A Temporal Comparison of AltaVista Web Searching

Bernard J. Jansen
School of Information Sciences and Technology, Pennsylvania State University, 329F Thomas Building, University Park, PA 16802. E-mail: jjansen@ist.psu.edu

Amanda Spink
School of Information Sciences, University of Pittsburgh, 610 IS Building, 135 N. Bellefield Avenue, Pittsburgh, PA 15260. E-mail: aspink@sis.pitt.edu

Jan Pedersen
Overture Web Search Division, 1070 Arastradero Road, Palo Alto, CA 94304. E-mail: jan.pedersen@overture.com

Major Web search engines, such as AltaVista, are essential tools in the quest to locate online information. This article reports research that used transaction log analysis to examine the characteristics and changes in AltaVista Web searching that occurred from 1998 to 2002. The research questions we examined are (1) What are the changes in AltaVista Web searching from 1998 to 2002? (2) What are the current characteristics of AltaVista searching, including the duration and frequency of search sessions? (3) What changes in the information needs of AltaVista users occurred between 1998 and 2002? The results of our research show (1) a move toward more interactivity with increases in session and query length, (2) with 70% of session durations at 5 minutes or less, the frequency of interaction is increasing, but it is happening very quickly, and (3) a broadening range of Web searchers' information needs, with the most frequent terms accounting for less than 1% of total term usage. We discuss the implications of these findings for the development of Web search engines.

Introduction

Web searching has become a daily behavior for many people, with the Web now the first choice for many people seeking information (Cole, Suman, Schramm, Lunn, & Aquino, 2003; Pew Internet Project, 2002). Given the Web's importance, we need to understand how Web search engines perform (Lawrence & Giles, 1998), how people use and interact with Web search engines, and the Web-searching trends that are emerging over time. Examining Web searching over time is an important area of research that has the potential to increase our understanding of Web searching, to advance our knowledge of Web searchers' information needs, and to positively impact the design of Web search engines.

To our knowledge, there has been limited large-scale research examining Web-searching changes or trends. There is a body of research focusing on the Excite search engine (Jansen, Spink, Bateman, & Saracevic, 1998; Spink, Jansen, Wolfram, & Saracevic, 2002; Spink, Wolfram, Jansen, & Saracevic, 2001; Wolfram, Spink, Jansen, & Saracevic, 2001), along with studies of a few other systems (Cacheda & Viña, 2001b; Hölscher, 1998; Silverstein, Henzinger, Marais, & Moricz, 1999). Although these studies provide important insights into Web searching, further research is needed that validates these results across search engines and across time. This is especially important because Web information systems are continually undergoing incremental, and sometimes radical, changes. Research is needed to evaluate the effect of these changes on system performance and on user searching behaviors over time.

We address this need in the present study by examining logged Web search sessions of AltaVista,[1] a major U.S. Web search engine. In this article, we analyze general searching characteristics and changes, including session duration, query length, results pages viewed, and term usage. We address temporal issues by comparing data we collected in 2002 with similar data collected in 1998 by Silverstein, Henzinger, Marais, and Moricz (1999). We describe our research design and our analysis of the AltaVista Web-search-engine data, followed by a discussion of results. We then discuss the key findings and the implications of our research

Received September 23, 2003; revised January 13, 2004; accepted March 1, 2004

 ● Published online 7 February 2005 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/asi.20145

[1] http://www.altavista.com

results for Web-search-engine users and designers. We conclude with directions for future research.

Related Studies

There is a growing body of research examining the use of Web search engines (Cacheda & Viña, 2001a; Hölscher & Strube, 2000; Jansen & Pooch, 2001; Jansen, Spink, & Saracevic, 2000; Montgomery & Faloutsos, 2001). Cacheda and Viña (2001a) report statistics from a Spanish Web directory service, BIWE.[2] The researchers report on page results, queries, query operators, and terms. Hölscher and Strube (2000) examine European searchers on the Fireball[3] search engine, a predominantly German search engine, reporting on the use of Boolean and other query operators. They note that experts exhibit different searching patterns than novices. Jansen and Pooch (2001) reviewed the Web-searching literature, comparing Web searchers with searchers of traditional information retrieval systems and online public access catalogues. The researchers report that Web searchers exhibit different search characteristics than do searchers of other information systems, and they call for uniformity in terminology and metrics for Web studies.

Jansen et al. (2000) conducted an in-depth analysis of the user interactions with the Excite[4] search engine, and reported that user sessions are short (i.e., few queries) and that Web queries are also short (i.e., few terms). Montgomery and Faloutsos (2001) analyze data from a commercial research service, also noting short sessions and queries. This stream of research provides useful snapshots of Web searching.

One limitation of these studies, however, is that they are snapshots with no temporal analysis comparing Web search engine usage over time. We could locate only two studies of major Web search engines that provided a temporal comparison. First, Spink, Jansen, et al. (2002) provided a four-year analysis of searching on the Excite search engine using three snapshots. They report that Web-searching sessions and query length have remained relatively stable over time, although they noted a shift from entertainment to commercial searching. The researchers show that on the Excite search engine, Web-searching sessions are very short, as measured by the number of queries. Users view a very limited number of result pages.[5] The majority of Web searchers, approximately 80%, view no more than 10 to 20 Web documents. These characteristics have remained fairly constant across the multiple studies.

Second, Jansen and Spink (2004) conducted a two-year study of AlltheWeb.com[6] users. The researchers noted even shorter sessions from this temporal analysis of searchers and a near total intolerance of viewing more than one results page. There has been little analysis of page-viewing characteristics of Web searchers at any finer level of granularity, although Jansen and Spink (2003) report that Web searchers of AlltheWeb.com view about five actual Web documents. The researchers also noted a shift toward commercial searching on AlltheWeb.com, although there is less of it than on the Excite search engine.

There are studies that examine searching on specific Web sites, rather than Web search engines. For example, Wang, Berry, and Yang (2003) analyzed 48 consecutive months of data from a university Web site. Analysis was at the query and term level. The researchers did not collect session level data. The results of the query analysis were similar to those reported in studies of Web search engines. The term analysis results were targeted, naturally, to the university domain rather than the more general searching environment of Web search engines.

These results are comparable to those obtained by Jones, Cunningham, and McNab (1998), who examined searches on a university online digital library over several months, and to results obtained by Croft, Cook, and Wilder (1995), who examined searches on a government Web site over several weeks. There are similarities in the results of these types of studies when compared with results from studies of major Web search engines, but there are also differences, due in part to distinctions in information content. Analyses of searching on Web search engines and individual Web sites are certainly complementary, but these are also distinct research areas.

We center our research analysis on the interactions between the user and the search engine. Interaction has several meanings in information searching, although the definitions generally encompass query formulation, query modification, and inspection of results lists, among others. Belkin and colleagues (1995) extensively explored user interaction within an information session and proposed 16 information-seeking strategies (ISSs) that users can employ. These strategies focus on what information the user wants, unlike the research reported here, which analyze what the user does to acquire the information.

Bates (1990) presents four levels of interaction: *move, tactic, stratagem,* and *strategy*. Using Bates's classification and definitions, this research primarily focuses on levels one and two (*move* and *tactic*) and provides glimpses of level three (*stratagem*). Efthimiadis and Robertson (1989) present and categorize interaction at various stages in the information retrieval process from information-seeking research. Beaulieu (2000) identifies three aspects of interaction: interaction within and across tasks, interaction as task sharing, and interaction as a discourse.

Saracevic (1997) views interaction as the exchange of information between user and system, with increases in interaction resulting from increases in communication content. Lalmas and Ruthven (1999) assert that interaction results in users learning how to search on a particular system, how to

[2]http://www.biwe.com/index.html

[3]http://www.fireball.de

[4]http://www.excite.com

[5]When a Web search engine user submits a query, the search engine returns the results in blocks, or *chucks*, usually of about 10 results. These are also referred to as *results pages* and are presented to the user sequentially from the highest-ranked results page to the maximum number of results retrieved by the search engine.

[6]http://alltheweb.com, a predominantly European search engine.

express their information needs, and how to perform relevance feedback. The researchers present two groups of interaction: that which occurs across sessions and that which occurs within a session. They group within-session interactions into *query formulation interaction* and *postquery interaction*. This within-session category is the type of interaction that we examine in this study.

In this research, we address the issue of how little temporal analysis of interactions on major Web search engines has been conducted. We investigate the session durations of AltaVista searchers. We also examine the information needs of AltaVista searchers. Finally, we present how or if the searching patterns of AltaVista Web searchers have changed over time.

Research Questions

More specifically, the research questions driving this study are

1. What changes in AltaVista Web searching occurred from 1998 to 2002?
2. What are the current characteristics of AltaVista searching, including the duration and frequency of search sessions?
3. What changes in the information needs of AltaVista users occurred from 1998 to 2002?

The broader goal of our study is to compare our findings with previous studies to identify overall changes that have occurred for Web searching on AltaVista and to present the current state of Web searching by AltaVista users.

Research Design

AltaVista

In 1998, AltaVista had a document collection of more than 140 million Web pages (Sullivan, 1998b) and nearly 21 million unique visitors[7] per month. AltaVista supported several query operators, including *AND, OR, NOT, NEAR, MUST APPEAR, MUST NOT APPEAR,* and *PHRASE* operators (Silverstein, Henzinger, Marais, & Moricz, 1999). In 1998, AltaVista was the fifth most popular search engine (Sullivan, 1998a).

In 2002, AltaVista was the ninth most popular search engine (Sullivan, 2002), had a content collection of 550 million Web pages (Sullivan, 2000), and had approximately 5.6 million unique visitors per month. The drop-in Web site traffic reflects the intense competition and consolidation that has occurred in the Web search engine industry. AltaVista supported the same query operators as in 1998 (AltaVista, 2003). We see from this information that AltaVista offers a full range of searching options, has an extremely large content collection, and has millions of unique visitors per month. After being an independent company for several years, Overture Services purchased AltaVista in 2003 (Morrissey, 2003). The drop-in number of visitors, an increase in the content collection, and modifications to algorithmic operations are environmental changes that may affect the comparison of the two studies over this five-year period.

At the time of this study, data from other search engines (e.g., Google, Inktomi) were not available.

Data collection. To address our research questions, we obtained and quantitatively analyzed actual queries submitted to AltaVista in 2002. Using our analysis from 2002 and the results from a previously published study of AltaVista searchers from 1998 (Silverstein et al., 1999), we can determine the general changes in Web searching on AltaVista over a period of approximately five years. In utilizing the 1998 research (Silverstein et al., 1999), we used the reported results in cases where the analysis was identical or comparable. In other cases, we calculated the results based on the data that the researchers reported.

The queries examined for this study were submitted to AltaVista over a 24-hour period on Sunday, September 8, 2002. We checked to see if any news stories from this day looked as if they may have influenced the investigation, namely the term analysis. There did not appear to be a major news stories occurring on this date. However, the date is near the one-year anniversary of the September 11, 2001, terrorist attacks.

We recorded the queries in a transaction log that represents a portion of the searches executed on the Web search engine on this particular date. The original general transaction log contains approximately 3,000,000 records. Each record contains three fields:

1. Time of day: measured in hours, minutes, and seconds from midnight of each day as recorded by the AltaVista server
2. User identification: an anonymous user code assigned by the AltaVista server
3. Query terms: terms exactly as entered by the given user

This record structure and format corresponds well with that of the 1998 study. Silverstein and colleagues (1999) used a transaction log with several fields including these three:

1. Time stamp: measured in milliseconds from January 1, 1970
2. Cookie: the cookie file name used to identify a user computer
3. Query: terms exactly as entered by the given use

The queries in the 1998 study were submitted to AltaVista during the period August 2 to September 13, 1998. The total

[7] *Unique visitors* is a standard metric in reporting traffic to Web sites and is based on the Internet Protocol address of the client computer visiting a Web site within a certain time period (in this case, one month). The number of unique visitors is a far more accurate measure of Web site traffic than a hit or a download.

transaction log contained 993,208,159 requests, just under a billion records. Of this total, nearly half were empty queries, so much of the reported results use the 575,244,993 nonempty queries (Silverstein et al. 1999).

Data analysis. The transaction log of the 2002 data is a flat ASCII file, which we imported into a relational database to conduct the analysis. We generated a unique identifier for each record. Using the three fields (time of day, user identification, and query terms) and the corresponding fields from the 1998 study (Silverstein et al., 1999), we located the initial query and then recreated the chronological series of actions in a session. In this study, we adopted the terminology outlined by Jansen and Pooch (2001):

- A *term* is any series of characters separated by white space or other separator.
- A *query* is the entire string of terms submitted by a searcher in a given instance.
- A *session* is the entire series of queries submitted by a user during one interaction with the Web search engine.
- An *initial query* is the first query submitted in a session.
- An *identical query* is a query within a session that is a copy of a previous query within that session.
- A *repeat query* is a query submitted more than once, irrespective of the user.

The transaction log contained searches from both human users and agents. We were interested in only those queries submitted by humans, rather than by some automated process. Because there is no way to accurately distinguish between human and nonhuman searchers, most researchers utilizing transaction logs for data collection must either ignore this classification (Cacheda & Viña, 2001a) or assume some temporal or interaction cutoff (Montgomery & Faloutsos, 2001; Silverstein et al., 1999).

We chose the latter approach by separating all sessions with 100 or fewer queries into an individual transaction log. We chose this cutoff because it is almost 50 times greater than the reported mean number of queries per search session (Jansen, Spink, & Saracevic, 2000) for human Web searchers, which assured that we were not excluding any human searches. Although this cutoff probably introduced some agent or common user terminal sessions, we were satisfied that we had retrieved a subset of the transaction log that contained queries submitted primarily by human searchers, yet was broad enough not to introduce bias by imposing a cutoff threshold that is too low.

When a searcher submits a query, views a document, then returns to the search engine, the AltaVista server logs this second visit with the same user identification and query, but with a new time (i.e., the time of the second visit). This information is beneficial in determining how many of the retrieved results pages the searcher visited from the search engine, but unfortunately it also introduces duplicate queries.

To address this duplication issue, we collapsed the four data sets by combining all identical queries submitted by the same user to give us the unique queries for analyzing sessions, queries and terms, and results pages viewed. We used the complete noncollapsed sessions to obtain an accurate measure of the session duration and the number of results pages visited. When we collapsed the sessions, we recorded the number of identical queries by the same user in a separate field within the remaining records.

In addition to the fields for unique identifier and number of identical queries, we included a field within each record containing the length of the query, measured in terms. We also generated two other tables for the collapsed data set, one for term data and one for co-occurrence data. The term table contains fields for a term and the number of term occurrences in the complete data set. The co-occurrence table contains fields for term-term pairs and the number of pair occurrences within the data set, irrespective of order.

Our database now contains four tables: noncollapsed data set, collapsed data set, terms, and co-occurrence). We analyzed the data from these four tables to investigate our three research questions. We conducted the analysis using queries, usually a series of layered queries, Visual Basic for Applications scripts, or a combination of the two. We report the results of our analysis in the follow section.

Results

Overall Results

We present the aggregate results for the analysis in Table 1.

Overall, we see a move toward greater interactivity between the user and the search engine. The percentage of three-term queries has increased from nearly 28% in 1998 to 49% in 2002. The proportion of users who modified queries increased by approximately 32 percentage points, from 20% of all users in 1998 to 52% in 2002. Spink, Ozmutlu, Ozmutlu, and Jansen (2002) had previously found that European AlltheWeb.com users were also more interactive than AltaVista users were in 1998.

Examining session length, the percentage of longer sessions increased, with 32% of users submitting three or more queries (see Table 1) per session in 2002 compared with 7% in 1998. The mean session length from the 2002 study is similar to results reported on the Excite (Jansen et al., 1998) and AlltheWeb.com (Jansen & Spink, 2004) search engines. There was an increase in the percentage of users viewing more than the first results page, which when combined with other increased interactions, may indicate greater user persistence in locating relevant results.

Overall usage of Boolean and other query operators was similar to results from both studies, about 20%. It is difficult to determine the changes for Boolean usage only because the 1998 study combined Boolean and other operators, reporting an overall usage of 20.4%. In the 2002 study, Boolean language was used for 6% of queries (61,065 queries), which is comparable to that reported by Spink, Jansen, et al. (2002) for the Excite search engine, but higher than that reported in searching on European search engines (Hölscher & Strube, 2000; Jansen & Spink, 2004).

TABLE 1. Aggregate results for data analysis of 1998 and 2002 AltaVista data sets.

	AltaVista 1998		AltaVista 2002	
Sessions	285,474,117		369,350	
Queries	993,208,159		1,073,388	
Terms				
Unique			369,350	9.5%
Total			1,073,388	100%
Mean terms per query	2.35 (sd = 1.74)		2.92 (sd = 1.91)	
Terms per query				
1 term	39,640,423	25.8%	218,628	20.4%
2 terms	39,947,713	26.0%	330,875	30.8%
3+ terms	42,406,034	27.6%	523,885	48.5%
Mean queries per user	2.02 (sd = 123.4)		2.91 (sd = 4.77)	
Users modifying queries	34,416,491	20.4%	193,468	52.4%
Session length				
1 query	119,228,559	77.6%	175,882	47.6%
2 queries	20,742,082	13.5%	75,343	20.4%
3+ queries	13,674,409	6.9%	118,125	32.0%
Results pages viewed				
1 page	718,615,763	85.2%	781,483	72.8%
2 pages	63,258,430	7.5%	139,088	13.0%
3+ pages	13,674,409	7.3%	150,904	14.1%
Boolean queries and other operators	202,614,464	20.4%	210,717	20.0%
Terms not repeated in data set			176,196	5.6%
Use of 100 most frequently occurring terms			592,699	18.9%

Notes: For the 1998 study, percentage and figures use the 575,244,993 nonempty queries.
The 1998 data from this and other tables is from or calculated from data reported in (Silverstein, Henzinger, Marais, & Moricz, 1999).

Unfortunately, we cannot examine the usage differences based on the term analysis because the 1998 study presented few term analysis results. Comparing the 2002 results with other published research, the use of unique terms (9.5%), the terms not repeated in the data set (6%), and the use of the most frequently occurring terms (19%) are similar to that reported for users of Excite (Spink, Jansen, et al., 2002) and AlltheWeb.com (Spink, Ozmutlu, et al., 2002).

In the following sections, we examine the results of our analysis in more detail at three levels of granularity: session, query, and term levels of analysis.

Sessions

Session length. We see from the results reported in Table 2 that there was a sharp decease in the percentage of one-query sessions and a corresponding increase in the percentage of longer sessions.

This development is counter to that reported in other analyses of search engines trends. For example, Spink, Jansen, et al. (2002) reported a move toward greater simplicity in searching with shorter sessions on the Excite search engine. Jansen and Spink (2004), in their analysis of European searching, note a similar inclination. Perhaps AltaVista users are most sophisticated in their searching characteristics. However, the number of one-query sessions in 1998 was much higher than that reported in similar Web studies from that time period (Hölscher, 1998; Jansen et al., 1998). The number of one-query sessions from the 2002 study is more in line with results from other studies within this time period (Cacheda & Viña, 2001a; Spink et al., 2002). This discrepancy, when compared with other studies, may indicate that the temporal cutoff (Silverstein et al., 1999) used may have been too low. He et al. (2002) report that the average Web session is approximately 15 minutes.

Session duration. Table 3 presents the session durations for the 2002 data set.

TABLE 2. Occurrences and percentages of session length for 2002.

Session length	1998		2002	
	Occurrences	Percentages(%)	Occurrences	Percentages(%)
1	119,228,559	77.6	175,882	47.62
2	20,742,082	13.5	75,343	20.40
3	6,760,382	4.4	40,445	10.95
4	6,914,027*	4.5	23,463	6.35
5			14,719	3.99
6			9,726	2.63
7			6,664	1.80
8			4,731	1.28
9			3,481	0.94
> = 10			14,896	4.03

Notes: For the 1998 study, percentage and figures use the 575,244,993 nonempty queries.
*Number and percentages are for sessions of 4 and more queries.

TABLE 3. Occurrences and percentage of session duration for 2002.

Session duration	Occurrences	Percentages(%)
Less than 5 minutes	264,492	71.6
5 to 10 minutes	22,374	6.1
10 to 15 minutes	12,573	3.4
15 to 30 minutes	18,794	5.1
30 to 60 minutes	13,259	3.6
1 to 2 hours	8,952	2.4
2 to 3 hours	4,317	1.2
3 to 4 hours	2,671	0.7
More than 4 hours	21,916	5.9

Note: For 1998 study, all sessions were 5 minutes or less by definition.

Silverstein et al. (1999) assigned a temporal cutoff of 5 minutes as the maximum session duration. For the analysis of the 2002 data set, we measured the session duration from the time the first query was submitted until the user departed the search engine for the last time and did not return.

With this definition of search duration, we can measure the total user time on the search engine and the time spent viewing the first and all subsequent Web documents, except the final document. This final viewing time is not available because the Web search engine server records the time stamp. Naturally, the entire time between visits from the Web document to the search engine may not have been spent viewing the Web document. However, this may not be a significant issue, as the results in Table 3 show a large percentage of very short session durations.

The mean session duration was 58 minutes and 10 seconds, with a standard deviation of 3 hours, 34 minutes, and 12 seconds. However, we see that the longer session durations skewed our result for the mean and masked significant details. Fully 81% of the sessions were less than 15 minutes. Perhaps more surprisingly, nearly 72% of the sessions were fewer than 5 minutes. This is substantially shorter than earlier reported research on Web session length (Cyber Atlas, 2002; He et al., 2002). He et al. (2002) estimated that the average Web session is approximately 15 minutes, based on analysis of Excite and Reuter transaction logs. The percentage of sessions of 5 minutes or less that we observed was nearly three times that reported for AlltheWeb.com searchers (26%) (Jansen & Spink, 2004). Assuming that session duration held constant, the 1998 study probably artificially shortened about 30% of the sessions by using a 5-minute cutoff.

Queries

Query length. Table 4 shows that queries composed of one, two, or three terms accounted for 67% of all queries in 1998 and 73% of all queries in 2002.

The percentage of one-term queries decreased by 5% during the same time period. For queries with three terms, there is a sharp decline in the frequency of occurrences, dropping to a minimal percentage for five-terms queries. The number of one-term queries was notably lower than has been reported elsewhere (Cacheda & Viña, 2001a; Spink, Ozmutlu, et al., 2002). As in the two other published temporal analyses (Jansen & Spink, 2004; Spink, Jansen, et al., 2002), query length moved slowly upwards.

TABLE 4. Query lengths.

Query length	1998		2002	
	Occurrences	Percentages(%)	Occurrences	Percentages(%)
0	31,650,880	20.6	301	0.03
1	39,640,423	25.8	218,628	20.4
2	39,947,713	26.0	330,875	30.8
3	23,046,758	15.0	244,777	22.8
4	19,359,276*	12.6	128,485	12.0
5			63,025	5.9
6			27,125	2.5
7			12,432	1.2
8			5,636	0.5
9			37,487	3.5
> = 10			4,617	0.4

Notes: For the 1998 figure, calculated based on 153,645,993 distinct queries only.

*Number and percentages are for queries of 4 and more terms.

Top queries. Table 5 displays the top repeat queries in the two data sets.

As the table shows, there are only four query terms that appear in the 25 most frequently searched in both 1998 and 2002 (yahoo, hotmail, porn, and sex). The percentage of total queries accounted for by the top 10 queries in 2002 were approximately half of the corresponding percentage in 1998. This would indicate a broadening of information needs, which confirms findings from other studies (Jansen & Spink, 2004; Spink, Jansen, et al., 2002). We see a decrease in sexual queries from 1998 to 2002 and an increase in searches for general entertainment and alternate information sources during the same time period. The decrease in searching for sexual sources parallels an increase of nonsexual content on the Web (Lawrence & Giles, 1999). Of course, other factors may have influenced this trend, including other methods of locating online sexual material. The trend to locate other information sources may correspond with the move toward increased e-commerce searching (Jansen & Spink, 2004; Spink et al., 2002).

There also appears to be greater use of search engines not to search for information but as a shortcut for navigation. Some Web users appear to submit the name of a particular Web site to the search engine and just click on the uniform resource locator (URL) in the results page rather than type the URL in the address box of the browser or locate a bookmark, favorite, or shortcut. If the Web page's URL appears in the search engine's first page of results, this method requires less effort than other methods of accessing a particular URL. To determine likely navigation queries, we assumed queries that were complete or partial URLs were navigation queries. We also submitted the remaining top repeat queries to

TABLE 5. Top 25 queries with frequency of occurrence and percentage of queries.

1998			2002		
Query	Occurrences	Percentages(%)	Query	Occurrences	Percentages(%)
sex	1,551,477	0.27	google	837	0.09
applet	1,169,031	0.20	yahoo	727	0.08
porno	712,790	0.12	ebay	720	0.08
mp3	613,902	0.11	sex	412	0.05
chat	406,014	0.07	yahoo.com	395	0.04
warez	398,953	0.07	dictionary	374	0.04
yahoo	377,025	0.07	hotmail	336	0.04
playboy	356,556	0.06	translator	324	0.04
xxx	324,923	0.06	hotmail.com	308	0.03
hotmail	321,267	0.05	thumbzilla	306	0.03
(non-ASCII query)	263,760	0.05	www.yahoo.com	305	0.03
pamela anderson	256,559	0.04	lyrics	278	0.03
p****	234,037	0.04	maps	276	0.03
sex	226,705	0.04	babelfish	267	0.03
porn	212,161	0.04	mapquest	264	0.03
nude	190,641	0.03	porn	260	0.03
lolita	179,629	0.03	kazaa	241	0.03
games	166,781	0.03	translate	238	0.03
spice girls	162,272	0.03	nfl.com	234	0.03
beastiality	152,143	0.03	literotica	232	0.03
animal sex	150,786	0.03	nfl	226	0.03
SEX	150,699	0.03	weather	219	0.02
gay	142,761	0.02	search engines	213	0.02
titanic	140,963	0.02	www.hotmail.com	211	0.02
bestiality	136,578	0.02	google.com	210	0.02

Note: For the 1998 figure, calculated based on 153,645,993 distinct queries only.

AltaVista, classifying the query as a navigation query if the URL of the top result contained only the query term(s), a domain suffix, and possibly a "www" prefix.

In the 2002 data set, the top 15 queries—*google, yahoo, ebay, yahoo.com, hotmail, hotmail.com, thumbzilla, www.yahoo.com, babelfish, mapquest, nfl.com, nfl, weather, www.hotmail.com*, and *google.com*—were very likely the result of using the search engine as a navigation tool. This search engine usage appears only three times in the 1998 data set (i.e., warez, yahoo, hotmail). Several of the other most frequent repeat queries were also most likely navigation-related queries (i.e., *ask jeeves, MSN, richard realms, sublime directory, www.google.com, warez*, and *babel fish*).

Page results viewed. Reviewing the analysis of results pages viewed in Table 6, there is a sharp decrease in the number of

TABLE 6. Results pages viewed.

Number of results pages viewed	1998		2002	
	Occurrences	Percentages(%)	Occurrences	Percentages(%)
1	846,213,351	85.2	781,483	72.8
2	74,490,612	7.5	139,088	13.0
3	29,796,245	3.0	60,334	5.6
4	42,707,951*	4.3	27,196	2.5
5			16,898	1.6
6			11,646	1.1
7			6,678	0.6
8			4,939	0.5
9			3,683	0.3
10			4,074	0.4
>10			17,324	1.6

Notes: For the 1998 figure, calculated based on 153,645,993 distinct queries only.
*Number and percentages are for results pages of 4 and more.

viewings between the first and second pages of results and between the second and the third pages of results, with very few users viewing more than four or five pages of results.

The percentage of AltaVista searchers who viewed only one page of results decreased from 85% in 1998 to 72% in 2002, and there was a corresponding increase in searchers who viewed more results pages. The overall number of results pages viewed was still quite low, with approximately 85% to 92% of users viewing no more than the first two results pages. As with users of other Web search engines (Cacheda & Viña, 2001a; Hölscher & Strube, 2000; Jansen & Spink, 2004; Spink, Jansen et al., 2002), AltaVista users appear to have a low tolerance for reviewing large numbers of results.

Terms

We present a term analysis in Table 7.

Term analysis was not available for the 1998 data set. From the general transaction log, we extracted the top terms, removing the terms without content (*and, or, de, la, le*). Table 7 presents the top 25 terms. The term analysis presents some patterns. First, even the most frequently occurring terms represent a small percentage of overall term usage. The most frequently used term (*free*) accounted for only approximately 0.6% of all term usage. Second, the occurrence of sexual terms (*sex, nude, porn*) was lower than some might expect. Third, there was a significant variety of terms, indicating the diverse information needs of AltaVista users.

TABLE 7. Top occurring terms and frequencies for 2002.

Term	Frequency	Percentages(%)
free	18,404	0.6
sex	7,771	0.2
pictures	7,713	0.2
new	7,468	0.2
nude	5,363	0.2
music	5,358	0.2
school	5,160	0.2
how	5,148	0.2
lyrics	5,006	0.2
home	4,872	0.2
pics	4,788	0.2
download	4,715	0.2
online	4,365	0.1
american	4,206	0.1
state	4,179	0.1
county	4,109	0.1
university	3,765	0.1
car	3,762	0.1
texas	3,644	0.1
real	3,587	0.1
games	3,527	0.1
software	3,495	0.1
art	3,493	0.1
map	3,434	0.1
florida	3,417	0.1
world	3,412	0.1
college	3,405	0.1
video	3,374	0.1
city	3,355	0.1
history	3,299	0.1
search	3,188	0.1
web	3,149	0.1
porn	3,118	0.1
sale	3,082	0.1

Silverstein et al. (1999) did not report a direct term analysis. However, they reported the occurrences of the top 25 queries. Removing the one non-ASCII query, these queries represent 2% (8,734,653 queries) of the 575,244,993 queries in the database. Three of the remaining 24 queries are two-term queries (*pamela anderson, spice girls*, and *animal sex*). If one splits these three queries by term and disregards capitalization, there are 20 terms representing 9,304,270 occurrences. Of these 20 terms from the 1998 study, only three (*nude, porn*, and *sex*) appear in the top 25 terms in 2002.

Term co-occurrence. Although a term analysis is useful, it is sometimes difficult to determine the specific usage of a term intended by a searcher outside the framework of a particular query. In these cases, a term co-occurrence (Leydesdorff, 1989) is more helpful in determining usage. We present in Table 8 term co-occurrences for the 2002 data set in a correlation matrix fashion.

From the term co-occurrence analysis, the predominance of searching clustered around *free* is apparent. All of the occurrences of *free* appear to be cost related. We do not see any other noticeable clustering of term pairs, although several pairs are two-term phrases (*clip art, september 11, puerto rico* and *las vegas*). This diffusion of term pairs reinforces our findings of the term and the query analyses that these Web users are searching for an increasing variety of topics and information. In fact, the variety of information needs is greater than that reported in other studies (Jansen & Spink, 2004; Spink, Jansen, et al., 2002). In a study of Excite users, Wolfram (1999) noted high clustering of several term pairs around entertainment that we do not see in this analysis.

Silverstein et al. (1999) report the co-occurrence of the top 10,000 terms from approximately 313,000,000 million queries. None of the term pairs from the 1998 study appears as a top pair in 2002. Silverstein et al. (1999) also report highly correlated phrases. Again, none of these phrases from 1998 appears as a term pair in 2002.

Topical Query Classification

We qualitatively analyzed a random sample of approximately 2,600 queries from the 2002 data set in 11 nonmutually exclusive, general topic categories developed by Spink and colleagues (2002). Two independent evaluators manually classified each of the queries independently. The evaluators then met and resolved discrepancies. Table 9 displays the evaluation results.

Queries related to people, places, or things accounted for nearly half of the queries, with commerce, travel,

TABLE 8. Frequency of term co-occurrence for top 25 terms for 2002.

	11	angeles	art	cards	carolina	diego	download	engine	erotic	estate	free	high	jersey	las	new	puerto	sex	state(s)	windows
11																			
clip			706																
free							613												
games																			
greeting				606															
las		810																	
music											563								
new													653						
north					871														
nude											789								
pics											1,098								
pictures											642								
porn											1,037								
real										2,927									
rico																756			
san						758													
school												1,925							
search								592											
september	827																		
sex											1,310								
state																			
stories									619								893		
united																		1,127	
vegas														950					
xp																			688
york															2,782				

employment, the economy, and computers, and the Internet or technology accounting for another 25% of the queries.

The other eight categories account for the remaining 25% of queries. Combined with the evidence from the term occurrence and term co-occurrence results, this analysis confirms survey data that the Web is now a major source of information for most people (Cole et al., 2003; Fox, 2002), that there is a trend toward using the Web as an economic resource and tool (Lawrence & Giles, 1999; Spink, Jansen, et al., 2002), and that people use the Web for an increasing variety of information tasks (Fox, 2002; National Telecommunications and Information Administration, 2002). It also confirms that the topics of sex and pornography are major topics for search engine users (Jansen & Spink, 2004; Spink, Jansen, et al., 2002).

Silverstein et al. (1999) did not conduct a topic analysis. However, we can use the top 25 queries that they do report and classify these into corresponding categories. Of the 20 queries, 56% (14) are related to sex or pornography, 24% (6) to entertainment or recreation, 16% (4) to computers, the Internet, or technology items, and 4% (1) to other topics (i.e., queries containing non-ASCII characters).

TABLE 9. Comparison of general topic categories for 2002.

Rank	2002 (2,603 English Queries)	Percentages(%)
1	People, places or things	49.27
2	Commerce, travel, employment or economy	12.52
3	Computers or Internet or technology items	12.40
4	Health or sciences (physics, math)	7.49
5	Education or humanities	5.07
6	Entertainment or recreation (music, TV, sports)	4.57
7	Sex or pornography	3.26
8	Society, culture, ethnicity or religion	3.11
9	Government (or military)	1.57
10	Performing or fine arts (i.e., ballets, plays, etc.)	0.69
		100.00

Discussion

AltaVista Web Searching

The goals of this research were to identify the characteristics of AltaVista Web searching, compare these characteristics with searching on other Web search engines, and isolate the changes that occurred with AltaVista searchers between 1998 and 2002.

Based on our analysis, there appears to greater interactivity between users and the search engine. The increased interactivity is welcome news for search engine developers, as it may indicate a move by Web searchers to more carefully refine their information needs. Interaction in information searching is a process of query formulation, reformulation,

inspection of results, and judgment of results (Belkin, Cool, Stein, & Theil, 1995). The increase in query and session lengths and the increase in the number of results pages being viewed indicate this greater interactivity. Overall, the interactions between Web searchers and systems are still relatively simple, as evidenced by the low use of query operators. There are indications that the use of query operators has little effect on search engine results (Eastman, 2002; Jansen & Eastman, 2003) or topical relevance (Eastman & Jansen, 2003), which may account for their low usage.

There is a sharp decrease in the number of results pages viewed, especially between the first and the second results pages and between the second and the third results pages; very few users view more than four or five pages of results. AltaVista users have a low tolerance for reviewing large numbers of results, although the trend is to view more results. Given that more than 70% of Web users use search engines to locate other Web sites (Alexa Insider, 2000), the implications are rather clear for content and service providers. Certainly for those publishing on the Web or engaged in Web e-commerce, the need to be ranked within the top 10 or 20 results remains critical to direct visitors to one's Web site.

For our second research question, the duration of the average session was 58 minutes and 10 seconds, with a standard deviation of approximately 3.5 hours. However, most sessions (81%) were less than 15 minutes in duration and nearly 72% of the sessions were less than 5 minutes. Although the frequency of interaction (i.e., the number of queries per session) may be increasing, the duration of these sessions is very short, and the interactions happen very quickly. This session duration of approximately 15 minutes seems common across search engines, with almost exact session durations reported for Excite users (He et al., 2002) and for AlltheWeb.com users (Jansen & Spink, 2003).

This observation presents a challenging predicament for search engines and Web information providers. First, Web search engine users are making quick judgments of retrieved results based on the description the search engines present in the results pages. However, these descriptions are usually very short, and they may not adequately reflect the content of the specific Web site.

Additionally, given the short session duration of about 15 minutes, AltaVista Web users are not spending much time on specific Web sites. In other research, AlltheWeb.com searchers spent about 30 seconds on Web sites (Jansen & Spink, 2003). For content providers and Web site designers, the fact that users spend little time reviewing or searching a site for relevant information indicates a need for very well-designed Web sites and relevant information on those sites that is extremely easy for users to locate.

For research question 3—What changes in the information needs of AltaVista users occurred from 1998 to 2002?—the range of information needs appears to have broadened. We base this conclusion on the high percentage of unique terms, the large number of terms not repeated in the data set, and the wide spread of query topical classifications. Other Web studies also report a trend toward broadening information needs (Jansen & Spink, 2004; Spink, Jansen, et al., 2002).

Only four query terms are common to both the 1998 and 2002 lists. This illustrates that the information needs of search engines users changed or became more diverse. The percentage of total queries accounted for by the top 10 terms in 2002 was approximately half of the percentage of the top 10 terms in 1998, indicating a broadening of information needs among AltaVista users. There was a decrease in sexual queries and an increase in both general entertainment and alternate information sources during the same time period. The broadening of information needs is a challenge for search engine designers as they attempt to provide relevant results to searchers in a timely manner.

At the term level of analysis, the most frequently occurring terms represented a small percentage of overall term usage. The most frequently used term (*free*) in 2002 accounted for only approximately 0.6% of all term usage. The use of sexual terms was extremely low, and the diversity of terms was quite large. From a term co-occurrence analysis, the only noticeable clustering was around cost (*free*). Again, this diversity reinforces our initial term analysis findings—that these Web users are searching for an increasing variety of information topics.

Implications

This study contributes to the Web searching literature in several important ways. First, the data come from users who submitted real queries and viewed actual Web pages. Accordingly, the data provide a realistic glimpse into how Web users search without the self-selection issues or altered behavior that can occur with lab studies or survey data. Second, the sample is quite large, with approximately 285,000,000 sessions from the 1998 study (Silverstein et al., 1999) and more than 350,000 sessions from 2002. Third, we obtained data from a very popular search engine, as measured by both document collection and the number of unique visitors, to ensure that our results were generalizable. Fourth, our analysis is one of the few trend comparisons of Web searching, providing valuable insight into the changing patterns of Web searching interaction.

Limitations

As with any research, there are limitations to our study. The sample data come from one major Web search engine, over a 24-hour period for the 2002 data set, introducing the possibility that the queries do not represent the queries submitted by the broader Web-searching population. We also compare a 24-hour sample with a sample of five weeks. However, Jansen and Pooch (2001) suggest that the characteristics of Web sessions, queries, and terms are very consistent across search engines. The number of viewers unique to AltaVista experienced a significant decline from 1998 to 2002, while the document collection underwent a fivefold

increase. These factors may affect the numerical data; however, it is reasonable to expect that the percentages would be fairly stable across the time period. Additionally, we compare the change analysis with two other publications that analyze Web-searching trends, highlighting similarities and differences. Web-searching change appears to be an incremental process, so we would not expect dramatic day-to-day changes.

We do not have information about the demographic characteristics of the users who submitted queries, so we must infer their characteristics from the demographics of Web searchers as a whole. In addition, because the data were collected during a one-day and a five-week period within the five-year period, the possibility of bias exists if Web-searching activity on these dates was not representative. However, a comparison of the collected body of Web transaction log research (Abdulla, Liu, & Fox, 1998; Cacheda & Viña, 2001a; Hölscher & Strube, 2000; Jansen & Spink, 2004; Montgomery & Faloutsos, 2001; Selberg & Etzioni, 1997; Spink, Jansen, et al., 2002) shows a great deal of similarity among Web searchers, indicating that particular dates have little effect on session length, query length, Boolean usage, and similar variables. However, particular dates do have an effect on the usage of popular terms (Searchwords.com, 2003; Wolfram, 1999; Wolfram, Spink, Jansen, & Saracevic, 2001). Finally, there were also differences in measures between the 1998 (Silverstein et al., 1999) and the 2002 data analyses, primarily with the definition of a session. We highlight this aspect in the results section where appropriate. With the exception of session definition, the remainder of the analysis was generally comparable.

Conclusion and Further Research

The results of our research provide important insights into the current state of Web searching and Web usage for information consumers, developers of search engines, and Web site designers. Concerning methods, the study illustrates that transaction log analysis is a viable method for analyzing real users interacting with real systems in the complex environment of the Web. This complex environment is difficult to recreate in a laboratory setting (Dumais, 2002).

There are several avenues for future research. Certainly, we need more analysis in this field on a wider variety of Web search engines, ideally on the most popular search engines, such as Microsoft Search, Google, America Online, and Yahoo!. However, access to the user data and the willingness of search engines to provide the access hampers this area of research. Additionally, because Web search engines introduce more changes in searching interfaces, researchers should continue to evaluate these changes to gauge their effect on Web searchers by using either transaction log analysis or laboratory studies. Finally, we must continue the trend analysis of Web searching to predict future behavior and identify future user needs.

Acknowledgments

We thank AltaVista for providing the Web search engine data set, without which we could not have conducted this research. We also thank the two anonymous reviewers whose comments significantly improved this manuscript.

References

Abdulla, G., Liu, B., & Fox, E. (1998). Searching the World-Wide Web: Implications from studying different user behavior. In Proceedings of the World Conference of the World Wide Web, Internet, and Intranet (pp. 1–8). Orlando, FL.

Alexa Insider. (2000). Alexa insider's page. Retrieved March 30, 2000, from http://insider.alexa.com/insider?cli=10

AltaVista. (2003). Special search terms. Retrieved May 16, 2003, from the AltaVista Web site: http://www.altavista.com/help/adv_search/syntax

Bates, M.J. (1990). Where should the person stop and the information search interface start? Information Processing and Management. 26(5), 575–591.

Beaulieu, M. (2000). Interaction in information searching and retrieval. Journal of Documentation, 56(4), 431–439.

Belkin, N., Cool, C., Stein, A., & Theil, S. (1995). Cases, scripts, and information-seeking strategies: On the design of interactive information retrieval systems. Expert Systems With Applications, 9(3), 379–395.

Cacheda, F., & Viña, Á. (2001a). Experiences retrieving information in the World Wide Web. Proceedings of the 6th IEEE Symposium on Computers and Communications (pp. 72–79). Hammamet, Tunisia.

Cacheda, F., & Viña, Á. (2001b). Understanding how people use search engines: A statistical analysis for e-business. Proceedings of the e-Business and e-Work Conference and Exhibition 2001 (pp. 319–325). Venice, Italy.

Cole, J.I., Suman, M., Schramm, P., Lunn, R., & Aquino, J.S. (2003, February). The UCLA Internet report surveying the digital future year three. Retrieved February 1, 2003, from UCLA Center for Communication Policy Web site: http://www.ccp.ucla.edu/pdf/UCLA–Internet–Report–Year–Three.pdf

Croft, W., Cook, R., & Wilder, D. (1995). Providing government information on the Internet: Experiences with Thomas. Proceedings of the Digital Libraries Conference (pp. 19–24). Austin, TX.

Cyber Atlas. (2002). November 2002 Internet usage stats. Retrieved January 1, 2003, from the Nielsen//NetRatings Inc. Web site: http://cyberatlas.internet.com/big_picture/traffic_patterns/article/0,,5931_1560881,00.html

Dumais, S.T. (2002, May 7–11). Web experiments and test collections. Retrieved April 20, 2003, from http://www2002.0rg/presentations/dumais.pdf

Eastman, C.M. (2002). 30,000 hits may be better than 300: Precision anomalies in Internet searches. Journal of the American Society for Information Science and Technology, 53(11), 879–882.

Eastman, C.M., & Jansen, B.J. (2003). Coverage, ranking, and relevance: A study of the impact of query operators on search engine results. ACM Transactions on Information Systems, 21(4), 383–411.

Efthimiadis, E.N., & Robertson, S.E. (1989). Feedback and interaction in information retrieval. In C. Oppenheim (Ed.), Perspectives in information management (pp. 257–272). London: Butterworths.

Fox, S. (2002, July). Search engines. Retrieved October 15, 2002, from the Pew Internet & American life project Web site: http://www.pewinternet.org/reports/toc.asp

He, D., Göker, A., & Harper, D.J. (2002). Combining evidence for automatic Web session identification. Information Processing and Management, 38(5), 727–742.

Hölscher, C. (1998). How Internet experts search for information on the Web. Proceedings of the World Conference of the World Wide Web, Internet, and Intranet (pp. 1–6). Orlando, FL.

Hölscher, C., & Strube, G. (2000). Web search behavior of Internet experts and newbies. International Journal of Computer and Telecommunications Networking, 33(1-6), 337–346.

Jansen, B.J., & Eastman, C.M. (2003). The effects of search engines and query operators on top ranked results. Proceedings of the IEEE 4th International Conference on Information Technology (pp. 135–139). Las Vegas, NV.

Jansen, B.J., & Pooch, U. (2001). Web user studies: A review and framework for future work. Journal of the American Society of Information Science and Technology, 52(3), 235–246.

Jansen, B.J., & Spink, A. (2003). An analysis of Web information seeking and use: Documents retrieved versus documents viewed. Proceedings of the 4th International Conference on Internet Computing (pp. 65–69). Las Vegas, NV.

Jansen, B.J., & Spink, A. (2004). An analysis of Web searching by European Alltheweb.Com Users. Information Processing and Management, 41(6), 361–381.

Jansen, B.J., Spink, A., Bateman, J., & Saracevic, T. (1998). Real life information retrieval: A study of user queries on the Web. SIGIR Forum, 32(1), 5–17.

Jansen, B.J., Spink, A., & Saracevic, T. (2000). Real life, real users, and real needs: A study and analysis of user queries on the Web. Information Processing and Management, 36(2), 207–227.

Jones, S., Cunningham, S., & McNab, R. (1998). Usage analysis of a digital library. Proceedings of the Third ACM Conference on Digital Libraries (pp. 293–294). Pittsburgh, PA.

Lalmas, M., & Ruthven, I. (1999). A framework for investigating the interaction in information retrieval. Proceedings of 9th European-Japanese Conferences on Information Modeling and Knowledge Bases (pp. 222–239). Iwate, Japan.

Lawrence, S., & Giles, C.L. (1998). Searching the World Wide Web. Science, 280(3), 98–100.

Lawrence, S., & Giles, C.L. (1999). Accessibility of information on the Web. Nature, 400, 107–109.

Leydesdorff, L. (1989). Words and co-words as indicators of intellectual organization. Research Policy, 18, 209–223.

Montgomery, A., & Faloutsos, C. (2001). Identifying Web browsing trends and patterns. IEEE Computer, 34(7), 94–95.

Morrissey, B. (2003, February 18). Overture to buy AltaVista. Retrieved 16 May, 2003, from the Internet advertising report Web site: http://www.internetnews.com/IAR/article.php/1587171.

National Telecommunications and Information Administration. (2002). A nation online: How Americans are expanding their use of the Internet. Washington, DC: U.S. Department of Commerce.

Pew Internet Project. (2002, July 3). Search engines. Retrieved from the Pew Internet Research Web site: http://www.pewinternet.org/reports/toc.asp?Reports=64

Saracevic, T. (1997). Extension and application of the stratified model of information retrieval interaction. Proceedings of the Annual Meeting of the American Society for Information Science (pp. 313–327), Washington, DC.

Searchwords.com. (2003). Top search words. Retrieved May 30, 2003, from http://www.searchwords.com

Selberg, E., & Etzioni, O. (1997). The metacrawler architecture for resource aggregation on the Web. IEEE Expert, 12(1), 11–14.

Silverstein, C., Henzinger, M., Marais, H., & Moricz, M. (1999). Analysis of a very large Web search engine query log. SIGIR Forum, 33(1), 6–12.

Spink, A., Jansen, B.J., Wolfram, D., & Saracevic, T. (2002). From e-sex to e-commerce: Web search changes. IEEE Computer, 35(3), 107–111.

Spink, A., Ozmutlu, S., Ozmutlu, H.C., & Jansen, B.J. (2002). U.S. versus European Web searching trends. SIGIR Forum, 32(1), 30–37.

Spink, A., Wolfram, D., Jansen, B.J., & Saracevic, T. (2001). The public and their queries. Journal of the American Society for Information Science and Technology, 52(3), 226–234.

Sullivan, D. (1998a). Ratings of most visited search engines. Retrieved May 16, 2002, from the Searchenginewatch.com Web site: http://searchengineguide.org/classi2.htm

Sullivan, D. (1998b, April 30). Search engine sizes scrutinized. Retrieved 16 May, 2003, from the Searchenginewatch.com Web site: http://www.searchenginewatch.com/sereport/article.php/2166151

Sullivan, D. (2000). Search engine sizes. Retrieved August 30, 2000, from http://searchenginewatch.com/reports/sizes.html

Sullivan, D. (2002, 23 Feb). Nielsen / Netratings search engine ratings. Retrieved January 6, 2002, from the SearchEngineWatch.com Web site: http://www.searchenginewatch.com/reports/netratings.html

Wang, P., Berry, M., & Yang, Y. (2003). Mining longitudinal Web queries: Trends and patterns. Journal of the American Society for Information Science and Technology, 54(8), 743–758.

Wolfram, D. (1999). Term co-occurrence in Internet search engine queries: An analysis of the Excite data set. Canadian Journal of Information and Library Science, 24(2/3), 12–33.

Wolfram, D., Spink, A., Jansen, B.J., & Saracevic, T. (2001). Vox populi: The public searching of the Web. Journal of the American Society of Information Science and Technology, 52(12), 1073–1074.