour that we studied was a comparison of the ratio of the number of viewed result lists (i.e., the clicked tags) to the number of visited content pages (individual results). The factors that affected this ratio depended on the type of task. For more complex tasks, independent of the interface and of the cognitive abilities, more result lists were examined than individual results. This could be explained, in part, by the task design. But it may also suggest that because more queries were tried and more results list examined, the searchers were less likely to invest more effort in visiting content Web pages. More interestingly, for simpler tasks the interface type made a difference. In particular, this difference was observed for users characterized by higher levels of cognitive abilities on simpler tasks performed in the Overview interface. These searchers visited more individual results than result lists than they did in the List interface and than did the searchers with low levels of cognitive abilities. The smaller number of queries (and thus the examined result lists) in the Overview as compared to the List interface can be plausibly explained by the role the overview tag cloud played in the query formulation and in the assessment of the search results list. First, it is possible that better visibility of terms available for query refinement in the Overview supported reformulation of better queries. Second, the summary function of the Overview allowed searchers to assess search result lists better and consequently they could examine fewer result lists.

We did not observe a speed-accuracy trade-off in this study. The differences in behaviour were mainly observed in effort and time on task. Task outcomes varied little and were mostly good to very good.

Conclusions and future outlook

We set out to investigate the relationships between the users' cognitive abilities, the search interface and the task. Our goal was to understand better the effects of the searchers' cognitive characteristics on their performance and preference for presentation and browsing of search results. The underlying objective of this line of research is to understand how to match search result presentations with different tasks and user contexts, while at the same time minimizing the cognitive load imposed by the search system on its users.

Searchers seemed to have benefited from the overview of search results in several ways. The overview enable them to complete tasks faster. However, we observed that the increase in efficiency was mainly the result of the faster performance on simpler tasks in the Overview interface. The differences in performance in the List interface were minor. The Overview interface allowed users to be more efficient in terms of the number of queries used. The smaller proportion of the number of queries (and, thus, the examined result lists) to the number of visited content pages in the Overview interface led us to conjecture two more benefits of the results overview: 1) overview makes query refinement easier and 2) overview better supports assessment of the search result set. As a result searchers became more effective and examined fewer result lists.

Cognitive abilities made a difference. The difference depended on the type of the interface and on the type of task. The Overview interface helped searchers with higher levels of cognitive abilities to become faster. This finding demonstrates another benefit of the search results overview. The List interface did not significantly differentiate performance between the two task types. However, the List interface differentiated between the user groups characterized by different levels of cognitive abilities. The importance of this finding comes from the fact that the List interface was similar to a typical Web search engine results interface. It suggests a significant role that cognitive abilities play in user interaction with search results.

Searchers with higher cognitive abilities expended more search effort in more demanding situations (in List interface and on more complex tasks). However, an improvement of task outcomes was not observed. It is thus unclear whether the higher levels of cognitive abilities brought any advantage to searchers in these task/interface combinations. One could speculate whether the higher effort invested by these users in search might have led them to some other benefits. Perhaps they gained more knowledge that was not immediately needed for the search tasks and that was not measured in this study. It is also possible that the higher cognitive capacity allowed these searchers to take advantage of serendipitous information encounters. The interaction data that has been collected cannot answer

these questions and it does not allow us to identify all cognitive moves performed by the searchers. We plan to examine these aspects in future studies by using a wider range of interaction data. In particular, we plan to combine mouse clicks with eye-tracking data.

The limitations of the current study include a homogenous participant group, a data set that contained socially-generated descriptive terms, a limited range of task topics and complexity and task performance in the laboratory environment. In particular, participants' relatively young age and their frequent use of Web search could have influenced, at least in part, the results. However, it is worth noting that others found similar effects of using an overview tag cloud. In one case, Kuo *et al.* (2007) found similar effects of an overview tag cloud created from PubMed metadata terms in a study that used a different population. In another case, Sinclair and Cardew-Hall (2008) found that a tag cloud served as a visual summary of a document collection and made the search task easier by giving users starting points. Hence, we believe there is a possibility of generalizing our results beyond the current study conditions. In the future, we plan to introduce more variety among participants and search tasks. Additionally, we plan to explore how multiple tag clouds and a cloud history could be used to support information search and browsing (Gwizdka and Bakelaar 2009); the prototype interface (called Tag Trails) is available on-line.

Our findings have implications for the design of search interfaces. They suggest benefits of providing result overviews. They also suggest the importance of considering cognitive abilities in the design of search results presentation and interaction.

Acknowledgements

About the author

Jacek Gwizdka is an Assistant Professor in the Department of Library and Information Science, in the School of Communication and Information at Rutgers, the State University of New Jersey, USA. He received his Master of Engineering degree in Electrical Engineering from the Technical University of Łodż, Poland, Master of Applied Science and PhD in Industrial Engineering from the University of Toronto, Canada. He can be contacted at informationr.net@gwizdka.com

References

- Allen, B. (2000). Individual differences and the conundrums of user-centered design: two experiments. *Journal of the American Society for Information Science*, **51**(6), 508-520.
- Baddley, A. (1986). *Working memory*. Oxford: Clarendon Press.
- Bateman, S., Gutwin, C. & Nacenta, M.A. (2008). Seeing things in the clouds: the effect of visual features on tag cloud selections. In P. Brusilovsky and H.C. Davis (Eds.), *Hypertext 2008: Proceedings of the 19th ACM Conference on Hypertext and Hypermedia Pittsburg, PA, USA, June 19-21, 2008*. (pp. 193-202). New York, NY: ACM Press. Retrieved 31, October 2009 from http://hci.usask.ca/publications/2008/fp039-bateman.pdf (Archived by WebCite® at http://www.webcitation.org/5kwguG7KB)
- Belkin, N.J. (1980). Anomalous states of knowledge as a basis for information retrieval. *Canadian Journal of Information Science*, **6**(5), 133-143.
- Belkin, N.J. (2008). Some(what) grand challenges for information retrieval. *SIGIR Forum*, **42**(1), 47-54.
- Bhavnani, S. K. (2001). *Important cognitive components of domain-specific search knowledge*. Paper

presented at *TREC'2001: the Tenth Text REtrieval Conference*. Retrieved 1 March, 2009 from http://trec.nist.gov/pubs/trec10/papers/TREC-2001-04.pdf (Archived by WebCite® at http://www.webcitation.org/5kunWWLIv)

- Borgman, C. L. (1989). All users of information retrieval systems are not created equal: an exploration into individual differences. *Information Processing and Management, 25*(3), 237-251.
- Borlund, P. (2003). The IIR evaluation model: a framework for evaluation of interactive information retrieval systems. *Information Research*, **8**(3), paper 152. Retrieved 1 March, 2009 from http://informationr.net/ir/8-3/paper152.html (Archived by WebCite® at http://www.webcitation.org/5kunkcfpC)
- Buntine, W.L., Valtonen, K. & Taylor, M.P. (2005). *The ALVIS document model for a semantic search engine.* Paper presented at the 2nd Annual European Semantic Web Conference (ESWC'05), Heraklion, Crete. Retrieved 1 March, 2009 from http://cosco.hiit.fi/Articles/eswc_buntine.pdf (Archived by WebCite® at http://www.webcitation.org/5kunt2FNq)
- Card, S.K., Moran, T.P. & Newell, A. (Eds.). (1983). *The psychology of human-computer interaction.* Hillsdale, NJ: Lawrence Erlbaum.
- Carroll, J.B. (1993). *Human cognitive abilities: a survey of factor analytic studies.* Cambridge: Cambridge University Press.
- Chen, S.Y. & Ford, N. (1998). Modelling user navigation behaviours in a hypermedia-based learning system: an individual differences approach. *Knowledge Organization*, *25*(3), 67-78.
- Chen, S.Y., Magoulas, G. D. & Dimakopoulos, D. (2005). A flexible interface design for web directories to accommodate different cognitive styles. *Journal of the American Society for Information Science and Technology*, **56**(1), 70-83.
- Chen, S.Y., Magoulas, G.D. & Macredie, R.D. (2004). Cognitive styles and users' responses to structured information representation. *International Journal on Digital Libraries*, *4*(2), 93-107.
- Duggan, G. B. & Payne, S. J. (2008). Knowledge in the head and on the Web: using topic expertise to aid search. In *Proceeding of the twenty-sixth annual SIGCHI conference on Human Factors in Computing Systems, Florence, Italy, April 05-10, 2008.* (pp. 39-48). New York, NY: ACM Press.
- Dumais, S., Cutrell, E. & Chen, H. (2001). Optimizing search by showing results in context. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems: CHI'08, Florence, Italy, April 05-10, 2008.* (pp. 277-284). New York, NY: ACM Press.
- Dziadosz, S. & Chandrasekar, R. (2002). Do thumbnail previews help users make better relevance decisions about Web search results? In *Proceedings of the 25th annual international ACM SIGIR conference on Research and Development in Information Retrieval, Tampere, Finland.* (pp. 365-366). New York, NY: ACM Press.
- Egan, D.E. (1988). Individual differences in human-computer interaction. In M. Helander (Ed.), *Handbook of human-computer interaction*. (pp. 543-568). New York, NY: North-Holland.
- Ekstrom,R. B. (1976). *Manual for kit of factor-referenced cognitive tests, 1976*. Princeton, NJ: Education Testing Service
- English, J., Hearst, M., Sinha, R., Swearingen, K. and Yee, K. (2002). Hierarchical faceted metadata in site search interfaces. In *Extended Abstracts on Human Factors in Computing Systems (Minneapolis, Minnesota, USA, April 20 - 25, 2002). CHI '02* (pp. 628-639). New York, NY: ACM Press.
- Fallows, D. (2008). *Search engine use*. Washington, DC: Pew Internet & American Life Project. Retrieved 31 October, 2009 from http://www.pewinternet.org/Reports/2008/Search-Engine-Use.aspx?r=1 (Archived by WebCite® at http://www.webcitation.org/5kwhyYCMg)
- Ford, N. (1986). Pyschological determinants of information needs. *Journal of Librarianship*, **18**(1), 47-62
- Ford, N. and Chen, S. Y. (2000) Individual differences, hypermedia navigation and learning: an empirical study. *Journal of Educational Multimedia and Hypermedia*, *9*(4), 281-312.
- Ford, N., Miller, D. & Moss, N. (2001). The role of individual differences in Internet searching: an empirical study. *Journal of the American Society for Information Science and Technology, 52*(12), 1049-1066.
- Ford, N., Miller, D. & Moss, N. (2005a). Web search strategies and human individual differences: a combined analysis. *Journal of the American Society for Information Science and Technology*, **56**(7), 757-764
- Ford, N., Miller, D. & Moss, N. (2005b). Web search strategies and human individual differences: cognitive and demographic factors, Internet attitudes and approaches. *Journal of the American Society*

for Information Science and Technology, **56**(7), 741-756

- Francis, G. & Neath, I. (2003). *CogLab on a CD.* 3rd ed. Boston, MA: Wadsworth Publishing.
- Golder, S.A. & Huberman, B.A. (2006). Usage patterns of collaborative tagging systems. *Journal of Information Science*, **32**(2), 198-208.
- Greene, S., Marchionini, G., Plaisant, C. & Shneiderman, B. (2000). Previews and overviews in digital libraries: designing surrogates to support visual information seeking. *Journal of the American Society for Information Science*, **51**(4), 380-393.
- Gwizdka, J. (2008). Cognitive load on Web search tasks. In JoséJ. Cañas & Herre Van Oostendorp (Eds.) *Proceedings of Workshop on Cognition and the Web, Information Processing, Comprehension and Learning, Granada, 24-26 April 2008.* (pp. 83-86). Granada, Spain: University of Granada
- Gwizdka, J. (in press). Assessing cognitive load on Web search tasks. To appear in *The Ergonomics Open Journal*.
- Gwizdka, J. & Bakelaar, P. (2009). Tag trails: navigating with context and history. *In Proceedings of the 27th international Conference Extended Abstracts on Human Factors in Computing Systems (Boston, MA, USA, April 04 - 09, 2009). CHI EA '09.* (pp. 4579-4584). New York, NY: ACM Press.
- Gwizdka, J. & Chignell. M.H. (2004). Individual differences and task-based user interface evaluation: a case study of pending tasks in e-mail. *Interacting with Computers*, **16**(4), 769-797.
- Gwizdka, J. & Cole, M. (2007). Finding it on Google, finding it on del.icio.us. In L. Kovács, N. Fuhr & C. Meghini (Eds.), *Research and advanced technology for digital libraries* (pp. 559–562). Berlin and Heidelberg: Springer-Verlag. (Lecture Notes in Computer Science LN4675)
- Halvey, M. & Keane, M. T. (2007). *An assessment of tag presentation techniques.* Paper presented at the 16th International World Wide Web Conference, WWW2007, May 8-12, 2007. Retrieved 21 July 2009 from http://www2007.org/htmlposters/poster988/ (Archived by WebCite® at http://www.webcitation.org/5kurART33)
- Harper, S., Michailidou, E. & Stevens, R. (2009). Toward a definition of visual complexity as an implicit measure of cognitive load. *ACM Transactions on Applied Perception (TAP)*, **6**(2), 1-18.
- Harrison, A.W. & Rainer, R.K. (1992). The influence of individual differences on skill in end-user computing. *Journal of Management Information Systems*, **9**(2), 93–111.
- Hearst, M. A. (1995). TileBars: visualization of term distribution information in full text information access. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (Denver, CO) CHI'95*, (pp. 59-66). New York, NY: ACM Press.
- Ingwersen, P. (1996). Cognitive perspectives of information retrieval interaction: elements of a cognitive IR theory. *Journal of Documentation*, **52**(1), 3-50
- Ingwersen, P. & Järvelin, K. (2005). *The turn: integration of information seeking and retrieval in context*. Dordrecht, Netherlands: Springer Verlag.
- Joho, H. & Jose, J. M. (2008). Effectiveness of additional representations for the search result presentation on the Web. *Information Processing & Management*, **44**(1), 226-241.
- Kim, K.S. (2001). Implications of user characteristics in information seeking on the World Wide Web. *International Journal of Human-Computer Interaction*, **13**(3), 323-340.
- Kim, K.S. & Allen, B. (2002). Cognitive and task influences on Web searching behavior. *Journal of the American Society for Information Science and Technology*, **53**(2), 109-119.
- Kline, P. (2000). *The new psychometrics*. London; Philadelphia, PA: Routledge.
- Krishnan, A. & Jones, S. (2005). TimeSpace: activity-based temporal visualisation of personal information spaces. *Personal and Ubiquitous Computing*, **9**(1), 46-65.
- Kules, B., Capra, R., Banta, M. and Sierra, T. (2009). What do exploratory searchers look at in a faceted search interface? In *Proceedings of the 9th ACM/IEEE-CS Joint Conference on Digital Libraries, Austin, TX, USA, June 15 - 19, 2009: JCDL '09* (pp. 313-322). New York, NY: ACM Press.
- Kuo, B.Y-L., Hentrich, T., Good, B.M. & Wilkinson, M.D. (2007). Tag clouds for summarizing web search results. In *Proceedings of the 16th international conference on World Wide Web (Banff, Alberta, Canada, May 8-12, 2007).* (pp. 1203-1204). New York, NY: ACM Press. Retrieved 31 October, 2009 from http://www2007.org/posters/poster1046.pdf (Archived by WebCite® at http://www.webcitation.org/5kxV7Gwqn)
- Lam, T.H.W., Liu, J.N.K. & Lee, R.S.T. (2009). Master: an intelligent ontology-based multi-agent system for sightseer. *International Journal of Software Engineering and Knowledge Engineering*, **19** (2), 137-147.
- Lieberman, H. (1997). Autonomous interface agents. In *Proceedings of the SIGCHI Conference on*

Human Factors in Computing Systems: CHI'97, Atlanta, Georgia, United States (pp. 67-74). New York, NY: ACM Press.

- Lohse, G. (1997). The role of working memory on graphical information processing. *Behaviour & Information Technology*, **16**(6), 297-308.
- Marlow, C., Naaman, M., Boyd, D. & Davis, M. (2006). HT06, tagging paper, taxonomy, Flickr, academic article, to read. In *Proceedings of the Seventeenth Conference on Hypertext and Hypermedia, Odense, Denmark, August 22 - 25, 2006: HYPERTEXT '06* (pp. 31-40). New York, NY: ACM Press.
- McGrew, K.S. (1997). Analysis of the major intelligence batteries according to a proposed comprehensive Gf-Gc framework. In D. P. Flanagan, J. L. Genshaft & P. L. Harrison (Eds.), *Contemporary intellectual assessment: theories, tests and issues* (pp. 151-179). New York, NY: Guildord.
- Messick, S. (1970). The criterion problem in the evaluation of instruction: assessing possible, not just intended outcomes. In M.C. Wittrock & D.E. Wiley (Eds.), *The evaluation of instruction: issues and problems*. (pp. 183-220). New York, NY: Holt Rinehart and Winston.
- Miller, G. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, **63**(2), 81-97.
- Modjeska, D. & Chignell, M. (2003). Individual differences in exploration using desktop VR. *Journal of the American Society for Information Science and Technology*, **54**(3), 216-228.
- Moulaison, H.L. (2008). Social tagging in the Web 2.0 environment: author vs. user tagging. *Journal of Library Metadata*, **8**(2), 101-111.
- Nov, O., Naaman, M. & Ye, C. (2008). What drives content tagging: the case of photos on Flickr. In *Proceeding of the Twenty-Sixth Annual SIGCHI Conference on Human Factors in Computing Systems, Florence, Italy, April 05 - 10, 2008: CHI '08* (pp. 1097-1100). New York, NY: ACM Press.
- Palmquist, R.A. & Kim, K.S. (2000). Cognitive style and on-line database search experience as predictors of Web search performance. *Journal of the American Society for Information Science*, **51**(6), 558-566.
- Parry, D. (2008). 'Tell me the important stuff': fuzzy ontologies and personal assessments for interaction with the semantic Web. In *Fuzzy Systems 2008, FUZZ-IEEE 2008 (Hong Kong)* (IEEE International Conference on Computational Intelligence). (pp. 1295-1300). Edison, NJ: IEEE.
- Riding, R. & Rayner, S. G. (1998). *Cognitive styles and learning strategies*. London: David Fulton Publisher.
- Rivadeneira, A.W., Gruen, D.M., Muller, M.J. & Millen, D.R. (2007). Getting our head in the clouds: toward evaluation studies of tagclouds. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems: CHI'07* (pp. 995-998). San Jose, CA: ACM Press.
- Rosenholtz, R., Li, Y. & Nakano, L. (2007). Measuring visual clutter. *Journal of Vision*, **7**(2), 1-22. Retrieved 1 March, 2009 from http://journalofvision.org/7/2/17/article.aspx (Archived by WebCite® at http://www.webcitation.org/5kus136qz)
- Saracevic, T. (1991). Individual differences in organizing, searching and retrieving information. In *ASIS '91: Proceedings of the 54th ASIS Annual Meeting, Washington, DC, October 27-31, 1991.* Medford, NJ: Learned Information Inc. for the American Society for Information Science.
- Sein, M.K., Olfman, L., Bostrom, R.P. & Davis, S.A. (1993). Visualization ability as a predictor of user learning success. *International Journal of Man-Machine Studies*, **39**(4), 599-620.
- Shiri, A.A. & Revie, C. (2003). The effects of topic complexity and familiarity on cognitive and physical moves in a thesaurus-enhanced search environment. *Journal of Information Science*, **29**(6), 517-526.
- Sinclair, J. & Cardew-Hall, M. (2008). The folksonomy tag cloud: when is it useful? *Journal of Information Science*, **34**(1), 15-29
- Sweller, J., van Merrienboer, J. & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, **10**(3), 251-296.
- Toms, E., O'Brien, H., Mackenzie, T., Jordan, C., Freund, L., Toze, S. and others. (2008). Task effects on interactive search: the query factor. In *Focused access to XML documents* (pp. 359-372). Berlin: Springer. (Lecture notes in computer science).
- Weller, H.G., Repman, J. & Rooze, G.E. (1994). The relationship of learning, behavior and cognitive styles in hypermedia-based instruction: implications for design of HBI. *Computers in the Schools*, **10** (3/4), 401-420
- Westerman, S.J. (1995). Computerised information retrieval: individual differences in the use of spatial

and non-spatial information. *Perceptual and Motor Skills*, **81**(3, part 1), 771-786

- Westerman, S.J., & Cribbin, T. (2000). Mapping semantic information in virtual space: dimensions, variance and individual differences. *International Journal of Human-Computer Studies*, **53**(5), 765-787
- White, R. W., Jose, J. M. & Ruthven, I. (2003). A task-oriented study on the influencing effects of query-biased summarisation in web searching. *Information Processing and Management*, **39**(5), 707-733
- Wickens, C.D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, **3**(2), 159-177.
- Witkin, H.A., & Goodenough, D.R. (1981). Cognitive styles: essence and origins. Field dependence and field independence. New York, NY: International Universities Press.
- Woodruff, A., Faulring, A., Rosenholtz, R., Morrison, J. & Pirolli, P. (2001). Using thumbnails to search the Web. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems: CHI'01* (pp. 198-205). New York, NY: ACM Press.
- Woodruff, A., Rosenholtz, R., Morrison, J. B., Faulring, A. & Pirolli, P. (2002). A comparison of the use of text summaries, plain thumbnails and enhanced thumbnails for Web search tasks. *Journal of the American Society for Information Science and Technology*, **53**(2), 172-185
- Zhang, H. & Salvendy, G. (2001). The implication of visualization ability and structure preview design for Web information search tasks. *International Journal of Human-Computer Interaction*, **13**(1), 75-95

How to cite this paper

Find other papers on this subject

Scholar Search   Google Search   Bing

■ Bookmark This Page

Appendix 1

The four search tasks

Task id: **FF₁**. Finding a Pub in London

Scenario: Your friend is a vegetarian and also enjoys beer. In particular, he likes the British ales. He will be visiting London the next week and has asked you to find a London pub with a vegetarian cuisine. He is travelling now and does not have access to the Internet. You will need to send a text message to him with the information that you find.

Your task: Please find a pub in London where vegetarian dishes are served. When you are done, please text him the name of one such pub.

(*Initial query*: london pub)

Task id: **IG₂**. Collections in London museums

Scenario: Your high school friend has recently got married. He and his bride are travelling through European cities

(Venice, Vienna and Paris). Because of the airport staff strike in Paris they have unexpectedly re-routed to London. They called you and asked for information about museums where they can see collections of interest to them. You know they are interested in anthropology, contemporary paintings and photography. You will search on the Internet and send them your recommendations by texting the names of three museums.

Your task: Please find three museums in London that will allow them to see such collections. When you are done, please text him the names of the museums you found.

(*Initial query*: london museums)

Task id: **FF₃**. Locating London hotel

Scenario: Your relative is coming back from India. She will have a one night stopover at Heathrow airport near London. But she forgot to make hotel arrangements earlier and has asked you to find a hotel is close to Heathrow. You will need to send her a text message what you find to his cell phone.

Your task: Please find a hotel located around the Heathrow airport. When you are done, please text her the name of this hotel.

(*Initial query*: london hotel)

Task id: **IG₄**. London Shopping

Scenario: Your significant other has asked you for information about some stores in London. She or he will be travelling to London next week and wants to get some LPs, books and t-shirts. You need to find information on where to buy British LPs (a.k.a. vinyl records), design books and cool t-shirts. You will search on the Internet and send your recommendations by texting the names of three stores.

Your task: Please find the three stores in London. When you are done, please text him the names of the stores you found.

(*Initial query*: london shopping)

Appendix 2

Details of the use of result lists vs. content pages

This appendix presents detailed results from the analysis of viewed result lists vs. visited individual content Web pages. We divided all cases into three groups according to the ratio of visited URLs to tags clicked as shown in Table 14.

URLTagRatio Value	Description
<1	**tags –** more tags clicked than visited URLs
=1	neutral – the same number of tags and URLs
>1	**URLs** - more visited URLs than tags clicked

Table 14: Grouping of URL to tag ratio (URLTagRatio)

We conducted a series of non-parametric tests to assess how the URL to tag ratio differed between the levels of task types, interfaces and cognitive abilities. We did it for all pairs of task and interface conditions. Tables 15, 16 and 17

show significant differences in URLTagRatio between all pairs of task and interface conditions. The results should be read from a table row to a column. Each cell represents a comparison of URLTagRatio between task and interface combination described by a row and task and interface combination described by a column. The word **tags** in a table cell indicates that more searchers used tags in the row condition than in the column condition. The word **URLs** indicates the same kind of relationship for URLs. For ease of reading, we have given the significant differences for the relevant table cells. However, one will note that these tables are symmetric with respect to the main diagonal.

		Task			
		Simple task		Complex task	
Task	User interface	List	Overview	List	Overview
Simple Task	List	n/a		URLs*	
	Overview		n/a	URLs**	URLs***
Complex Task	List	tags*	tags**	n/a	
	Overview		tags***		n/a
* p<0.05; ** p<0.01; *** p<0.001					

Table 15: Summary of Task/UI effects on URL/Tag Ratio

		Task			
		Simple Task		Complex Task	
Task	User Interface	List	Overview	List	Overview
Simple Task	List	n/a	tags^{b}	URLs^{b}	
	Overview	URLs^{b}	n/a	URLs**	URLs$^{**, L\,*}$
Complex Task	List	tags^{b}	tags**	n/a	
	Overview		tags$^{**, L\,*}$		n/a
* p<0.05; ** p<0.01; b: p=0.053 (borderline) L* Also significant for people with low *verbal closure*, p<0.05					

Table 16: Summary of Task/UI effects on URL/Tag Ratio for CV high

		Task			
		Simple Task		Complex Task	
Task	User Interface	List	Overview	List	Overview
Simple Task	List	n/a			
	Overview		n/a	URLs**	URLs$^{*, L\,**}$
Complex Task	List		tags**	n/a	
	Overview		tags$^{*, L\,**}$		n/a
* p<0.05; ** p<0.01; L** Also significant for people with low memory span, p<0.01					

Table 17: Summary of Task/User Interface effects on URL/Tag Ratio for high memory span

2 8 7

© the author, 2009.
Last updated: 30 October, 2009

Contents | Author index | Subject index | Search | Home