The Affects of Task Difficulty on Medical Searches

Anushia Inthiran
School of Information Technology
Monash University
Sunway Campus
anushia.inthiran@monash.edu

Saadat M. Alhashmi
School of Information Technology
Monash University
Sunway Campus
alhashmi@monash.edu

Pervaiz K. Ahmed
School of Business
Monash University
Sunway Campus
pervaiz.ahmed@monash.edu

ABSTRACT
In this paper, we analyze medical searching behavior performed by a typical medical searcher. We broadly classify a typical medical searcher as: non-medical professionals or medical professionals. We use behavioral signals to study how task difficulty affects medical searching behavior. Using simulated scenarios, we gathered data from an exploratory survey of 180 search sessions performed by 60 participants. Our research study provides a deep understanding of how task difficulty affects medical search behavior. Non-medical professionals and medical professionals demonstrate similar search behavior when searching on an easy task. Longer queries, more time and more incomplete search sessions are observed for an easy task. However, they demonstrate different results evaluation behavior based on task difficulty.

Categories and Subject Descriptors
H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval – relevance feedback, search process

General Terms
Human Factors, Experimentation

Keywords
Task difficulty, interactive search behavior, medical searching, user study

1. INTRODUCTION
Medical search engines provide a wealth of information and enable medical searches to be performed quickly. However, non-medical professionals (NMP) and medical professionals (MP) report many challenges when searching for medical information [3,6]. We investigate the affects of task difficulty on the search behavior of a typical medical searcher: NMP and MP. A participant qualifies as a MP if he/she has completed undergraduate medical education and is working as a physician, resident, medical officer or doctor. A participant qualifies as a NMP if he/she has not completed or pursuing any formal medical education. Participants are placed into one of these categories based on education and occupation. Besides occupation and education, perception of task difficulty for these categories of users could further vary based on knowledge and previous experiences. Upon completing each task, we asked participants to specify their subjective post-search task difficulty using a 3 point scale: easy (E), neutral (N) - neither easy nor difficult or difficult (D). With this information we analyze how participants differed in search behavior.

Previous work on task difficulty indicate when users find a task difficult, they are unable to locate all necessary information [2] and spend more time to complete a task [5]. They also demonstrate high search effort but low search efficacy [4]. While previous research is not specific to medical searching, in our study, we conduct an in-depth study to understand how medical task difficulty affects medical search behavior.

2. EXPERIMENTAL DESIGN
We performed an exploratory interactive user survey on a convenience sample of 60 (30 NMP, 30 MP) participants in a university setting. The domain used for this study is *MedlinePlus*[1]. *MedlinePlus* is used as the domain of study because it only utilizes basic text based matching hence we are able to observe for 'true' interactive behavior. All participants except 2 (NMP) had prior experience conducting medical based searching. NMP perform medical searching as and when there is a need to do so. MP performed medical searching on an average of 2 hours per week.

A total of 6 tasks were developed: 3 for NMP and 3 for MP. Participants are asked to carry out 3 search task each. The task was developed to fit the concept of a simulated scenario [1]. We did not develop similar tasks for both categories as it would be unrealistic to have NMP and MP search on the same task. Topics covered in the simulated task are based on prevalent cases observed in a public hospital. Tasks were verified by 2 independent medical doctors for clarity, accuracy and to incorporate varying levels of difficulty. Clinical based tasks are developed as this type of search is popularly conducted by NMP and MP [3,6]. All searches are confined to *MedlinePlus*. Participants are told that they could stop the search once they have found satisfactory result/s or could not find satisfactory result/s (incomplete search).

3. RESULTS AND DISCUSSION
As shown in Table 1, perceived task difficulty for both categories of participants is dissimilar. These values allow a baseline for us to compare aspects of a search session. Data is not normally distributed hence a H-test is conducted to evaluate statistical significance. For NMP, a value of H=6.47 (p=0.00392) is obtained for the number of queries issued per task.

WWW 2012 Companion, April 16–20, 2012, Lyon, France.
ACM 978-1-4503-1230-1/12/04.

[1] http://www.nlm.nih.gov/medlineplus

Table 1. Difference in search behavior by task difficulty for NMP and MP. The table shows the sum for each category, mean values in bracket and percentage values (%) are provided where applicable, # denotes number.

Type	Difficulty (#)	# of Queries [Mean]	Query Length [Mean]	Boolean [Mean]	Medical[2] (%)	Querying Assistance (%)	# of Results Clicked [Mean]	Results Refining (%)	Completion Time [Mean]	Incomplete Session (%)
NMP	E (21)	98[4.4]	204[10.9]	19[3.1]	8	30.0	35[2.1]	40.0	199[9.5]	11.7
	N (15)	53[3.6]	140[10.3]	16[2.6]	0	6.6	28[2.0]	10.0	124[7.0]	2.5
	D (54)	153[2.6]	351[6.4]	10[2.5]	0	1.8	103[2.2]	15.0	232[4.2]	7.5
MP	E (24)	75[3.0]	189[10.1]	2[1]	15	3.0	67[1.6]	0	91[3.8]	38.4
	N (13)	23[2.0]	99[6.0]	4[1]	7	7.6	20[2.0]	0	40[3.2]	20.8
	D (53)	97[1.8]	273[5.4]	2[1]	3	8.3	34[1.5]	33.3	172[3.2]	11.3

H=6.47 (p=0.0356) is obtained for query length per task and H=33.14 (p<0.001) is obtained for task completion time per task. For MP, a value of H=3.27 (p=0.195) is obtained for the number of queries issued per task, H=7.87 (p=0.0195) is obtained for query length per task and H=13.72 (p=0.001) is obtained for task completion time per task. We use mean, percentage and H-values to interpret data.

In an easy task NMP on average issue more queries (p=0.0392), longer queries (p=0.0356) and took more time to complete the task. Queries in easy tasks are complimented with the use of more Boolean operators (AND, +, "") and medical queries. Medical queries are not issued in neutral or difficult task. The average number of results clicked is similar regardless of task difficulty. NMP use more querying and results refining assistances (features provided by *MedlinePlus* that users have to explicitly use) yet experience more incomplete search sessions in an easy task. Low task difficulty caused participants to increase querying effort and refining of search results. Although more effort is observed in an easy task, this still led to a high number of incomplete sessions. Querying and results evaluation behavior of NMP differ in relation to task difficulty.

MP searching on an easy task on average issue longer queries (p=0.0195) and took more time to complete the task (p=0.001). There is no significant difference in the number of queries issued. More medical terms are used in queries for an easy task in comparison to a difficult or neutral task. However, the usage of Boolean operators did not differ across task. While more querying effort is observed in an easy task, this task had the most number of incomplete search sessions. MP used more querying and results refining assistance in a difficult task but clicked on more results in a neutral task. MP demonstrate different querying and results evaluation behavior in relation to task difficulty.

NMP demonstrate high search effort in an easy task yet experience more incomplete search sessions. This same behavior is observed for MP. NMP rely on querying and results refining assistance in an easy task. However the usage of Boolean operators, results clicked did not differ based on task difficulty. MP spend more time and effort to refine results in a difficult task but clicked on more results in a neutral task. For MP the number of queries issued, usage of Boolean operators did not differ based on task difficulty. NMP and MP have different perceptions of task difficulty. Our results show task difficulty can substantially impact search behavior of NMP and MP. Interestingly, medical task difficulty affects medical searching behavior differently in comparison to non-medical searching. Our findings can inform the development of better medical retrieval strategies to be adaptive to search behavior based on task difficulty. This will help increase search satisfaction and efficacy.

4. ACKNOWLEDGEMENTS

We thank non-medical professional participants of this survey and medical professionals from University Hospital Kuala Lumpur, Sultanah Fatimah (Muar) Specialist Hospital and Monash University Sunway Campus.

5. REFERENCES

[1] Borlund P. (2003). The IIR evaluation model:A framework for evaluation of interactive information retrieval systems. Information research 8,(3), paper no 152 Retrieved from: http://information.net/ir/8-3/paper152.html

[2] Brystorm, K (2002) Information and information and sources in tasks of varying complexity Journal of the American Society for Information Science and Technology, 53(7), pp. 581-591

[3] Gonzalez-Gonzalez AI., Dawes, M., Sanchez-Mateos, J., Riesgo-Fuertes, R., Escortell-Mayor, E., Sanz-Cuesta, T and Hernandez-Fernandez, T (2010). "Information Needs and Information Behavior of Primary Care Physicians", Annals of Family Medicine, (5:4) pp. 345-352

[4] Gwizdka, J (2008).Cognitive load on web search task. In Proceedings of the Workshop on Cognition and the Web, Retrieved:http://rutgers.academia.edu/JacekGwizdka/Papers/939608/Cognitive_load_on_Web_search_tasks

[5] Liu, J., Cole, M., Liu C., Bierig, R., Gwizdka, J., Belkin NJ., Zhang, J and Zhang, X. (2010). Search Behaviors in Different Task Types, JDCL, 2010, pp. 69-78

[6] Spink A., Yang Y., Jansen, J., Nykanen P., Lorence, D.P., Ozmutlu, S. & Ozmutlu, H.C. (2004). A Study of medical and health queries to web search engines, *Health Information and Libraries Journal*, Vol. 21, pp.44-51

[2] Term used is found the *MeSH* database