Effects of Working Memory Capacity on Users’ Search Effort

Jacek Gwizdka
School of Information
University of Texas
1616 Guadalupe St
Austin, TX 78701 USA
midi2013@gwizdka.com

Paper presented at MIDI 2013 Conference, 24-25.06.2013, Warsaw, Poland.

Abstract
We examined user behavior on information search tasks at two levels of complexity. Users were divided into two levels of working memory span (WM). The results show that in more demanding task conditions both user groups change behavior, but they differ in how they change it. High-WM user performed more actions to find more information, while low-WM users and switched their search tactic by significantly decreasing individual documents they visited.

Author Keywords
Search User Interface; Cognitive Abilities; Working Memory; Cognitive Load.

ACM Classification Keywords
H3.3 Information Search and Retrieval: Search process; H.5.4 Information Interfaces And Presentation: Hypertext/Hypermedia – Navigation.

Introduction
Information search requires evaluating documents, making relevance decisions and deciding when to stop the search process. In an idealized situation, a searcher could be assumed to possess perfect information and infinite resources (e.g., time, mental capacity). In reality, the available resources are limited. Evaluation of search results and encountered documents, decisions

what to examine and what is relevant are constrained by limited mental capacity. A limited number of “slots” in working memory is used to explain some constraints on human capacity for cognitive processing [1]. This limited capacity was first described by George Miller in his famous paper on “the magic number seven plus or minus two” [14]. Newer findings suggest the actual limit of working memory capacity may be lower or that it’s nature is different and lies in the control of attention required for effective information processing [11]. Nevertheless the fundamental notion of limited mental capacity is commonly agreed upon.

Understanding what contributes to a user’s mental load during a search process is important for identifying cognitive demands imposed by search tasks, user interfaces, and information displays. We know, for example, that users may be avoiding some interactions due to heightened mental requirements. In one information retrieval system a user relevance feedback feature was avoided by users due to the heightened cognitive load [2].

Study of mental demands and effort can involve an assessment of users’ mental load, a control of mental demands imposed by a task, by a system, or characterization of users by their levels of mental capacity. These methods are often used in combination.

The body of work related to mental load in human-computer interaction is quite extensive, both in terms of the studied effects as well as in terms of mental load measurement methods used (for example, for a review of measurement methods see [5]; for a review of studies see [9]). Here we only mention work most closely related to project presented in this paper. Prior work in the area of mental load and information search has established differences between task stages [7][12][9]. However, there is a general lack of work that would examine effects of cognitive abilities on information search tasks.

In the study reported in this paper we controlled search user interface, mental demands imposed by a task (by varying its level of complexity) and we assessed participants’ mental capacity by testing them on a working memory task. In this paper, we focus on the effects of working memory capacity and its interaction with task complexity, while the effects of user interface are reported elsewhere [10].

Methodology

37 undergraduates (8 females & 29 males) participated in an information search study conducted in a controlled experimental setting. The tasks were performed on a Windows XP desktop computer with a 17” LCD monitor equipped with eye tracking hardware (Tobii T60).

User Interfaces

We created two search interfaces: one with a results overview and one without. Both interfaces displayed a list of results similar to a traditional textual result list (Fig.2). Each result contained: title, URL and a list of descriptive tags. URL in each search result was linked to an external website.

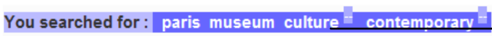
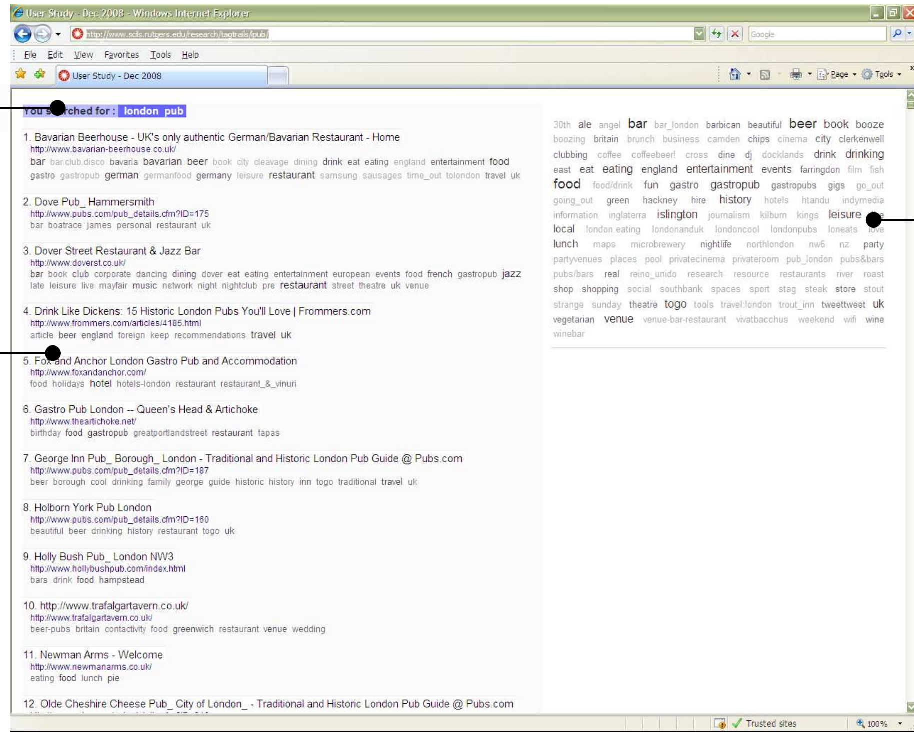


Figure 1. Interface detail: Query terms area. All search terms are shown. User added terms can be removed by clicking on "-".



Individual search results. Each consists of a title, a URL, and a list of descriptors (tags).

The interactive overview was presented as an alphabetical tag cloud that contained tags from all returned results. Tags of frequency lower than 2 were not used.

Tags in the cloud were organized alphabetically as it was expected to help finding a specific tag and to aid in information finding.

Figure 2. Search results list with overview (O+L UI). The "baseline" interface (List UI) contained search result list only and did not include the overview tag cloud.

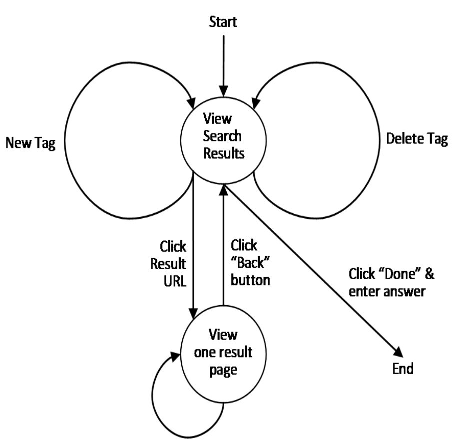


Figure 3. State diagram of user states and transitions.

The interface with overview (O+L UI) displayed an overview of the returned results (in a form of a tag-cloud) in addition to the list (Fig. 2).

Documents (web pages returned as search results) could be evaluated based on the descriptive words available in the search results or based on reading the document by visiting a web page. Clicking on a tag (keyword) in any of the two search interfaces added the tag to a search query, and thus narrowed down the results. The interface allowed for selective removal of tags. Query terms area is shown in Fig. 1. The possible user actions are shown Fig. 3.

The experiment's data set on topics related to travel, sightseeing and shopping was obtained by crawling the Delicious social bookmarking site. We collected approximately 18,000 unique bookmarks along with associated user-assigned tags (created by 600,000

FF	IG	FF	IG
FF	IG	IG	FF
IG	FF	FF	IG
IG	FF	IG	FF

Table 1: Search task order. FF – fact finding; IG – information gathering

users who added a total of 380,000 tags). We used the study task topics (related to London and Paris) to further select approximately 100,000 tagging instances (combinations of unique URL-tag pairs) that were applied to 1,700 bookmarks. The data was cleaned to remove non-Latin alphabets, references to browsers, and to personal collections as the source of bookmarks (e.g., mylibrary).

Tasks

Travel, sightseeing and shopping were selected as everyday search topics familiar to the general public. Task scenarios were constructed to present realistic situations and to provide participants with the search context and the basis for relevance judgments [1]. In all tasks participants were motivated to look for information for their friend. Tasks were designed at two levels of complexity: simple tasks involved finding a fact that satisfied specified criteria (e.g., name of a hotel located close to an airport), while more complex tasks involved information gathering about several items and selecting those that satisfied several criteria (e.g., finding three museums that collectively carried collections of three different kinds). This increased complexity by creating multiple (interdependent) paths to the task outcome. Users started tasks from the same initial list, which showed results from a two-word query. Users looked for information within these results. Narrowing search results was obtained by clicking on a tag, and was conceptually equivalent to adding a term to the search query. To broaden results a user could delete a tag (Fig. 1 & Fig. 3). This is equivalent to removing a term from a search query.

The number of tags added and deleted by a user correspond 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, possibly quite long, Web page). We considered participants using tag deletion as a cognitive move (i.e., formulating a new query) or as a "physical" move (i.e., navigating back to a previous results list). At the completion of each search task, participants entered their answers by composing a message, in which they made a sightseeing recommendation to their friend.

Study Design and Measures

The study had a within-subject design with task complexity and user interface as two main controlled factors. Each participant performed four tasks, two in each of the interfaces; the interface was switched after the second task. 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 (Table 1) and the interface. Four possible orders of two task complexities and two interfaces yielded a total of eight different task and interface rotations. Participants also performed working memory span (WM) task. Working memory (WM) reflects the ability to temporarily store and perform a set of cognitive operations on information that requires attention and the management of the limited capacity resources of short-term memory. Participants were divided into two groups, high-WM and low-WM, by splitting them at the median value of memory span score.

Task	Low WM	High WM
Simple	233	**156**
Complex	241	270

Table 2: Effect of task and WM on task duration [seconds]. Interaction effect (Task x WM): F(144,1)=4.2; p=.042. Task effect for High WM only: t(60.6)=-3.3; p=.002

Task	Low WM	High WM
Simple	5.1	3.7
Complex	7.2	**8.5**

Table 3: Effect of task and WM on the number of queries. Interaction effect (Task x WM): F(144,1)=3.1; p=.08. Task effect for High WM only: t(61.9)=-4.6; p<.001

Task	Low WM	High WM
Simple	8.5	7.1
Complex	**3**	5.2

Table 4: Effect of task and WM on the number of opened individual documents. Borderline task effect for Low WM only: t(47.1)=2; p=.051.

Task	Low WM	High WM
Simple	399	278
Complex	396	469

Table 5: Effect of task and WM on the number eye fixations. Interaction effect (Task x WM): F(144,1)=; p=.048; ask effect for High WM only: F(67,1)=8.6; p=.005

Task	Low WM	High WM
Simple	138	93
Complex	139	160

Table 6: Effect of task and WM on total fixation duration [sec]. Borderline Interaction effect (Task x WM): F(144,1)=; p=.051; Task effect for High WM only: F(67,1)=21; p<.001

Eye-tracking Derived Measures

Eye tracking data has been used in information search studies as a source of additional evidence to indicate relevance of search results. The focus has been on eye fixations, for example to indicate which items are considered in ranked search results pages or in identifying words useful for relevance feedback [4][13][15]. In contrast, we use eye tracking data to calculate objective measures of a user's cognitive processing related to reading and interacting with text. We implemented a simple, line-oriented reading model influenced by E-Z Reader – a cognitively-controlled, serial-attention model of reading eye movements [17]. We used eye fixations longer than 113ms to select those likely to result in word understanding. [6] provides implementation details.

We assessed cognitive search effort using the observed search and navigation decisions and using reading effort derived from eye-tracking data. The former were expressed as user actions: selection/de-selection of search terms and selection of documents to view (visits to result web pages). These selection/de-selection actions were equivalent to the number of query reformulations. Taking into account the possibility of deleting tags and document visits to be "physical" moves, we used 50% of these actions as cognitive actions.

Eye-tracking derived data served to operationalize several indicators of cognitive effort due to interacting with text. The number of regressions in a reading sequence and the fixation durations of the regression fixation have been associated with the difficulty of reading passages, resolution of ambiguous (sense) words, conceptual complexity of text, parsing difficulties and the reading goal [16].These indicators included, fixation duration, the existence and number of regression fixations in the reading sequence and the spacing of fixations in the reading sequence. The eye-tracking measures included also reading model parameters, such as counts and probabilities of scanning vs. reading.

Results

Our main interest is to investigate effects of working memory capacity and task complexity on the search process.

Task Effects

As expected, the more complex tasks required significantly more time (255s vs. 195s) and more effort. The increased effort on the more complex tasks was reflected in more actions performed (mean 7.8 vs. 4.5), in longer maximum reading fixation length and in more reading fixation regressions. Somewhat surprisingly, task outcomes did not differ significantly between the task and WM conditions (there was only a borderline difference p=.066 in task outcome for high WM users).

Task and WM Interaction Effects

Interaction effects of working memory span (WM) and task complexity on task duration and the number of tags clicked (number of queries entered and reformulated) were statistically significant. On the simpler tasks, high WM users spent overall less time than low WM (Table 2, They also clicked fewer tags, and performed fewer cognitive search actions overall (Table 3&4). Overall, high WM users performed actions faster than low WM (Table 8).

In contrast, the situation was different for the more complex tasks. High WM users tended to take more time than low WM users (Table 2). They also performed more search actions (Table 3&4). In this condition, high WM users also performed cognitive actions faster than low WM (Table 8).

Task	Low WM	High WM
Simple	9.35	7.25
Complex	8.70	11.11

Table 7: Number of cognitive actions (queries + 50% of individual document visits). Combines the results from Table 3 and 4.

Task	Low WM	High WM
Simple	24.9	21.5
Complex	27.7	24.3

Table 8: Time per action [sec]. Combines the results from Table 2, 3 and 4.

Examining the effect of task complexity and WM on the number of visited individual documents (Table 4), we observe that this number drops significantly for low WM searchers on complex tasks, while for high WM users this number also is lower, but not significantly.

Eye-tracking data (Tables 5&6) show somewhat similar patterns for the number of lexical fixations and the total fixation duration (calculated for search interface pages only). What's interesting is lack of difference for low WM, while the eye fixation measures increase for high WM in higher complexity tasks.

Discussion

The main effects of task complexity were as expected. The interaction effects between task complexity and WM showed that the relationship between effort and task complexity was different for low and high WM users. Low WM users visited fewer documents in more complex tasks and it is plausible they were satisficing [18]. Compared to the low WM group, high WM users tended to perform more actions than under conditions of increased difficulty. Higher complexity tasks required finding more information and thus it was expected that these tasks would required entering more queries and, possibly, visiting more individual documents. This expectation was generally confirmed. However, it was not significant for low WM users. Yet, task outcomes of the high WM group did not show the benefit of the extra effort invested in the task performance. The task outcomes of low WM users were about the same as those of high WM users.

For visits to individual documents the trend to perform more cognitive actions (as reflected by issuing more queries) reversed in more complex tasks. The low WM users visited much fewer individual documents. While high WM also visited fewer documents in high complexity tasks, this relationship was not significant for them, but it was statistically for low WM.

One could speculate that, in order to "save" mental effort, lower mental capacity of low WM users forced to drop more visits to individual documents than in the case of high WM users. Perhaps this effect can be plausibly explained by the principle of least effort [18]. User spent the minimum amount of effort that was just sufficient for completing tasks (to maximizing the outcomes).

As demonstrated by the eye-movement data, high WM people did more reading on the more complex tasks.

Perhaps they gained knowledge that was not immediately needed and thus was not measured in this study. Perhaps the higher cognitive capacity allowed these searchers to take advantage of serendipitous information encounters [8]. Perhaps they applied a more complete and global evaluation of information that did not result in a measurable improvement of task outcomes, and that did not follow the principle of least effort. The presented results do not fully explain why the principle of least effort was "violated" in certain conditions.

Conclusion

In this study, we examined user behavior on information search tasks at two levels of complexity and for two users groups characterized by two levels of working memory span (WM). The results show that in more demanding task conditions both user groups change behavior, but they differ in how they change it. High-WM user performed more actions to find more information, while low-WM users and switched their search tactic by significantly decreasing individual documents they visited.

Study limitations include a) limited variation in search tasks, b) tasks were relatively easy – hence we did not see the significant differences in task outcomes, c) query "formulation" was done by clicking on words. We plan to address these limitations in further studies.

Acknowledgements

We gratefully acknowledge support from IMLS grants # LG-06-07-0105 and RE-04-11-0062-11.

References

[1] Baddeley, A. (1986). *Working memory*. Oxford: Clarendon Press.

[2] Back, J. & Oppenheim, C. (2001).A model of cognitive load for IR: Implications for user relevance feedback interaction. *Information Research*, 6(2), Article WS2. Retrieved from http://informationr.net/ir/6-2/ws2.html

[3] Borlund, P. (2003). The IIR evaluation model: a framework for evaluation of interactive information retrieval systems. *Information Research*, 8(3), paper 152/2003.

[4] Buscher, G., Dengel, A., & van Elst, L. (2008). Eye movements as implicit relevance feedback. *CHI'2008*. pp. 2991-2998. New York, NY: ACM.

[5] Cegarra, J., Chevalier,A. (2008).The use of tholos software for combining measures of mental workload: Toward theoretical and methodological improvements. *Behavior Research Methods*, 40(4), 988–1000

[6] Cole, M. J., Gwizdka, J., Liu, C., Bierig, R, Belkin, N. J., Zhang, X.(2011). Task and User Effects on Reading Patterns in Information Search. *Interacting with Computers*, 23(4), 346 - 362.

[7] Dennis, S., Bruza, P., & McArthur, R. (2002). Web searching: A process oriented experimental study of three interactive search paradigms. *JASIST*, 53(2), 120–133.

[8] Erdelez, S. (1997). Information encountering: a conceptual framework for accidental information discovery. Proceedings of *ISIC 1997*. 412–421. Tampere, Finland: Taylor Graham Publishing.

[9] Gwizdka, J. (2010). Distribution of cognitive load in web search. *JASIST*, 61(11), 2167 – 2187.

[10] Gwizdka, J, Cole, M. (2013). Does Interactive Search Results Overview Help? An Eye tracking Stud CHI'13 Extended Abstracts on Human Factors in Computing Systems. 1869-1874. ACM Press.

[11] Kane, M.J., & Engle, R.W. (2003).Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology General*, 132(1), 47–70.

[12] Kim, Y.M., & Rieh, S.Y. (2005). Dual-task performance as a measure for mental effort in library searching and web searching. Proceedings *ASIST'2005*.

[13] Loboda, T., Brusilovsky, P., & Brunstein, J. (2011). Inferring word relevance from eye-movements of readers. Proceedings of IUI'2011. pp. 175-184. ACM.

[14] 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.

[15] Pan, B., Hembrooke, H., Joachims, T., Lorigo, L., Gay, G. & Granka, L. (2007). In Google we trust: Users decisions on rank, position, & relevance. *JCMC*, 12, 801-823.

[16] Rayner, K. and Pollatsek, A. (1989). The psychology of reading. Lawrence Erlbaum Associates, Mahwah, New Jersey.

[17] Reichle, E. D., Rayner, K., and Pollatsek, A. (2004). The EZ Reader model of eye-movement control in reading: Comparisons to other models. *Behavioral & Brain Sciences*, 26, 04, 445-476.

[18] Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, 63, 129–138.