

**THE PERCEIVED AFFORDANCES OF
WEB SEARCH ENGINES:
A COMPARATIVE ANALYSIS**

by

Barbara M. Wildemuth* and Andrea Rene' Carter†

*School of Information and Library Science,
University of North Carolina at Chapel Hill

†North Carolina A&T State University

SILS Technical Report 2002-02

**School of Information and Library Science
University of North Carolina at Chapel Hill**

INTRODUCTION

What search capabilities are available in a Web search engine and how does a searcher find out about them? For the vast majority of people, their understanding of the capabilities of the search engine they are using has been formed through trial and error rather than through systematic study of its capabilities. Such exploratory methods of learning can be greatly aided if the designers of the search engine have made the available features apparent in the interface.

One way to evaluate the interfaces of search engines is to analyze the perceived affordances offered by each. Gibson (1977) developed the idea of affordances, in relation to the physical world and the objects in it. In his words,

“If an object that rests on the ground has a surface that is itself sufficiently rigid, level, flat, and extended, and if this surface is raised approximately at the height of the knees of the human biped, then it affords sitting-on. We call the object a seat, stool, bench, or chair. It affords support for the rump, whether or not it affords support for the back. If these five properties coexist the object is in fact sit-on-able; they combine to yield a higher-order property without much attention being paid to the five properties in isolation.” (68)

This concept can be translated into the virtual world of computer systems, with little modification. This translation requires that virtual objects, such as buttons, sliders, icons, etc., be understood as affording particular functions. In other words, the virtual objects must be perceived by the potential user as affording particular behaviors. Effective interface design relies on the assumption that the user can understand the system’s functionality based on the user’s perceptions of the system’s affordances. From the perspective of interface usability, those actions that the user *believes* can be performed are more important to accomplishing the user’s goals than the *actual* features available in the system (Norman, 1999).

In the ideal world, system users would accurately perceive the entire set of affordances represented in the interface and, thus, have access to the full functionality of the system. However, in the real world, the affordances of a particular interface object may or may not be accurately perceived by the user. For example, a Web site user may believe that a piece of blue underlined text affords clicking (and a consequent jump to another page or site), while the author of the Web site may have used the blue underlining only to emphasize the text. This situation is an example of a *false* affordance. Affordances may also be *hidden* if the system has a particular

function, but the availability of this functionality is not visible to the user (Fitzgerald & Goldstein, 1999).

Perceived affordances are part of everyday life, and may play a key role in a searcher's success with a Web search engine. For example, if a search engine provides one text box that is only 10 characters long, users may perceive that it affords the searching of terms that are 10 characters or less; by contrast, if a search engine provides multiple text boxes, each of which is 25 characters long, users may perceive that each query may (should?) include multiple terms that can be long phrases. This study analyzed and compared the perceived affordances of nine of the most popular Web search engines.

METHODS

The analysis was begun in July 2000, with all data reviewed and updated again in September 2001. Nine search engines were selected for inclusion in this study based primarily on two criteria: database size and audience reach. These search engines were AltaVista, Ask Jeeves, Excite, Google, Hotbot, LookSmart, Lycos/Open Directory, Northern Light, and Yahoo. If a search engine had both basic and advanced interfaces, each was analyzed. The criteria for analysis included characteristics of the text box for entering terms, characteristics of the search button, search syntax, the availability and placement of help for entering search terms, methods for limiting the search results, support for modifying a query, features of the directory structure, characteristics of results displays, and methods for setting user preferences. The analysis was conducted by directly examining the interface of each search engine for each feature or characteristic. Each of these aspects of the research methods are described in more detail below.

Selection of the search engines

Two criteria were used to select the search engines to be included in the analysis. The first was the size of the database indexed by the search engine. Why is a search engine's database size so important? When a surfer is using the Internet, he or she wants to be able to find the relevant items. It doesn't matter if a search tool has outstanding relevance ranking technology or a variety of search features and capabilities if the most relevant items are not in its database. The larger the database, the more likely it is that the searcher will be able to find information, particularly if the search focuses on subjects that are somewhat obscure.

Most search engines report the size of their own databases, as summarized by Sullivan (2001b). Notess (2001) went further, conducting searches whose results were then used to

estimate the total size of each database (see Table 1). Based on the data on search engine size, Google, Inktomi (used for iWon), AltaVista, Fast, Northern Light, Hotbot, Snap, and Excite were all candidates for further analysis. In addition, the largest directories were added to the list of possibilities. These include AskJeeves, at 128 million links, Lycos/Open Directory, at 2.6 million links, LookSmart, at 2.5 million links, and Yahoo, at 1.8 million links (Sullivan, 2001a).

Table 1. Database size and audience reach of search engines/directories included in the study

	DATABASE SIZE (MILLIONS)			AUDIENCE REACH		
	Pages indexed ¹ (Notess)	Pages indexed ² (Sullivan)	Directory links ²	MediaMetrix ³ (Aug 2001)	NetRatings ³ (Dec 2000)	StatMarket ³ (Apr 2000)
AltaVista	331	550		7%	8%	18%
Ask Jeeves			128	10%	3%	1%
Excite	159	380		15%	10%	7%
Google	355	1,000		17%	6%	6%
Hotbot	280				3%	
LookSmart			2.5	6%	8%	
Lycos/Open Directory			2.6	30%	12%	4%
Northern Light	282	350			1%	
Yahoo			1.5	65%	48%	53%

¹ Data estimated as of July 2000 (Notess, 2001).

² Search engine size from Sullivan (2001b); directory size from Sullivan (2001a).

³ Audience reach data from Sullivan (2001c, 2001d, 2000, respectively).

The second criterion used to select the search engines to be analyzed was their audience reach. Audience reach can be defined as “the percentage of web surfers estimated to have visited each search engine or portal during the month” (Sullivan, 2001c). Audience reach data from three different sources are provided in Table 1. Jupiter MediaMetrix (Sullivan, 2001c) and Nielsen/NetRatings data (Sullivan, 2001d) are derived from the navigation behaviors of panels of thousands of Web users, and include all visits to the sites listed (whether for the purpose of searching or for some other purpose). The StatMarket data (Sullivan, 2000) is based on searches conducted on sites that use the HitBox tracker; thus, it is search-specific but may underestimate search activity by missing searches if the site has not installed HitBox on all its pages. Based on the audience reach data available from these sources, candidate search engines and directories include Yahoo, AltaVista, Excite, Infoseek/Go, Lycos, Snap, GoTo, MSN, Netscape, and NBCi.

Four search engines or directories were included in the analysis based on *both* criteria (size and audience reach): Yahoo, AltaVista, Excite, and Lycos/Open Directory. LookSmart was also included, because of its large directory size. Ask Jeeves, Google, Hotbot, and Northern Light were also included because they meet at least one of the criteria (size) and also have some unique aspects to their interface designs.

Selection of interface objects to be analyzed

In order for users to be aware of a particular feature of a search engine, that functionality must be represented by some object in the interface. The current study was oriented to the search engines' perceived affordances as represented by these interface objects. The set of interface objects to be included in the analysis was empirically derived. As one search engine was analyzed, a comprehensive list of its interface objects (excluding those unrelated to searching) was developed. Each other search engine interface was then analyzed based on this list. If a new object was encountered in an interface, it was added to the list and *all* search engines were analyzed in relation to that interface object. Many of the search engines have multiple interfaces (usually described as “basic” and “advanced”). Both interfaces were included in the analysis.

The final set of interface objects included in this analysis included:

Characteristics of the text box for entering the search: location on the Web page, length of the box (in number of n's that will fit in the box), whether the text box scrolls (i.e., is expandable), ability to enter multiple search terms (number of terms/text boxes, whether the number can be expanded)

Characteristics of the “search” button: type and form of object (button, hyperlink), location in relation to the text box, title

Search syntax: phrase searching, negation, proximity operators, word stemming, case sensitivity, stop words

Availability/placement of help for entering search terms: help link title, button vs. hyperlink, distance to help page from search page (in number of links), whether query examples are provided (yes/no)

Methods for limiting a search: by collection, by language (number of languages available, position of English on that list), by source location (yes/no, number of locations available), by media type (yes/no, number of media types), by field (title, text, url, person), by date, by Internet domain, by depth of a page within a site (top level, number of levels, all pages)

Support for modifying a query: relevance feedback for modifying a query, searching within the retrieved set, suggestions of alternative query terms, suggestions of alternative search engines

Combinations of search engines and directories: availability of a directory of links, emphasis on directory versus search

Characteristics of the results display: whether the total number of hits retrieved is reported, the order in which the retrieved items are presented, which fields were displayed for each item, the number of retrieved items displayed on each page, whether multiple hits from a site were grouped, whether the query terms were highlighted in the display

Setting of preferences for searching and displaying results: type and form of object for access to this capability, the types of preferences that can be specified

As noted above, each of the nine search engines was analyzed for these perceived affordances. Both “basic” and “advanced” interfaces were analyzed, when available. The analysis was first completed in the summer of 2000, but was completely updated in September 2001.

RESULTS

Characteristics of the textbox and search button

The most critical object in a search engine interface is the text box in which the query is to be entered. This object is universally used for this purpose. In all the interfaces examined, the text box for entering queries was near the top of the page and in the center or just left of center. The text boxes varied in length (see Table 2), ranging from 20 to 140 characters (mean = 42 characters). All of the text boxes allowed longer entries, with horizontal scrolling.

Table 2. Characteristics of the text box(es) and search button(s) for entering the search

	Textbox length	# of search buttons	Search button color	Search button shape	Search button title
AltaVista	53	1	Grey	Rectangle	Search
AltaVista – advanced	140	1	Grey	Rectangle	Search
AltaVista – Search Asst	87	1	Grey	Rectangle	Search
Ask Jeeves	50	1	Red	Oval	Ask
Excite	22	2	Grey Grey	Rectangle Rectangle	Search Zoom In
Excite – advanced	20	1	Grey	Rectangle	Search
Google	50	2	Grey	Rectangle	Google Search, I'm Feeling Lucky
Google – advanced	43	3	Grey Grey Grey	Rectangle Rectangle Rectangle	Google Search Search Search
Hotbot	35	1	Lime	Rounded rect.	SEARCH
Hotbot – advanced	32	2	Lime Grey	Rounded rect. Rectangle	SEARCH SEARCH
LookSmart	28	1	Grey	Rectangle	Search
Lycos	30	1	Grey	Rectangle	Search
Lycos – advanced	32	1	Gold	Rounded rect.	Go Get It!
Northern Light	40	1	Blue	Rounded rect.	Search
Northern Light – power	40	1	Blue	Rounded rect.	Search
Yahoo	30	1	Grey	Rectangle	Search
Yahoo – advanced	30	1	Grey	Rectangle	Search

All but two of the search engines provide only one textbox for entering a query. Excite advanced provides three textboxes, with the capability to expand that number indefinitely (empirically tested up to 25 textboxes). Hotbot advanced provides two textboxes, and the user can expand that number up to four.

Most of the search engines had a single button for executing the query. However, some (Excite, Google, Google advanced, and HotBot advanced) had more than one button. The most common form of the button was the standard grey rectangle, titled “Search.” Some search engines used color and shape to make this button more noticeable. A few (notably Ask Jeeves, Excite, Google, and Lycos advanced) also used different titles for the search button. The search button was almost always placed just to the right of the text box. The only alternative placement was at the bottom of the search form (in AltaVista advanced and Search Assistant, Excite advanced, and the second button in HotBot advanced).

Search syntax

The basic syntax associated with Web searching has quickly become relatively standard. A plus symbol (+) is used to indicate that a term *must* appear in each item retrieved; a minus symbol (-) is used to indicate that items containing the term should be excluded; and enclosing quotes (“phrase”) are used to indicate that items should only be retrieved if they include the exact phrase. Unless the search engine provides examples of this syntax on the search page (as is the case with Ask Jeeves and Lycos advanced) or the user moves to the “help” pages, this syntax is not represented to the user via the interface.

A few of the search engines do use other interface objects to represent these search capabilities to their users. AltaVista Search Assistant, Hotbot, Hotbot advanced, and Lycos advanced provide drop-down menus for the choices of “all the words,” “any of the words,” or “exact phrase”. Google advanced provides a separate textbox for each of these alternatives, plus one for negation (“without the words”). Excite advanced takes a unique approach; it provides a checkbox for the user to designate that the words in the textbox should be “in this exact order”, and provides radio buttons for specifying the importance of a set of terms and if they should be excluded from the search results.¹ AltaVista advanced requires Boolean operators for specifying these search capabilities, and also includes the proximity operator NEAR.

One search engine, Hotbot advanced, allows users to ask for word stemming or not (using a checkbox). Case sensitivity and the handling of stop words were also examined, though the search engines’ capabilities in these areas were not represented explicitly in any interface. Only AltaVista advanced was case sensitive. To test the handling of stop words, “a” and “the” were searched in each interface. Ten of the search engines processed them as search terms, and seven treated them as stop words (AltaVista Search Assistant, Excite advanced, Google advanced, Hotbot, Hotbot advanced, Yahoo, and Yahoo advanced).

Availability of help/search assistance

Almost all the search engines (excepting only LookSmart) provide some sort of assistance to the user in formulating a search strategy and using the correct query syntax. However, there is little standardization in the ways in which access to this help is represented in the interface (see Table 3).

¹ In Web forms, a set of check boxes allows multiple items to be selected from the set of options while a set of radio buttons forces the user to select only one of the options at a time.

Table 3. Availability of help in formulating a search strategy

	Title of link	Location of “help” link	# of clicks to access help
AltaVista	Help	above textbox	2
AltaVista – advanced	Help	above textbox	2
AltaVista – Search Asst	Help	above textbox	1
Ask Jeeves	Help	top right of page	1
	Tips	under textbox	1
Excite	More	under button	2
Excite – advanced	Help	top right of page	2
Google	Jobs, Press, Cool Stuff	bottom right of page	3
Google – advanced	Advanced Search Tips	top right of page	1
	All About Google	top right of page	3
Hotbot	Help	bottom left of page	2
Hotbot – advanced	Help	bottom left of page	2
Lycos	Help	bottom left of page	3
Lycos – advanced	Help	right of button	1
Northern Light	Help	left column	2
	Tips	right of button	1
Northern Light – power	Help	left column	2
	Tips	right of button	2
Yahoo	Help	top right of page	3
Yahoo – advanced	Search Tips	above button	1
	Advanced Search Syntax	above button	1
	Help	right of button	2

Most of the search engines provide help (i.e., suggestions for formulating successful search strategies) via a hyperlink titled “Help,” “Tips,” or “Search Tips”. However, alternate titles are also used, e.g., “All About Google” in Google advanced. The titles used in Excite (“More”) and Google (“Jobs, Press, Cool Stuff”) seem somewhat less likely than the alternatives to encourage users to follow those links when they need assistance. In almost all cases, the “help” link was the title text augmented with a hyperlink; however, Hotbot uses a button for this link and Yahoo uses a button augmented with hyperlinked text just below it.

Finding the “help” link on a search page was sometimes straightforward, e.g., when it is placed near the textbox or search button. Some search engines (Ask Jeeves, Excite advanced, Yahoo) place the “help” link in the upper right corner of the page, where it is somewhat less obvious but is visible without scrolling. A few of the search engines (Hotbot, Hotbot advanced, Lycos) place the “help” link at the bottom of the page, where it requires scrolling before it is

visible.² In five of the search engines, there were multiple links to search help. In each instance, each link went to a different Web page providing slightly different help. In just over a third of the search engine interfaces, clicking on a “help” link takes the user directly to a page that provides assistance in formulating a search strategy. However, in some search engines, two or three clicks are required to reach such a page.

The final characteristic of “help” access was whether example terms were provided on the same page on which the query would be entered (data not shown in table). There were only two interfaces (Ask Jeeves and Lycos advanced) that provided such examples. In all other cases, the user needed to leave the search page to see examples of queries.

Limiting the search results

Search engines provide a variety of ways in which to limit the search scope. Those included in this study allow users to limit their searches by collection, by language, by source location, by date, by media type, by field, and by the depth of a retrieved page within a site. The search engines analyzed varied in their offering of these search capabilities. If the capability was clearly available via an object in the interface, it is listed in Table 4 (next page).

Many of the search engines allow the search to move to a particular “collection” in which the search will be conducted. Hyperlinks from the collection names and radio buttons were the most common ways of making this feature visible in the interface. The most common collections include Web sites, usenet groups, and multimedia (providing one way to limit a search to a particular media type, such as photographs). The collections in Google advanced are topic-oriented (e.g., “Linux” or “universities”). In addition to the three collections provided in Northern Light, there are five additional, topic-oriented search pages, e.g., Investext Search or Geosearch, that are available via hyperlinks.

Many of the search engines allow the search to be limited to pages in a particular language. Almost always, the list of available languages is presented as a drop-down menu from which one or more languages can be selected; in Lycos advanced, it is presented as a list with radio buttons (allowing only one language to be selected for each search). The number of languages ranges from 6 to 29. In some, English is listed as the first option, while in others, it is listed in its alphabetical order. In some of the search engines (AltaVista basic and advanced and Google basic and advanced), the user can establish a “preference” for limiting their searches by

² While the Google “help” link is placed at the bottom of the page, no scrolling is required to see it.

language. In HotBot, any of the advanced features (including limiting the search by language) can be moved to the user's customized "basic" search page.

Table 4. Mechanisms for limiting a search

	By collection		By language		By source location	By field
	Number	Interface object	Number	Position of English	Number	Field names
AltaVista	7	Hyperlinks	26	First		
AltaVista – advanced	5	Hyperlinks	26	First		
AltaVista – Search Asst					10	Title, url, domain
Ask Jeeves			*			
Excite	4	Radio buttons				
Excite – advanced			12	First		url, domain
Google	4	Tabs				
Google – advanced	5	Hyperlinks	29	Alpha		Title, text, url, domain
Hotbot	6	Hyperlinks	12	Alpha		
Hotbot – advanced			12	Alpha	16	Title, url, domain, person
Lycos	9	Menu of hyperlinks				
Lycos – advanced	15	Radio buttons	25	Alpha		Title, url, domain
Northern Light	3	Radio buttons				
Northern Light – power	3	Radio buttons	6	First	26	Title, url
Yahoo	5	Menus of hyperlinks				
Yahoo – advanced	11	Drop-down menu				

* Spanish version of search engine available.

Only three of the search engines have an explicit mechanism for limiting the user's search by the location of the pages to be retrieved. The AltaVista Search Assistant and Hotbot advanced interfaces use a drop-down menu to allow the user to specify the region from which pages should be retrieved. The additional regions available in Hotbot are, in reality, specific domains within the United States. In both search engines, the capability to limit a search by region is provided in addition to searches by domain. Northern Light Power Search uses a drop-

down menu to show users that they can search for “Web sites from” a particular country. No additional capability for searching by domain is explicitly provided.

The date of the retrieved page can be used to limit the retrieved set in some search engines. Most of these emphasize the recency of the retrieved documents, such as limiting the search to those published within the past six months. AltaVista advanced and Search Assistant, Google advanced, Hotbot basic and advanced, and Yahoo advanced all provide this capability. Northern Light Power Search and AltaVista advanced and Search Assistant also provide textboxes where the user can key in a date range. Hotbot advanced allows the user to specify the oldest desired publication date of the documents to be retrieved.

The user can search for particular media types in some of the search engines, by specifying that the search should be limited to a particular collection, e.g., “Images”, “Video”, or “MP3”. The “collections” lists in AltaVista basic and advanced, Excite basic, Google basic, Lycos basic and advanced, and Yahoo advanced all include at least one option that specifies a particular medium. One search engine (Hotbot, basic and advanced) allows the user to limit their search to particular media using checkboxes, four in the basic interface and 14 in the advanced interface. As with other features of the advanced interface, the Hotbot basic interface can be “personalized” by adding any of the 14 checkboxes to it.

Some search engines allow the user to search a particular field within the record. AltaVista Search Assistant, Google advanced, Hotbot advanced, Lycos advanced, and Northern Light Power Search allow searches of the title field. In Google advanced, the user can specify searches in the text of the document. AltaVista Search Assistant, Excite advanced, Google advanced, Hotbot advanced, Lycos advanced, and Northern Light Power Search allow searches for a particular url. AltaVista Search Assistant, Excite advanced, Google advanced, Hotbot advanced, and Lycos advanced support domain searches. In all cases, a particular domain can be keyed in. Excite advanced also provides a dropdown menu listing suffixes used in the U.S. plus country names (as noted above). Hotbot also provides two other ways to limit a search. First, it provides the capability of limiting a search to a particular “page depth”. The choices, provided via radio buttons, are any page, top page, personal page, or a specified page depth. Second, it allows the user to specify that a search term relates to “a person”. This option, as an alternative to words or phrases, can be specified as part of the basic query. Northern Light Power Search also provides two additional ways to limit a search. Using a series of check boxes, the user can limit the documents retrieved to those representing one or more of a list of 16 subjects. In the same way, the user can limit the documents retrieved by document type. The list of 15 document

types include some implicit references to domain (e.g., Education web sites), but also includes some more specific document types (e.g., Job listings or Press releases).

A relatively new type of search capability is to be able to search the links associated with a particular page. Google advanced allows the user to specify a url, and it will retrieve the pages that link to that url. In its drop-down menu of search options (e.g., “look for all the words”), Hotbot advanced also includes “look for links to this url”. Lycos’ advanced search interface includes a “link referrals” tab, which leads to a page on which the user can search for the links to a specific url, search for links to a particular host, and/or exclude links to a particular host. While it is likely that only Internet-savvy users will understand the search capability that these interfaces afford, it seems likely that this capability will be added to more search engines in the future.

Support for modifying a query

Two of the search engines (AltaVista, all versions, and Google, all versions) support the use of relevance feedback to modify a query. In AltaVista, some (though not all) of the entries in the list of results include a hyperlink titled, “Related pages.” In Google, each of the entries in the list of results includes a hyperlink titled, “Similar pages.” In either case, clicking on the link launches a search for pages like the one selected. In addition, the Google advanced search page allows the user to type in a url and it will search for pages similar to that one. In none of these cases is “similar” or “related” defined for the user, so the mechanisms for executing the reformulated search are unknown.

Some of the search engines support the user in reducing the retrieved set by allowing them to use the retrieved set as the basis for the next query. In AltaVista basic, Hotbot basic and advanced, and Lycos basic, a checkbox allows the user to “Search within these results.” This checkbox is located next to the textbox that holds the text of the original query. In Google basic and advanced, this same capability is provided via hyperlinked text, “Search within results,” located at the bottom of the results page. Northern Light takes a novel approach by sorting the search results into “custom folders” by topic, and allowing the user to select one of these folders. The selected folder is then sorted into sub-folders, and the user can navigate within this unfolding hierarchy of topic specialization. This approach seems to be one implementation of the scatter/gather concept (Pirulli et al., 1996).

A number of the search engines suggest alternative terms for modifying a search. The most common approach is that used by AltaVista basic and Search Assistant, Ask Jeeves, Hotbot

basic and advanced, and Yahoo basic and advanced. Each of these search engines provides a list of alternative queries, each one a hyperlink. The suggested terms are generally placed at or near the top of the search results page (the one exception is in Yahoo, where they're placed at the bottom of the results page). Various phrases are used to suggest these alternative queries: "Others searched for..." in AltaVista, "Try these search terms to see more answers" in Ask Jeeves, "People who did this search also search for..." in Hotbot, and "Related searches" in Yahoo. Excite basic and advanced take a different approach. Next to the search button is a "Zoom in" button. If the user clicks it, rather than the button labeled "Search," a new window opens, containing alternative search term(s) that can be substituted for the original query term(s). Only one of these alternative terms can be selected for each query, but the terms themselves may contain multiple words. Several other search engines take a more directory-oriented approach, by providing links to categories in their directory. AltaVista, Google, and LookSmart provide category links with some of the retrieved items (called "see also" references in LookSmart). Yahoo provides a list of "Category matches" above the list of Web sites retrieved, and Ask Jeeves provides a list of questions similar to the one posed by the searcher.

Interestingly, many of the search engines also provide encouragement for users to search other search engines. In most cases, there are links from a search engine's results page that will automatically launch the same search in another search engine. Google advanced provides hyperlinks to AltaVista, Excite, Hotbot, Lycos, and Yahoo. Hotbot basic and advanced results pages include the Lycos logo, which serves as a link for getting a "Second Opinion". Conversely, Lycos refers its users back to Hotbot (and also to Web Directory and DirectHit, from Lycos basic). Yahoo provides hyperlinks to Google, Ask Jeeves, and "more..."

Additional searching capabilities

Some of the search engine interfaces analyzed provide additional affordances. A couple of the search engines offer a selection of pre-edited searches or current articles on popular topics. On its basic search page, Lycos presents its "Top 50" hot search topics for the week. Northern Light offers a somewhat similar service with its Special Editions, "concise and timely accounts of today's most important issues." The AltaVista and HotBot basic interfaces each provide access to "text-only searches." This link, located near the search textbox in AltaVista and at the bottom of the page in HotBot, leads to a graphics-free version of the basic search interface. The AltaVista advanced search interface allows the user to weight the search terms. The Boolean query is entered and, in a separate textbox, users may specify terms to "sort by." Ask Jeeves

provides a more communication-oriented approach, via a an “Ask Other People” tab. On that page, the user can perform keyword searches of past postings or post a question to the Ask Jeeves community. Each of these capabilities is afforded by a particular interface, but is not available in multiple search engines.

Combinations of search engines and directories

Many search engines are attached to or are augmented by directories of links. The previous section described how the directory structure is used in combination with the search results to encourage users to modify their query. In addition, some of the search engines (AltaVista basic, Ask Jeeves, Excite basic, Hotbot basic, LookSmart, Lycos basic, and Yahoo basic) provide a directory on the search page (see Table 5 for the details on each). None of the advanced search pages provides access to a directory.

Table 5. Directories available with search engines

	Number of main categories in directory	Placement of directory on search page
AltaVista	14	Begins midway down page
Ask Jeeves	17	Available via separate tab, titled “Browse by subject”
Excite	12	Begins midway down page
Hotbot	14	Begins near top of page
LookSmart	10	Begins near top of page
Lycos	14	Begins midway down page
Yahoo	14	Begins midway down page

The number of categories in the directory varied, from 10 to 17. None of the sets of categories is identical to any other set. It appears that Hotbot and LookSmart place the most emphasis on the directory, since it is placed nearer to the top of the search page. These two search engines differ in their approach, though, as evidenced by the contrast between Hotbot’s relatively-complete selection of search capabilities and LookSmart’s very sparse selection of search capabilities.

Characteristics of results displays

There is a fair amount of standardization in the way that results are presented. All except Ask Jeeves report the total number of items retrieved. All of the search engines implicitly provide the results sorted in relevance order; only Northern Light provided an explicit relevance

score. Northern Light advanced also allows the user to specify that the results be sorted by date, rather than relevance. Lycos basic and Excite basic place “sponsored” or “featured” sites at the top of the results list.

Most often, each item listing includes the page title, a 2-3 line description, and the page url (except in LookSmart). In Excite basic and advanced and Hotbot advanced, the user may customize the display slightly, from the search specification page. The choices include the full description (described above), a brief description (in Hotbot), titles only (in Excite), and url’s only. In Ask Jeeves, Excite, Hotbot, Lycos, Northern Light, and Yahoo, the retrieved items are numbered. The item listings in Northern Light also include the page’s relevance score, date, and an interpretation of the url domain, e.g., “commercial site”. AltaVista offers to translate the retrieved items using Babel Fish.

Most search engines present 10 items per page in the results listings. The only exceptions are Yahoo basic, with 20 items per page as the default, and LookSmart, which displays 12-15 items per page. The advanced versions of Excite, Google, Hotbot, and Yahoo allow the user to specify the number per page, using a drop-down menu on the search specification page. The Google preferences page also allows the user to specify how many results per page should be displayed.

Most of the search engines allow the user to view all the items retrieved. AltaVista limits the number of pages that can be viewed to 20. Excite, Google, LookSmart and Lycos basic allow the user to skip through the pages, while the rest of the search engines require that the user sequentially page through every page in the results set (i.e, the only results navigation tool is a “next page” link).

The majority of the search engines highlight the query terms in the results listings, but this is not a universal feature. AltaVista basic and Search Assistant, Excite, Google, Lycos, and Yahoo all use bold face to highlight the query terms. Northern Light uses bold face if the term appears in the page title, and blue text if the term appears in the item description.

Setting preferences for searching and displaying results

A few of the search engines allow the user to customize the search capabilities and results displays, and to save those preferences. Access to this capability via the interface is different in each case. AltaVista (all versions) has a hyperlink near the search textbox, titled “Customize Settings”. When this link is traversed, the user can then specify which Web page information should be shown in the results, which links from the results (translate, related pages, more pages

from the site) should be available, whether the query terms should be in bold face in the results display, the number of items that should be displayed on each results page, which languages to search in (selected by checking one or more of the 25 available), and the type of encoding for any non-Roman alphabets to be used in searching. Google basic has a hyperlink near the search textbox, titled “Preferences”. This link leads to a page on which the user can specify the interface language (selected from a drop-down menu including 66 choices, including some quasi-languages like Pig Latin and Hacker), the language for searching (selected by checking one or more of the 28 available), the number of items to be displayed on each results page, whether to open the results pages in a new browser window, and whether SafeSearch filtering should be set on or off. Hotbot basic includes a button, just below the search criteria, titled “Personalize These Settings”. When this button is selected, the customization page allows the user to move any of the advance search and display capabilities of the interface to the basic search page. These capabilities include phrase and field searching, the Hotbot word filter, limiting by language, limiting by date, limiting by media format, limiting by location/domain, limiting by depth within a particular site, word stemming, the level of detail for the results listings, and the number of results to be listed on each page.

Summary

In general, some aspects of Web search engine interfaces are becoming more standardized and other aspects vary widely across the search engines. All search engines provide a textbox and some type of accompanying button for entering a query. The textboxes varied widely in length, and some search engines had more than one search button. Almost all of the search engines provide assistance in specifying a query, but only two provide examples of queries on the search page itself. All of the search engines use the same basic syntax for specifying a query (e.g., a + symbol for requiring that a term be present in the retrieved items), but there is quite a bit of variation in the type and amount of assistance provided via drop-down menus or checkboxes in the interface. The mechanisms for limiting a search varied widely across search engines, though almost all provided a way to limit the search to particular collections of resources and/or resources in particular languages. Direct support for modifying a query (e.g., relevance feedback or limiting a search to the results set just retrieved) was available in only a few of the search engines. There are a number of hybrid search engine/directory interfaces, with directory topics reported as part of the results set or the basic directory sharing the search page with the search engine. The results of searches are reported in fairly standard

ways: brief summaries provided, usually 10 per page, in relevance order. Some of the search engines allow the user to customize the search interface or the presentation of results.

DISCUSSION

The “first wave” of end user searching occurred when personal computers and online access became available to individuals, in the early 1980’s. The audience consisted primarily of professionals or academics who were not able to access the databases with which they were familiar and which they had already been accessing via intermediaries. For example, chemists began to search *Chemical Abstracts* and physicians began to search *MEDLINE*. As librarians and other information professionals faced this first wave of people who wanted to search for themselves but had no prior knowledge of how to conduct a successful search, several approaches were taken in the attempt to improve the search outcomes of the end users. These included search training (both face-to-face and online tutorials) and, secondarily, the development of search interfaces developed specifically for end users. Eventually, most end users were able to distinguish the searches that were reasonably straightforward and in their own discipline (which they could conduct for themselves) from those that were more complex or in a new domain (for which they would seek intermediary assistance).

With the advent of the Internet and the World Wide Web, in the mid-1990’s, end user searching has entered a “second wave,” in which many more people are using online resources to locate information that is useful to them. This wave differs from the previous wave of end user searching in two important ways. First, the information needs presented by this new wave of searchers are much more diverse. For example, at the moment this paragraph was drafted, the most recent questions posed to Ask Jeeves included, “Where can I learn about the physics of a tsunami (tidal wave)?” and “Where can I find pictures of or information about the architectural structure of the Sears Tower?” Second, the information resources available on the World Wide Web are much more diverse than the well-structured bibliographic, statistical, and full-text databases available online in prior decades. Given these two changes, the Web engine search interface must do the bulk of the work in aiding users in conducting successful searches. As noted by Shneiderman, Byrd, and Croft (1997), “the future of the World Wide Web as a universally acceptable tool may depend on our ability to reduce the frustration and confusion of the masses of users, while enabling them to reliably find what they need.” It is for this reason that this paper has focused, not on the search engine’s capabilities, *per se*, but on its perceived affordances. The interface objects available on a search interface can communicate to the user

about the options available for searching and the limits on those options. This study was undertaken to understand the current state of the art in Web search engines and the perceived affordances that are available in their interfaces.

Based on the results of this study, it can be concluded that we have not yet made great progress in supporting end user searching on the Web. The standardization of the “basic” search interface is useful, in that searchers can use essentially the same syntax in essentially the same way to conduct a search on any Web search engine. In most cases, there is no need to learn Boolean logic. Any set of entered terms can be processed by the search engine. While this design direction allows any person to enter a search and retrieve something, it provides little or no guidance to searchers in formulating a query that will effectively fulfill their information needs. In most cases, guidance/help on search syntax and other ideas for improving query formulation is anywhere from one to three clicks away from the search page.

One area that could easily be redesigned is the size of the textbox for entering query terms. In most cases, a search can be improved by adding more terms to the query; while this will increase the size of the retrieved set, it is likely to provide more input for the search engine’s ranking algorithm, thus putting more relevant items higher on the list. Excluding the very large textbox offered by the AltaVista advanced search interface, the average length of the textboxes on the search interfaces examined is 39 characters. All of them allow the user to add more characters (i.e., the textbox supports horizontal scrolling), but this capability is not visible when examining the interface. Only when a novice searcher tries to push beyond the typical two-word query may this functionality be perceived. Thus, it seems likely that longer textboxes may result in longer and, thus, more effective searches.

Most of the search engines offer a second search interface that provides some hints about the availability of mechanisms for improving a query. Almost all of these are titled “Advanced” in the link for accessing them; such terminology will not encourage novice users to explore this interface for guidance in formulating their searches. If the user does adventure into this space, he or she will often be rewarded by several additional tools that provide some structure for formulating a well-defined query. For example, the “word filter” in Hotbot’s advanced search interface allows the user to specify multiple query terms, each in a separate textbox. Two drop-down menus next to each textbox allow the user to specify whether the term is a word, a phrase, or a person, and whether the retrieved items must contain, should contain, or must not contain the term. While adding the extra interface objects may make the interface more complex, it affords the user the opportunity to more completely specify the query. The same argument could

be made for such “advanced” search features as limiting a search by date, by language, or by type of media.

Only a few of the search engines provide additional mechanisms for searchers to modify their initial query. There is evidence that users can select additional terms from a suggested list, and do desire that type of control (Koenemann & Belkin, 1996). Yet only a few of the search engines provide such a facility. There is evidence that users understand and can use relevance feedback mechanisms, if they are available (Spink, Jansen, & Ozmultu, 2000). Yet only AltaVista and Google provide relevance feedback links in their results listings. There is evidence that users often want to narrow their searches (e.g., Wildemuth, de Bliet, & Friedman, 1992). Yet less than half the search engines provide the capability to conduct a follow-up query within the initial set of retrieved items. Today’s search engines could be improved by providing these types of support to users who wish to modify their initial query.

Once a query has been executed, the results are displayed. In order for a user to evaluate the results set and make decisions about browsing through the listed items, the search engine should provide some feedback about the characteristics of the retrieved items. Appropriately, almost all the search engines notify the user of the number of items retrieved. Most then display a short summary of each item, with ten items on the first page of the listing. If the user wants to look at more of the retrieved items, he or she must follow a link to the next page of results. In setting the default number of results to appear on the first page of the listing, the search engine interface designers have faced a tradeoff—to display more results might require a longer download time for the page, but to display less results requires that the user download more pages. Based on studies of OPAC use (e.g., Wiberley, Daugherty, & Danowski, 1990; Bates, 1984), it seems likely that the default number could be increased somewhat, particularly if the number and size of the graphics on the results page is kept to a minimum so that download times are reasonable.

Most of the problems identified in this study and discussed above involve missing functionality (i.e., capabilities that would be useful but for which there are no perceived affordances) or hidden affordances (e.g., the fact that the search textbox can scroll). In addition, a few of these problems could be described as false affordances (Fitzgerald & Goldstein, 1999). For example, no search engines explain the use of stop words, so users might expect to be able to search for “Vitamin A”. Yet, if a person enters “vitamin a” in Google’s advanced search interface, looking for “results with all of the words,” the search will yield the same results as if he or she had searched only “vitamin.” The unexpected behavior of this search engine on this

search could be considered a false affordance because the interface seems to afford the possibility of searching for both “vitamin” and “a”, yet appears to perform a search only on “vitamin.” It is likely that many of the search engines examined here include a number of false affordances that could not be identified with the methods used in this study. To accurately identify which of the perceived affordances of these search engines are “false,” a study should be designed to include direct observations of user interactions with each search engine interface.

In summary, it is clear that Web search engines are widely used by the general population, in response to a wide variety of information needs (Rainie & Packel, 2001). While the “basic” search engine interface is relatively simplistic and, in some sense, usable by its intended audience, there are many affordances that could be added, having the potential to improve the effectiveness of the searches being conducted.

CONCLUSION

The results of this study provide a snapshot of the state of the art of the perceived affordances of Web search engines. These findings have implications for end users, as they attempt to select a search engine that is likely to fulfill their information needs; for librarians and other educators, as they train their students to conduct effective Web searches; and for designers of Web search interfaces.

Anecdotal evidence would indicate that many end users are unaware that there are differences among the many search engines available on the Web. One explanation for this lack of awareness may be the standardization of the “basic” search interfaces—they all look alike. However, users may be able to conduct much more effective searches on some search engines than others. For example, if someone were searching for audio files, the Hotbot advanced interface provides a simple checkbox that affords this functionality. If all the interfaces converge in their perceived affordances, it will make little difference which engine is used for searching (other than database size and coverage). Until then, it can be argued that some search engine interfaces provide affordances that others do not.

Librarians and other educators are often in the position of introducing their students and other clients to Web searching. These instructional programs often include a discussion of term selection and query formulation. They should also include some review of the perceived affordances of Web search engines. While it would be impossible to include all the affordances offered by all the search engines, trainers should attempt to make their students aware that the search engines do differ in this regard and how to understand the functionality that is being

offered in any particular interface. In general, it is recommended that trainers encourage their students to explore the “advanced” search engine interfaces in order to become more familiar with the power that they offer in developing an effective query.

Interface designers and search engine developers have an opportunity to improve the design of Web search engine interfaces. Such improvements might include: extending the textbox length to encourage users to enter longer queries, renaming the link to the “advanced” search interface to encourage users to explore its capabilities, providing examples of well-formulated queries on the search page itself, and making the query syntax (e.g., +, =, “phrase”) more apparent to users via drop-down menus, check boxes or other interface objects. As more is known about user searching behaviors, additional recommendations can be formulated.

In order to move to the next steps of improving interface designs, additional research should be conducted. In particular, studies of actual user perceptions and behaviors should be undertaken. A more complete understanding of users’ perceptions of search engine interface affordances can be gained through direct observation of searching and concurrent think-aloud protocols. Once we have a more complete understanding of how searchers perceive and understand the interface features available, we can begin to study how design changes affect searcher behavior. While it can be expected that the perceived affordances of a Web search interface will affect users’ behaviors, we do not yet know how likely or strong these effects are. For example, we can expect that making the textbox longer will encourage users to generate longer queries, but we don’t yet know whether such an interface change actually will affect user behavior. Two types of studies are called for. The first, and simplest, is to gather transaction logs of existing Web search engines that are quite different in their perceived affordances, and compare current user behaviors. For example, queries from the transaction logs from the Excite advanced search interface (with a textbox 20 characters long) could be compared with those from the AltaVista Search Assistant interface (with a textbox 87 characters long). Our hypothesis that a longer textbox will result in longer queries could thus be tested. The second type of study is the controlled manipulation of particular interface objects. For example, different ways of presenting alternative search terms to users could be developed and compared. User interactions with each alternative design could be observed, the strengths and weaknesses of each design could be determined, and the study could result in the most effective mechanisms for presenting this functionality. It is hoped that researchers will continue to pursue these issues, since the results of their research could improve the search success of millions of people.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the assistance of Chang Su in updating the detailed review of the search engines in fall 2001. Financial support for this project was received from the University of North Carolina at Chapel Hill Summer Pre-Graduate Research Experience program.

REFERENCES

- Bates, M. J. (1984). The fallacy of the perfect thirty-item online search. *RQ*, 24(1), 43-50.
- Fitzgerald, W., & Goldstein, E. (1999). Honesty of affordance. In Marsh, J. P., Gorayska, B., & Mey, J. L., *Humane Interfaces: Questions of Method and Practice in Cognitive Technology*. Amsterdam: Elsevier, 179-189.
- Gibson, J. J. (1977). The theory of affordances. In Shaw, R., & Bransford, J., *Perceiving, Acting, and Knowing : Toward an Ecological Psychology*. Hillsdale, NJ: Lawrence Erlbaum, 67-82.
- Koenemann, J., & Belkin, N. J. (1996). A case for interaction: a study of interactive information retrieval behavior and effectiveness. *Human Factors in Computing Systems: CHI 96 Conference Proceedings (April 13-18, 1996, Vancouver)*, 205-212.
- Norman, D. A. (1999). Affordance, conventions, and design. *Interactions*, 6(3), 38-41,43.
- Notess, G. (2000). Search engine showdown [online]. <http://www.notess.com>.
- Pirolli, P., Schank, P., Hearst, M., & Diehl, C. (1996). Scatter/gather browsing communicates the topic structure of a very large text collection. In Tauber, M. J., et al., *Human Factors in Computing Systems, Common Ground: CHI '96 Conference Proceedings (Vancouver, April 13-18, 1996)*. New York: ACM Press, 213-220.
- Rainie, L., & Packel, D. (2001). More online, doing more. Washington, DC: Pew Internet & American Life Project.
http://www.pewinternet.org/reports/pdfs/PIP_Changing_Population.pdf. Last accessed March 12, 2002.
- Shneiderman, B., Byrd, D., & Croft, W. B. (1997, January). Clarifying search: a user-interface framework for text searches. *D-Lib Magazine [online]*.
<http://www.dlib.org/dlib/january97/retrieval/01shneiderman.html>. (Last viewed on November 15, 2001.)
- Spink, A., Jansen, B. J., & Ozmultu, H. C. (2000). Use of query reformulation and relevance feedback by Excite users. *Internet Research: Electronic Networking Applications and Policy*, 10(4), 317-328.

- Sullivan, D. (2000, July 17). StatMarket search engine ratings [online]. In *Search Engine Watch*. <http://searchenginewatch.com/reports/statmarket.html>. (Last viewed on November 15, 2001.)
- Sullivan, D. (2001a). Directory sizes [online]. In *Search Engine Watch*. <http://searchenginewatch.com/reports/directories.html>. (Last viewed on August 21, 2001.)
- Sullivan, D. (2001b, August 15). Search engine sizes [online]. In *Search Engine Watch*. <http://www.searchenginewatch.com/reports/sizes.html>. (Last viewed on August 21, 2001.)
- Sullivan, D. (2001c, July 6). Jupiter Media Metrix search engine ratings. In *Search Engine Watch*. <http://searchenginewatch.com/reports/mediamatrix.html>. (Last viewed on August 21, 2001.)
- Sullivan, D. (2001d, February 19). Nielsen//NetRatings search engine ratings [online]. In *Search Engine Watch*. <http://searchenginewatch.com/reports/netratings.html>. (Last viewed on November 15, 2001.)
- Wiberley, S. E., Jr., Daugherty, R. A., & Danowski, J. A. (1990). User persistence in scanning postings of a computer-driven information system: LCS. *Library & Information Science Research*, 12(4), 341-353.
- Wildemuth, B. M., de Blied, R., & Friedman, C. P. (1992). Search moves made by novice end users. *ASIS '92 Proceedings (Pittsburgh, PA, October 26-29, 1992)*, 29, 154-161.