Claremont Colleges
Scholarship @ Claremont

CGU Faculty Publications and Research | CGU Faculty Scholarship

1-1-2009

Persuading Consumers to Form Precise Search Engine Queries

Gondy Leroy
Claremont Graduate University

Recommended Citation
G. Leroy, "Persuading Consumers to Form Precise Search Engine Queries", American Medical Informatics (AMIA) Fall Symposium, San Francisco, November 14-18, 2009.

This Conference Proceeding is brought to you for free and open access by the CGU Faculty Scholarship at Scholarship @ Claremont. It has been accepted for inclusion in CGU Faculty Publications and Research by an authorized administrator of Scholarship @ Claremont. For more information, please contact scholarship@cuc.claremont.edu.

Persuading Consumers to Form Precise Search Engine Queries

Gondy Leroy, PhD
Claremont Graduate University, Claremont, CA

Abstract

Today's search engines provide a single textbox for searching. This input method has not changed in decades and, as a result, consumer search behaviour has not changed either: few and imprecise keywords are used. Especially with health information, where incorrect information may lead to unwise decisions, it would be beneficial if consumers could search more precisely. We evaluated a new user interface that supports more precise searching by using query diagrams. In a controlled user study, using paper-based prototypes, we compared searching with a Google interface with drawing new or modifying template diagrams. We evaluated consumer willingness and ability to use diagrams and the impact on query formulation. Users had no trouble understanding the new search method. Moreover, they used more keywords and relationships between keywords with search diagrams. In comparison to drawing their own diagrams, modifying existing templates led to more searches being conducted and higher creativity in searching.

Introduction

Search engines have become an essential part of our online experience and millions of people search online for health information. The information found online affects decisions about health, healthcare and visits to a healthcare provider for at least a third of the consumers[1], although numbers as high as 80% are reported for female health information consumers[2].

Enormous progress has been made over the years in the ability to store and index large collections, retrieve items in a very short time, and present results. In contrast to these backend processes, the user interface of search engines has not changed significantly with the exception of a few popular improvements such as spelling correction, phrases versus single words distinction, and term suggestion. From the first search engines, such as Archie, to today's Google, searching is based on forming a sequential string of words in a search text box. As a result, online search behavior has not changed either. We may type better or longer search phrases, but we essentially provide a 'bag of words' and expect a list of documents with those words highlighted.

The work described here explores the potential of a new interface for every day users. A user study was conducted to explore the new search paradigm and users' willingness and ability to use it and the impact on the quality of the search queries.

Search Engines and User interaction

User Queries

While the Internet is increasingly used for gaming, e-commerce, or gambling, information searching still forms a large portion of all online activities. Three types of searches exist: navigational, i.e., find a site, transactional, i.e., find activities, or informational, i.e., find information[3]. A 2001 survey indicated that almost half of the queries are informational, while later work showed a higher estimate of 80%[4]. Unfortunately, users do not form good and precise queries. Most queries contain only 2 or 3 words[5-9] regardless of the topic (medical or not)[10] and users vary widely in their ability to identify good versus poor keywords[11]. With so few words and billions of documents that contain them, returning the best matching documents is difficulty. Making matters worse is a finding by McCray and Tse[12] who suggested a correct alternative for a misspelled term, which does not contain any useful information, users accepted it in only 45% of cases. Furthermore, novices tend to start out with very general, imprecise queries[13] and consumers' mental models of the use of Boolean terms, stop word removal, and term order in a search engine are often incorrect[14]. With so few vague keywords, it is difficult for a search engine to retrieve the best matching documents.

Search Engine Advances

Most advances in search engines have been made in the backend processes. Efficient algorithms have been developed to assemble and provide access to increasingly large collections of documents, images, and video. Query results are fine-tuned and ranked according to a relevance criterion. Most algorithms are automated and do not require input from users. For example, the first ranking function consisted of simple tf\idf approaches[15, 16], but current algorithms are more advanced and take network or webpage characteristics into account, e.g., PageRank[17].

In addition to providing access, many algorithms focus on improving the user query. For example, query expansion aims to add keywords and so increase the query's precision. It can be manual or automatic. With automatic query expansion, term ranking functions select the additional terms for expansion. In contrast, manual expansion has now become readily available to users. Both Yahoo! and Google provide easy-to-use query expansion. When typing in the search box, a user can choose from suggested queries that complete the personal one. These suggestions are based on popularity calculations of terms from all searches being conducted with that search engine.

Most research, including our own, has focused overwhelmingly on backend algorithms and ignored the user interface. New input methods are seldom used by search engines. This work presents a first step toward a new, intuitive but more powerful search interface.

Diagram Queries

An affordance, first coined by Gibson[18], is a property of an entity or object that allows interaction with that object in a specific way. Manipulation of affordances has been used for decades to guide our interaction with the physical environment. For example, the type of door handle will influence whether you try to push or pull. Doors with a flat metal plate as door handle are meant to be pushed. On the Internet, such physical affordances are mimicked to help users behave in similar ways, e.g., buttons are *pushed*, users *check out* when buying books. A change in affordances can have large scale consequences. Baron[19] provides examples of behavioral and social change, sometimes unintended, as a result of changed affordances. For example, phones becoming mobile allowed us to roam (an affordance) and made us available everywhere.

Today's search engines provide only a single search box which allows users to submit a string of text. By not offering other options, no other type of search input is possible. The goal of this project is to improve the search interface. However, there are constraints. From personalization research, it is clear that users do not want to spend time training or filling out forms. Therefore, the interface should be effortless to use, should not require training, and should be simple to use, even for novices. This can be best accomplished by changes in the affordances.

Instead of a single search box, a 2-dimensional interface can be used that consists of multiple search boxes, connections between those searches boxes, and the ability to type in a search box and add additional boxes. Figure 1 shows an example query for the question "What medication treats depression in teenagers?" Each box represents a search term, e.g., "depression." The keyword can be a single word or an entire phrase. The labels on the arrows, "treats", and the directionality of these arrows show how the search terms need to be related to each other.

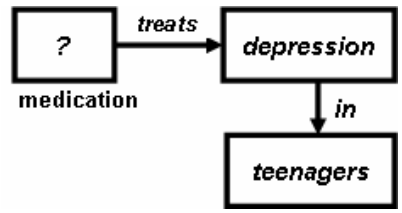

Figure 1. Example Query Diagram

In addition to encouraging more structured queries, this interface provides two additional, easy-to-use affordances. The first is the use of a question mark to specify the subject of the search, instead of a description of the needed information. The second is the option to add meta-information to search terms, for example "medication." This type of search will allow more precise searching. It will also help backend ranking algorithms by providing more structured information. To enable this type of searching, different pre-processing techniques and data structures will be needed but they are not the topic of this study.

User Study

Before developing an entirely new search engine user interface and the necessary required backend improvements, it was deemed prudent to design and prototype the interface. It is important to evaluate the willingness of everyday users to search differently without requiring a significant amount of training. This evaluation can be done by comparing paper-prototypes in a controlled manner.

Searching with Google, the most popular search engine in the United States, was compared to searching by query diagram. The study focused explicitly on user behavior during query formation, not the retrieved documents.

Methods

The *instructions* provided to participants included an explanation of the project and the request to write out the different searches they would perform to find the information. Since search diagrams are new, we also included an explanation of possible symbols and how they could be used to form a query. Participants were

asked to write out as many queries as they believed it would take, including as many keywords as they would have used online.

The study was conducted as a controlled laboratory experiment with one *independent variable*: the user interface. Three different user interfaces were compared: 1) a typical Google interface (control/baseline), 2) query templates that could be filled in or changed, and 3) constructing query diagrams without the use of templates. The conditions are referred to as the *Google*, *Template Diagram*, and *Blank Diagram* conditions. The order of the three conditions was randomized for the participants with one constraint: the Template condition always preceded the Blank condition. By first showing templates, implicit training was included for querying by diagram, making the condition more comparable to Google searching for which all users are already trained. This increases the validity of our conclusions.

Six different questions were used (Table 1), which were assigned to the different conditions in a balanced approach so that all questions were associated with the different conditions, i.e., not the same two questions per condition.

	Question
Q1	Do elderly women have a higher rate of dementia than elderly men?
Q2	Do men with sleep apnea tend to be overweight or have diabetes?
Q3	Do children in childcare centers have a higher rate of asthma?
Q4	Are the rates for recovery from alcoholism the same for men and women?
Q5	How common is removing multicystic dysplastic kidneys (MCDK) in children over the past 10 years?
Q6	Do children with autism have a higher rate of seizures?

Table 1: User study task: Questions posed to users

There were *two dependent variables*: user search behavior and a subjective user evaluation of the interface. To compare search behaviors, the number of searches, number of words per search, and number of words copied from the question were used. Since we used a paper-prototyping approach, paper copies of the interfaces were shown. Users were asked to write out each search separately and were informed they could write out as many searches with as many keywords as they thought they would need. The number of words per search was a simple count of each individual word in a search. Prepositions were included in this count. Although they are commonly ignored by search engines, they will play a significant role in the search diagrams. The number of words copied from the question or not will provide an indication of unique and creative searching.

The subjective evaluation was conducted with a short survey. Participants chose which input method, Google, Blank Diagrams, Template Diagrams, was fast, easiest, or best. We asked three questions:

1. Which version was easy to do (you can choose more than one)
2. Which version was fast to do (you can choose more than one)
3. Which version do you believe will give you the best results (you can choose more than one)

Study Results

Demographics

Twenty-two users participated in the study. Figure 1 shows an example user query using a template. The left side shows the original template, the right hand side user modification. The average age of the participants was 35; the youngest participant was 20 years old and the oldest was 72 years old. There were 6 different native languages represented, distributed as follows: 32% Arabic, 27% Chinese, 18% English, 14% Filipino, 5% Spanish (one participant), and 5% Tagalog (one participant). None reported that they did not understand the instructions.

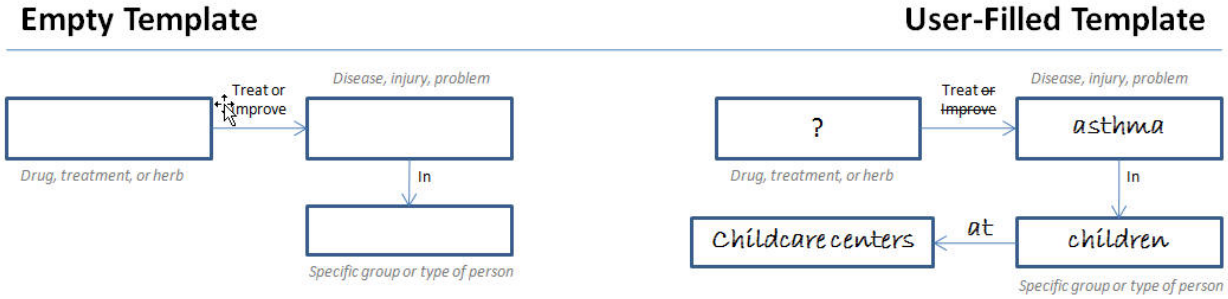

Figure 1.Example of template use in response to "Do children in childcare centers have a higher rate of asthma?"

Search Behavior Evaluation

The first comparison looked at the number of queries per question, number of words per question, and number of original words (not copied from the given task) per query. For word count, all words, including prepositions, were included for simplicity. Although search engines such as Google do not use prepositions (the term frequency of these terms is too high to make them useful), they will be important for the query diagrams and were therefore included. The meta-information was not included in this count, since it is not intended to be used as keywords.

Since all users completed the three conditions, a repeated-measures analysis was conducted (Table 2). There was a significant main effect ($F(2,78) = 32.52$, $p < .001$) for the number of words used. There were on average 4.8 words used in the Google queries, 7.9 words in the Template and 6.8 words in the Blank Condition. Post-hoc contrasts showed that the differences between the Google and Templates ($F(1,39) = 113.98$, $p < .001$), Google and Blank ($F(1,39) = 19.12$, $p < .001$), and Templates and Blank ($F(1,39)=7.30$, $p < .01$) were all significant.

A second significant main effect was found for the number of queries ($F(2, 84) = 7.163$, $p < .01$). On average, 1.4 queries were written for the Google condition. With Template Diagrams, the number was significantly higher with 1.9 queries, while the Blank Diagrams numbers, 1.1 queries, were similar to Google. Post-hoc contrasts showed a strong trend between Google and Template Diagrams ($p = .068$) and a significant difference between the Template and Blank Diagrams ($F(1,42) = 16.004$, $p < .001$).

N = 22 Averages per person:	Google	Template Diagram	Blank Diagram
Queries**	1.4	1.9	1.1
Words *	4.8	7.9	6.8
Words Copied from Question**	86%	61%	85%
Nodes	-	3.0	3.2
Labeled and Directional Arcs	-	1.9	1.6
Other Arcs	-	1.4	2.1

Table 2: Search Behavior (* $p < .01$, ** $p < .001$)

It has been shown that participants in controlled user experiments often will copy and paste words from the given search task[20]. This may partially explain the use of more words in our experiment compared to the numbers reported in the literature[5-9]. To investigate this, the percentage of words in user queries that were copied from the given task was calculated. The results were very surprising. Both conditions that allowed free query expression led to a lot of copying: 86% of the words in the Google queries and 85% in the Blank Diagrams. However, only 61% of words being copied in the Template condition (even though participants could and did change templates as they liked), which was significantly different from both other conditions ($F(1,39) = 33.20$, $p < .001$).

The last portion of the analysis focused on comparing the Template Diagram and Blank Diagram conditions. The use of nodes, labeled and directional arcs, and other arcs was very similar for both. An almost equal number of nodes were used in Template (3.0) and Blank (3.2.) Diagrams. Slightly more labeled and directional arcs were used in the Template (1.9) versus Blank (1.6) Diagrams. However, in the Blank Diagrams more unlabeled arcs (2.1) were used than in the Template Diagrams (1.4).

Different conditions let to different behaviors with new items. In the Blank Diagram condition, the question mark was used by two people. They each used it in one query. However, in the Template Diagram condition, six people used the question mark in a total of twelve queries.

User Interface Evaluation

Table 3 shows the subjective evaluation of each condition. Most users chose two options for each question: Google and another. Google received the highest numbers, indicating that participants found it easy and fast to use Google and also believed it to lead to good results. When using diagrams, participants preferred to draw their own diagrams: 45% chose their own diagrams as easy to do compared to only 18% choosing the templates, 41% chose their own diagrams as faster compared to only 27% choosing the templates, and 45% chose their own diagrams as providing better results compared to only 32% choosing the templates.

N = 22	Percentage of users choosing:		
Question posed:	Google	Template	Blank
Which version was easy to do:	82	18	45
Which version was fast to do:	68	27	41
Which version do you believe will give you the best results:	59	32	45

Table 3: Subjective evaluation

Conclusion

The goal of this study was to evaluate if health information consumers would be willing and able to use a different type of search input. Google's interface was compared to a diagram interface using a paper prototype. The results demonstrate that consumers were willing and able to use query diagrams, even with limited training. Moreover, query diagrams led to more search terms being used

in a more structured format. Using more terms makes it possible to match document more precisely. The meta-information makes it possible to fine-tune matches since it provides context for the document, e.g., treatments not appearance of asthma in childcare centers. Even so, users preferred Google, which was not unexpected given the years of training and comfort levels that have been achieved using Google. When comparing the two diagram conditions, users preferred to form their own diagrams but were more creative using templates.

A limitation that needs to be taken into account relates to paper-prototyping. Since users did not perform the actual search, they did not have the benefit of seeing results before modifying their queries. Actual searching may therefore differ.

The results show how searching for information can be improved in an intuitive manner. Unexpectedly, the results show that using Templates led to more unique queries (less copying). This provides a unique and exciting opportunity to the medical field. Many useful templates could be designed in advance by experts to help layperson find more precise information. Thesauri could help bridge the user query and the documents leading to much improved, semantically enriched searching.

Acknowledgements

The author thanks users for participating and M. Al-Karusi, A. Deonarine, C. Gonzales, J. Ku, MJ Kwak, and L. Sun, for helping distribute the study materials.

References

1. Baker L, Wagner TH, Signer S, Bundorf MK. Use of the Internet and E-mail for Health Care Information: Results from a National Survey. J. of the Am. Med. Ass. 2003 May 14;289(18):2400-6.
2. Warner D, Procaccino JD. Toward Wellness: Women Seeking Health Information. J. of the Am. Soc. for Inf. Sc. and Techn. 2004 ;55(88):709-30.
3. Broder A. A Taxonomy of Web Search. ACM SIGIR Forum. 2002 September;36(2):3-10.
4. Jansen BJ, Booth DL, Spink A. Determining the user intent of web search engine queries. 16th Int. Conf. on WWW; Canada; 2007. p. 1149 -50
5. de Lima EF, Pedersen JO. Phrase Recognition and Expansion for Short, Precision-biased Queries Based on a Query Log. 22nd Ann. Int. ACM SIGIR Conf. on Research and Dev. in Info. Retr.; Berkeley, CA USA; 1999. p. 145-52.
6. Ross NCM, Wolfram D. End User Searching on the Internet: An Analysis of Term Pair Topics Submitted to the Excite Search Engine. J. of the Ame. Soc. for Info. Sc.. 2000;51(10):949-58.
7. Spink A, Wolfram D, Jansen MBJ, Saracevic T. Searching the Web: The Public and Their Queries. J. of the Am. Soc. for Info. Sc. and Techn. 2001;52(3):226-34.
8. Toms EG, W.Kopak R, Bartlett J, Freund L. Selecting versus Describing: A Preliminary Analysis of the Efficacy of Categories in Exploring the Web. Tenth Text REtrieval Conference; Maryland; 2001.
9. Hölscher C, Strube G. Web Search Behavior of Internet Experts and Newbies. Ninth Int. WWW Conf.; The Netherlands; 2000. p. 337 - 46.
10. Lau T, Horvitz E. Patterns of Search: Analyzing and Modeling Web Query Refinement. Seventh Int. Conf. on User Modeling; 1998.
11. Ruthven I. Re-examining the Potential Effectiveness of Interactive Query Expansion. 26th Ann. Int. ACM SIGIR Conf. on R&D in IR; Toronto, Canada: 2003. p. 2113-220.
12. McCray AT, Tse T. Understanding Search Failures in Consumer Health Information Systems. AMIA; Washington, DC; 2003. p. 430-4.
13. Navarro-Prieto R, Scaife M, Rogers Y. Cognitive Strategies in Web Searching. 5th Conf. on Human Factors & the Web; Gaithersburg, Maryland; 1999.
14. Muramatsu J, Pratt W. Transparent Queries: Investigation Users' Mental Models of Search Engines. 24th Ann. Int. ACM SIGIR conf. on R. and D. in IR.; New Orleans, Louisiana; 2001. p. 217 - 24.
15. Salton G, The SMART Retrieval System: Experiments in Automatic Document Processing: Prentice-Hall; 1971.
16. Salton G, Buckley C. Improving Retrieval Performance by Relevance Feedback. J. of the Am. Soc. for Info. Sc.. 1990;41(4):288-97.
17. Page L, Brin S, Motwani R, Winograd T. The PageRank Citation Ranking: Bringing Order to the Web: Stanford Digital Library Technologies Project; 1998.
18. Gibson JJ. The Ecological Approach to Visual Perception. New Jersey: Lawrence Erlbaum Associates; 1986.
19. Baron NS. Always On: Language in an Online and Mobile World. New York: Oxford University Press; 2008.
20. Leroy G, Lally AM, Chen H. The Use of Dynamic Contexts to Improve Casual Internet Searching. ACM Trans. on Info. Syst. 2003 July;21(3):229-53.