This paper describes a research study done to detect differences in queries constructed in the traditional Google interface and the advanced search Google interface. Sixteen undergraduates at The University of North Carolina at Chapel Hill participated in this study.

Study participants were asked to complete a questionnaire of demographic information and then search for answers to four questions using either the traditional Google interface or the advanced Google interface.

This research indicated that users who use the advanced Google screen may construct queries with more terms. Additionally they use more Boolean logic and other search modifiers. More research needs to be done to substantiate this research due to the small sample size used.

**Headings:**

Online searching

User information-seeking behavior

Internet search engines
THE DIFFERENCES BETWEEN USERS’ QUERIES USING DIFFERENT SEARCH ENGINE INTERFACES

by
Ashley Richardson Langley

A Master's paper submitted to the faculty of the School of Information and Library Science of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Master of Science in Information Science

Chapel Hill, North Carolina
November, 2002

Approved by:

Advisor
Although the goal of much advanced IR research is to make computer do humanlike things, the possibility that people can be humanlike should not be overlooked . . . The potential for improving human interaction with recorded knowledge is immediate, but there seem to be few signs of interest in so practical a goal. Waiting for Godot, we fail to grasp what is now in reach (Swanson 1988).

You have just arrived at home with a new computer desk. The desk comes in a long, flat rectangular box with a photo on the top. When you open the box, you see various pieces and hardware, but there is no instruction sheet mixed in with the pieces. Perplexed, but determined, you start trying to assemble the desk without any directions. You have to put together and take apart components many times before you get something that looks like a desk. As you complete the final task of putting casters on the desk, you see a small piece of paper stapled to the bottom of the desk leg. You remove the folded paper and open to find a full set of directions. While relieved the directions exist, you are frustrated because they would have been useful about four hours ago when you started the assembly. Since you have completed the task, they are meaningless at this point.

In many ways using a search engine is similar to putting together a piece of furniture without directions. Users go to a search site and are
presented with something that does not look like much (in most cases, a rectangular text box) and asked to type something into it in the hopes that they get exactly the information they need after they click the “search” button. After they have been working with the system for a period of time, some stumble upon a “help” button to assist them with their query construction. Many never even find it.

Web search engines come with a plethora of special features to assist the user with information retrieval tasks. Most of the popular search engines support Boolean operators such as AND, OR and NOT. Phrase matching is another feature that almost all search engines embrace. Provided they know the syntax, users have the ability to limit searches by languages or domains with many search engines. Certain keywords can be omitted from the search results by many products, and a few search engines even utilize stemming to give users a more inclusive set of results (Notess 2002). The current study will explore if a more comprehensive interface will provide the users with the tools to create queries that are longer and make more substantial use of advanced search operators.
Background

Various studies have been done relating to query structure, query formation, and use of searching features of various interfaces. This rich body of research has data from an assortment of sources and interfaces.

Finechel studied users of mixed experience levels in their use of ERIC. She found that ERIC searchers with higher levels of experience used thesaurus terms extensively, modified their searches with greater frequency, and actively worked to limit the search costs (Fenichel 1981).

In 1989–1990, advanced humanities researchers were given introductory DIALOG training followed by unlimited access to the system. Analysis of the statements issued by these researchers showed that less than 37 percent of the queries issued used any Boolean operators. Despite training, users of the system reported confusion with syntax and operators while querying the system. Several users chose not to utilize the system despite going through training sessions (Siegfried, Bates, and Wilde 1993).

A study conducted in the late 1990s provides another source of data about searching while using DIALOG. Designed to be a study of users’ relevance judgments, data was collected using both logs, videotape and transcripts of users going through the search process. Later analysis of
this rich data set showed that the average query had a length of 15 terms with a standard deviation of 8.77 (Spink & Saracevic 1997).

More recently, Jansen, Spink, and Saracevic studied over 51,000 queries issued by over 18,000 users during a day using the Excite search engine (http://www.excite.com). During their analysis, they made several observations about users’ searching strategies. The average query length for this set of data was 2.21 terms with 85 percent of the queries consisting of 3 terms or less. These results were in stark contrast to earlier IR systems with their much higher term count (Jansen, Spink, & Saracevic 2000).

Excite supports Boolean operations; however, searchers must use capitalization for the search engine to differentiate between a Boolean operator and a search term. Looking at the Excite data, Jansen et al. (2000) found only 20 percent of the users tried any operator or modifier. Further, the modifiers were used incorrectly about half the time.

Additionally, these investigators chose to explore whether users reformulated queries while searching. They categorized queries as being unique, modified, or identical. The user issued unique queries once during a session whereas queries classified as identical were those using the same exact set of search terms multiple times during the session. Researchers considered a query to be modified if it contained the addition
and/or deletion of terms from an earlier search by the same person. This examination showed that 35 percent of the queries were unique while only 22 percent of them were modified (Jansen, Spink, & Saracevic 2000). This study paints the picture of web searches that are short, rarely reworked, and devoid of search modifiers.

Another study by Clarke, Cormack, and Tudhope corroborated the low term count finding when studying searching using the MultiText Netnews searcher. The average query for their web-based system was 2.9 terms (Clarke, Cormack, & Tudhope 2000). Additionally, researchers at Apple Computers found that 87% of the queries studied from three search engines had three or fewer terms (Rose & Stevens 1996).

While Excite was quite typical of search engines in 1997, other search engines have attempted to create an interface that is easier for users and solicits better queries. AskJeeves (http://www.askjeeves.com) is one of the newer attempts at soliciting better user input than keyword searching. It is designed for queries to be given in question form. By using natural language, the designers aim to create a more user-centric experience. Spink and Ozmultu (2002) conducted a study of AskJeeves logs and found several noteworthy results. Most of the users did not form queries that utilized question syntax. Moreover, those users who asked questions usually did not reformulate their queries. When users
chose to ask questions, they often did not directly address their information need. For instance, instead of asking “What were the original paint color schemes for a 1955 Chevrolet Bel Air?”, the user question would be “Where would I find information about color schemes for 1955 Chevrolet Bel Airs?”

While the mean number of search terms was 6.4 for the entire data set, the distribution of terms showed two peaks, one at 3 terms and another at 8 terms. Spink and Ozmultu concluded that the first peak corresponded to non-question queries similar to the results found in the Excite research, while the second peak was reflective of queries in a question format. Since AskJeeves solicits natural language queries, use of “and” and “or” were in context rather than as operators.

Jansen(2000) studied the effect of query complexity on the search results. This study focused on fifteen typical web queries ranging from two to four terms. These simple queries were run five search engines to give a base result. The appropriate modifiers such as quotations and Boolean logic were added to the queries which were submitted again to the five search engines. Jansen’s work showed that there was a statistically significant difference in the first ten results reported between the simple and modified queries. However, while statistically significant,
Jansen’s work showed that almost all of the simple queries retrieved at least half of the documents that the more elaborate query found.

Search engine interfaces vary, and the types of retrieval tasks users perform do as well. Questions come in a variety of shapes and forms. Some questions may be straightforward and only need a one word answer while others may be more conceptual in nature. Often users structure their information search differently depending on the information for which they are being asked. White and Iivonen (2001) researched how the type of question impacts a user’s strategies for finding the answer. Users were asked what tools they would use to answer sixteen questions. They could choose to use a search index, to go to an authoritative site, or to use a web search engine. The questions asked by the researchers were classified as either predictable or unpredictable based on the extent that the source of the information can be anticipated. Each question was also classified as open or closed based on the breadth of answers that would be applicable to the question. Some questions may have only one appropriate answer whereas others may have a number of acceptable responses.

Overall, White and Iivonen found that users chose to employ a search engine over browsing an index or going straight to an authoritative site. Specifically, 43 percent of the respondents stated they would use a
search engine to initiate an information seeking task. If a question was judged to have an unpredictable source for the answer, then participants relied on search engines from 60 to 66 percent of the time (White & livonen 2001).

Thus, this literature shows that web searchers use fewer terms than searchers utilizing other systems. Boolean operators and phrase searching continue to be parts of many systems but they rarely see use. Attempts to have users create natural language queries have been mixed because users do not ask their questions in a direct manner and often do not take the time to create a full statement of their need. Despite these issues, users tend to prefer using web searches more than going directly to an informative site or using a web directory such as Yahoo provides.

**Research Hypotheses**

The above research suggests that users do not use the special features built into many search engines. Many of the queries are short and constructed quite broadly. Since all of these web searches have used a standard interface for a given system, a study comparing a basic interface with a more advanced interface that does not require specialized syntax will enhance the understanding of the retrieval process. Therefore this study will test the following hypotheses:
Hypothesis 1: Participants who use the advanced search screen will create longer queries than the participants who use the regular search screen.

Hypothesis 2: Participants who use the advanced search screen will reformulate their queries less than those who use the regular search screen.

Hypothesis 3: Participants who use the advanced search screen will use Boolean logic and other modifiers more than those who use the regular search screen.

Methodology
In order to test these hypotheses, sixteen undergraduate students were recruited to perform four specified search tasks based on White and Iivonen’s (2001) study. Half used the Google basic search interface and half used the Google advanced search interface. Based on the transaction logs, each query was analyzed for length, use of special Boolean operators, and presence of other query modifiers. The study methods are described in more detail below.

The Search Interface. Google was chosen to be the search engine utilized for this study. Like most search engines, Google (http://www.google.com) does Boolean searching, phrase matching, and keyword exclusion. Additionally, one can limit search results to specific
domains and designated languages while searching with Google. It searches a variety of files including those in the portable document format (commonly referred to as pdf files). This site made its mark using a plain text box and little else on its main page at a time when many of the major competitors (such as Yahoo, AltaVista, Excite and Lycos) were putting long lists of categories on their entry page to promote browsing.

Figure 1: The Default Google Search Screen
Figure 2: The Advanced Google Search Screen

Google also has an advanced search feature, which goes far beyond most search engines in terms of query construction. This search page, complete with drop-down boxes, multiple choices and text boxes,
provides users with another way of searching for information. Unlike the regular interface, users do not have to be familiar with Boolean operators or specialized syntax to construct complex queries. Rather, all the user has to do is read the screen and make a few choices to request a specific set of documents.

Study Participants. Participants consisted of 16 undergraduates enrolled at The University of North Carolina at Chapel Hill. Six of the participants were enrolled in an introductory course on retrieving and analyzing information (INLS 40), and were recruited via an announcement during a class session. The other participants were recruited through an announcement in a computer literacy course offered through the computer science department or brought along by friends who were participating. The average age of the participants was twenty-three, and the group consisted of seven female and nine male subjects. Participants were offered non-monetary compensation in the form of a candy bar for their participation in the study which took between thirty and forty five minutes.

Study Procedures. Participants were assigned to use the simple Google interface (http://www.google.com) or the advanced Google interface (http://www.google.com/advanced_search?hl=en) based on their participant number. Odd-numbered participants used the advanced interface while even-numbered participants used the simple interface.
Participants completed an informed consent form (see Appendix A) and an initial questionnaire containing basic background/demographic information as well as information about familiarity with various search engines (see Appendix B). After completing the questionnaire, participants were assigned to a computer workstation running Red Hat Linux 7.3 and using the Mozilla browser. Each participant searched for answers to four questions using the designated interface. After each query, the users were instructed to click on the “Home” icon in the browser to return to the designated search screen.

The four search questions were chosen from White and Iivonen’s (2001) work. Each question covered one of the four facets of questioning studied by them. The questions chosen were:

1. What is the World Health Organization doing to stop river blindness in Africa? (closed/predictable)
2. I am looking for a copy of the multinational treaty banning land mines that was signed shortly after Princess Diana’s death, the one that US and Finland refused to sign. What is it and where can I find a copy? (closed/unpredictable)
3. Who are the current members of NATO, the North Atlantic Treaty Organization? (open/predictable)
4. What are considered to be the causes of hooliganism or fan violence at World Cup soccer games? (open/unpredictable)

The participants listed one or two of the best URLs for finding this information on a sheet that was returned to the researcher. The participants were instructed that they could run as many queries as
necessary to yield a set of information that is satisfactory to them. After answering the four questions to their own satisfaction, the participants were compensated and released.

**Data Analysis.** The browser log file was downloaded for analysis after each trial. Each URL for Google was viewed to determine the underlying query, and that information was recorded. Then, each query was analyzed for length, use of special Boolean operators, and presence of other query modifiers. The participants’ actual answers to the four study questions were not processed. All data was analyzed using MS Excel and MS Access.

**Research Findings**

The questionnaire given to the participants solicited information about the participant’s searching experience. Since the students’ backgrounds were varied, information gathering about the participants was warranted.

All but one of the sixteen participants answered that they had a favorite search engine. Ten participants (62.5%) stated that Google was their favorite search engine while five (31.3%) of the participants preferred Yahoo! All of the participants reported familiarity with at least two search engines (see Table 1).
<table>
<thead>
<tr>
<th>Search Engine Used</th>
<th>Number of Participants Reporting Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>AltaVista</td>
<td>14</td>
</tr>
<tr>
<td>Excite</td>
<td>11</td>
</tr>
<tr>
<td>Google</td>
<td>16</td>
</tr>
<tr>
<td>HotBot</td>
<td>7</td>
</tr>
<tr>
<td>Lycos</td>
<td>12</td>
</tr>
<tr>
<td>Yahoo</td>
<td>15</td>
</tr>
<tr>
<td>Other</td>
<td>11</td>
</tr>
</tbody>
</table>

Table 3: Experience of Participants with Major Search Engines

Participants’ reports of formal training in Internet searching were varied. Only three of the participants reported having any formal training in Internet searching outside of the INLS 40 class. Two of those students had some training in another college-level course while one of the participants reported that instruction had been provided in high school. No participants reported receiving instruction in a high school or college library setting.

Each participant was asked if he/she had read any of the help pages for a search engine. Only the six participants who were also enrolled in INLS 40 had read a help page. Many of the other students expressed visible amusement (i.e., they laughed out loud) when they read the question, and answered that they had not read the help pages.

While answering the four searching questions, the sixteen participants in the study issued a total of 262 queries. Overall, the study participants used an average of 4.9 terms for each query with a standard
deviation from the mean of 2.4 terms. Over 82 percent of the queries contained six terms or less. Figure 3 shows the distribution of query length for the entire data set.

![Figure 3: Number of Terms per Query for Entire Data Set](image)

Looking at the two groups, the eight participants who were using the traditional Google interface issued a total of 132 queries with an average number of 4.7 terms per query with a standard deviation of 2.0 terms. The smallest query consisted of one term while the largest query in the group had thirteen terms. Figure 4 shows the query distribution for the participants utilizing the traditional interface.
Figure 4: Number of Terms per Query for Regular Google Search

For the advanced search interface group, the numbers were slightly different. This group used 130 queries with an average term count of 5.1 terms with a standard deviation of 2.8 terms. The queries ranged in length from one term to seventeen terms. These results are shown in Figure 5.

Figure 5: Number of Terms per Query for Advanced Google Search
The first hypothesis stated that participants utilizing the advanced Google screen would create longer search queries than those using the regular Google interface. To test whether the advanced group had longer queries, the mean query lengths for each group were tested for a significant difference. A two-sample t-test of hypothesis one revealed a p-value of .07. Therefore one must reject that hypothesis with caution.

To test whether the advanced group reformulated less than the regular Google group, a t-test comparing the numbers of reformulations for each question among the two test populations was used. The decision was made to test each question separately since the land mine question was reformulated more than others in both sets. Since the variability within each group differed, the t-test for unequal variances was used. The only question that showed a significant difference was question one with a p-value of 0.03. All other questions had a p-value in excess of 0.30.

Due to a small data set, hypothesis three could not be formally tested. However, there were notable differences in the use of advanced operators between the regular Google users and the advanced Google users. None of the queries on the simple interface used the OR operator while fourteen queries, or 10.7%, issued on the advanced interface used
the OR. Nine of the queries (6.8%) on the simple interface employed phrase searching while 52 queries (40%) on the advanced interface used phrase searching. Only one query on the advanced interface used the exclusion operator while no queries on the simple interface used it. One search on each interface used the site limitation function within Google. Two searches on the advanced interface were limited to English only sites while no searches on the simple interface used that option. Further, searches on the advanced interface used the file type limitations, time limitations, and text only searches, but each was only used by one of those participants.

**Additional Observations**

While not part of the formal study, several ancillary findings are worth noting. Problems associated with phrase searching and a special syntax anomaly on the advanced search page became apparent during data analysis. These particular observations may inspire future studies to delve into these issues.

The misuse of phrase searching occurred with great frequency among those who used the phrase operator. Many users chose to enclose a single word in quotations. Since Google already uses keyword searching and does not employ truncation, use of quotations in this situation provides no help to the search. Of the 66 queries that utilized
phrase searching, thirteen (19.6%) of those contained at least one one-word phrase.

Interestingly, the advanced search interface seemed to cause one particular search string error. The advanced screen contains the “with at least one of the words” option to capture OR operations; several users chose to enter phrases in this area instead of solitary keywords. Therefore a phrase like “multinational treaty ban” would be parsed into “multinational OR treaty OR ban” by the Google search engine if the user did not use quotes in the advanced screen. This need for specific syntax is in direct contrast to the rest of the advanced interface. Based on the questions and other successful query strings, it is improbable that this is the search that the user desired. Clearly, the user would have to know some specialized syntax to keep a phrase from being searched as a string of terms joined with an OR.

Another problem that arose with the queries submitted by users concerned the excessive length of some queries. Google truncates queries at ten terms. Several users kept creating queries that exceeded this length despite a warning on the screen. This artificial constraint on the length of queries could skew the data collected for other studies since they relied on server logs which would not capture the extraneous terms instead of local browser logs which capture exactly what the user typed.
**Discussion**

Since the hypothesis that the mean number of search terms between groups would be different could not be clearly rejected, one must conclude that additional studies need to be conducted to determine if there is really a difference between the numbers of search terms between users of the two interfaces. With such a small sample set, one cannot discount that the means are not different with a $p$-value of 0.07.

Testing of hypothesis two did not provide conclusive results either. While three of the questions had non-significant results, question one provided a significant result. Thus, one has to conclude that, for some questions, the advanced search interface cuts down on the reformulations needed for the first question. However, the newness of the task could have contributed to this result since this only occurred on the first question. More research is called for to investigate the possible effects of question order and learning.

This study supports the third hypothesis that advanced search interface users construct queries that are more complex. In every category, the advanced users had more queries with special operators than the regular users.

In summary, there may be a difference in the length of users’ queries based upon which search engine interface used. The results form
this study indicate there may not be a difference, but the value is too
close to reject outright. More research is needed. For one of the
questions there was a difference in the number of reformulations based on
the interface used; however, the other three questions showed no
measurable difference in the number of reformulations. When looking at
query complexity, those users who were assigned to the advanced Google
screen showed a much greater use of special operators and modifiers.

**Conclusions**

There are several weaknesses in this study. The initial recruitment
goal was to include a large sample of INLS 40 students after they had
finished some introductory web searching instruction. Additionally the
study anticipated that each INLS 40 student would bring a non-INLS 40
student with him/her to diversify the population. However, recruitment
had to be expanded to other populations once it became evident that the
initial population was not very responsive to pleas for participation. With
this expansion, some of the desired balance for the study was lost. Even
with the recruitment issues, limiting such a study to college students
provides a skewed population sample. Ideally this study would have been
more inclusive of Internet users from all demographic groups.

The statistical analysis methods used for this paper assume that a
random unbiased sample was taken and that each observation (i.e., search)
was independent. Clearly, these assumptions have not been met by this study, primarily because there was no statistical tool that could easily accommodate the study design. Nevertheless, the results do provide a preliminary view of the effects of the search engine interface on users’ searching behaviors.

While searching for answers to the questions, several participants commented that they had never looked for answers to questions such as these. In a recent study of the users’ search interests, Computers/Networking ranked as the most popular category followed by Adult, Entertainment, Recreation/Chat and Shopping (Pu, Chaung, and Yang 2002). None of the questions asked would easily place terms in these categories. Therefore, these questions may be far out of the norm and could result in atypical queries. This deserves further research as well.

Unfortunately, no definitive conclusions can be drawn from this study. With its small sample size and skewed recruitment, it is not indicative of most user populations. However it does hint that there may be some merit to having users query using a screen that does not require any special syntax. Several of the study participants who were assigned to the advanced screen commented that they thought the advanced screen would be something quite complicated so they had refrained from using it
in the past. Therefore, I would suggest that another study be run comparing the two search engine interfaces. In addition to using the four questions tested here, new questions that cover areas more often searched should be introduced.

This research adds to the body of research supporting the need for better search interfaces. Unfortunately Google’s advanced screen is not a panacea to the problem of how to help users write the queries to best provide them with information they deem relevant. However we do now have another tool to explore how users develop queries. While we may be waiting for Godot, we are not waiting aimlessly.
Appendix A – Informed Consent Form

Introduction to the Study:

- We are inviting you to be in a research study of people searching for information on the World Wide Web.
- Ashley Langley, a Master’s student in the School of Information and Library Science of the University of North Carolina at Chapel Hill, is doing this study, under the supervision of Dr. Barbara Wildemuth.

Purpose:

- The purpose of this study is to see how users construct queries to different types of questions using a search engine for the World Wide Web.
- We hope to use what we learn from the study to make suggestions for later search engine development.

What Will Happen During the Study:

1. We will ask you to complete a one-page questionnaire with some information about you and your knowledge of search engines.
2. We will ask you to develop queries to find the answers to four questions.
3. This study should take between thirty minutes to an hour to complete.

If you have any questions or concerns about being in this study, you should contact Ashley at alangley@email.unc.edu, or her advisor, Dr. Barbara Wildemuth at wildemuth@ils.unc.edu.

Your Privacy is Important:

- We will make every effort to protect your privacy.
- We will not use your name in any of the information we get from this study or in any of the research reports.
- Any information we get in the study will be recorded with a code number that will not be tied back to you in any way.

Risks and Discomforts:

We do not know of any personal risk or discomfort you will have from being in this study.

Your Rights:

- You decide on your own whether or not you want to be in this study.
- If you decide to be in the study, you will have the right to stop being in the study at any time.

Institutional Review Board Approval:

- If you have any concerns about your rights as a participant, you may contact the Academic Affairs Institutional Review Board, which has approved this study, at (919) 962-7761, or at aa-irb@unc.edu.

I have had the chance to ask any questions I have about this study, and they have been answered for me. I have read the information in this consent form, and I agree to be in the study. There are two copies of this form. I will keep one copy and return the other to the investigator.

______________________________                                                  __________________________
(Signature of Participant)                                                                     (DATE)
Appendix B – Research Questionnaire Annotated with Response Totals

1. Identifier: __________

2. Age: __________

3. Gender:
   (9) Male
   (7) Female

4. Have you had formal training in Internet searching other than INLS 40?
   (3) yes
   (0) no
   (10) I haven’t taken INLS 40.

5. If so, where did you receive this training?
   (1) High school
   (0) Public library
   (0) UNC-CH library
   (2) UNC-CH class
   which class? __________

6. Which of the following search engines have you used?
   Yahoo!
   Excite
   Lycos
   Hotbot
   AltaVista
   Other

7. If you have a favorite search engine, what is it? __________

8. Have you ever read the “help pages” associated with a search engine?
   (6) Yes
   (10) No

9. How often do you use a search engine for something other than INLS 40?
   (0) Never
   (0) About Once a Semester
   (2) About Once a Month
   (4) About Once a Week
   (10) About Once a Day

10. On a scale of 1 to 5, how familiar are you with the Google search engine (http://www.google.com)?
    (0) 1 – Not Familiar at All
    (2) 2 – Vaguely Familiar
    (3) 3 – Somewhat Familiar
    (10) 4 – Familiar
    (1) 5 – Extremely Familiar

11. How often do you use Google?
    (0) Never
    (1) About Once a Semester
    (1) About Once a Month
    (8) About Once a Week
    (6) About Once a Day

Thank you for your participation in this study!
Works Cited


Siegfried, Susan, Marcia J. Bates, and Deborah N. Wilde. “A profile of end-user searching behavior by humanities scholars: The Getty Online


