Exploring the Realities of Interaction and Search Success

Elisabeth Logan and Kristen Jacobson

Florida State University

Abstract

Investigating skilled and novice interactive searching in a menu-based environment, little or no significant difference was found between the two groups in the number of cycles, number of entries, or in the percent of positive interactive responses during a search session. Despite high percentages of appropriate entries in response to system results, only half the searchers were able to use iterative cycles to obtain additional useful information, although skilled searchers were more successful than novice searchers and interim successes were greater than final success. Whether these findings can be generalized to all searching situations or are particular to menu-based systems merits further investigation.

1. Background

Interactive searching and the associated integral role of the system user has been the focal point of a substantial segment of the online field since Bates[1], Belkin[2], and Saracevic[3] among others challenged the widely-held concept that ideal searchers' clearly stated information needs were matched to relevant documents by information retrieval systems. These systems were then best evaluated on how well the mechanics of the search engine was able to make an effective match. In a review of interactive information retrieval, Savage-Knepshield and Belkin[4] note that early information retrieval systems focused on improvements in the design and implementation of search mechanisms rather than on users and what they needed to interact effectively with retrieval systems. Saracevic, in championing the role of the user and the importance of an effective user interface, argues that relevance and the necessary involvement of human

judgements as users interact with IR systems, involves an exchange of information between users and systems. However, there still remains a division between those disciplines which focus on the system and system performance and those which recognize the importance of the interactivity between systems and users.

Many others have addressed the importance of user interaction from a variety of viewpoints, Marchionini[6], Fidel[7], Xie and Cool[8], Shaw et al[9] and Hseih-Yee[10] However, despite arguments over whether the First Digital Libraries Initiative and TREC research sufficiently addresses the role of the user, it is clear that in the field of information science, user-system interaction is of considerable importance.

Because of the commonly recognized importance of user-system interaction, and despite intuitive assessments of search results which appear to demonstrate improved retrieval results with continued iteration, there is little in the current literature to evaluate this premise. The need for more specific information on the nature and results of interactive searching merits further investigation[11]. In an article by Wolfram and Dimitroff[12] in which searches of Boolean and hypertext environments were investigated, for both novice and experienced searches they report among other things that novice searchers" spent more time searching the hypertext system...[but] this extra time did not result in better performance" .

With this as background, our study investigates interactive search activity of novice and skilled searchers in a menu-based environment and its relation to both perceived and actual search success.

2. The Study

The study investigates interactive behavior of novice and skilled searchers as they seek answers to assignment-related questions. Searching in the Florida State University School of Information Studies' Usability Lab, videos and online logs were made of each searcher' s actions. Search results were analyzed for both short run and long run success and compared with related iterative actions. Short run " success" was noted if an entry was determined to be an

appropriate response to the previous search result and served to advance the search process. That is, if a searcher broadened a search appropriately, it was considered a successful iteration. Likewise, if an action was inappropriate or an unnecessary repeat of a previous move, it was considered unsuccessful. Interim success was defined as an iterative cycle anywhere in the search session which produces useful or positive results. Long run success was noted if a search was terminated at the point where useful information was found after an iterative cycle.

Both novice and skilled searchers looked for answers in the Lexis-Nexis Academic Universe Databases available to all students at Florida State University on the campus website. This format was chosen for several reasons. First, it is a user-friendly version of the traditional Lexis-Nexis system and as such should require little previous knowledge of how the system operates. This made it an ideal vehicle for true novice users. Second, because of the important position menu-based interfaces have been accorded, it seemed appropriate to look more closely at whether this example would be able to elicit appropriate searching action from novices as well as skilled searchers.

3. Methodology

The novice searchers for the study are from an undergraduate class, Information Needs and Preferences, in the school of Information Studies at Florida State . For many this was the first class taken in their major. All participants ranked themselves as “ novice” on the pre-search questionnaire although several indicated an “ expert” level of familiarity with the WWW. Students participated in the study as part of their first assignment, so were novice searchers in the true sense of the word. The searchers were assigned individual searching times in the Usability Lab where their actions and search logs could be captured. All were introduced to the facility and given a brief introduction to the Lexis-Nexis Academic Universe environment. Each searcher was given a printed statement of the question and basic guidelines, asked to speak aloud as they worked, and encouraged to search until useful information had been found. Searchers were

advised to print out any material considered helpful in answering the assigned question.

The skilled searcher group (note that we chose the term “ skilled” rather than “ expert” to reflect more accurately the nature of their ability), is actually composed of two subgroups: undergraduate students and graduate students. Both the undergraduate and graduate groups were members of classes in electronic searching and had completed close to an entire semester of course work. Most of this group indicated they were “ knowledgeable” or “ skilled” in searching menu-based systems such as Lexis-Nexis Academic Universe. One in each subgroup, however, did indicate the “ novice” category.

The skilled searchers were asked to participate as part of a regular classroom activity and searched for documents useful in answering a question posed as part of a homework assignment for which credit was given. As in the case of the novice searchers, the skilled searchers were introduced to the lab facility and to the Lexis-Nexis Academic Universe environment and given a printed sheet containing the question and basic guidelines. All were encouraged to speak aloud and to search until they felt useful information had been located. They were to print out their results and use them to fulfill their homework assignment. As each participant searched, an audio/video tape was made of the process and search interactions logged electronically. After the search session, all participants were asked to complete a brief questionnaire evaluating the search process and their results. Questions for each group were pre-tested to insure that appropriate material was available. The questions were not difficult but did require that appropriate sources be selected, dates checked, and keywords be entered in logical combinations. Participants were shown were to find the readily available Academic Universe help screens.

4. Data Analysis

Audio/video tapes, search logs, and printed results were transcribed and analyzed by the investigators. Since in this study we were primarily interested in searcher iteration and its relationship to long run and short run measures of

search success, we examined each search session for several factors: 1) number of cycles in each search session, 2) number of entries per session and per cycle, 3) number of positive (appropriate) or negative (inappropriate) responses per session and per cycle, 4) number of cycles where interaction produced positive or negative results, 5) number of sessions in which iteration produced useful results, 6) relations between perceived and objective evaluations of success.

A search session was completed when the participant indicated that for whatever reason, the search was over. Each time during a search the searcher stopped entering terms and printed out material was determined to be a cycle. Perceived success is an evaluation by the searcher that retrieved material is useful; objective success is an evaluation by the investigators that the retrieved material is relevant to the question. Positive moves were those considered appropriate to system responses. That is, if the search was broadened or narrowed when appropriate to the situation and advanced the search process, the move or action was considered positive and a " short run success" . The positive moves included broadening or narrowing by any means as well as lateral actions (those using the same or similar terms but in different sources). Negative moves were those in which the above actions were inappropriately taken or where moves were repeated unnecessarily. Interim success was defined as an iterative cycle anytime in the search session which produced useful or positive results, and long run success noted when a search was terminated at the point where useful material was located through an iterative cycle.

5. Results

5.1 Novice Searchers - Cycles and Entries

The group of ten novice searchers used from 1 to 5 cycles during their search sessions. Three searchers used 2 cycles; three searchers used 4 cycles; two searchers used 3 cycles and one searcher used 5 cycles and another used a single cycle.

Most of the novice searchers used a large number of entries during each search session. Half of the group used more than 15 search entries per session and four of the ten used 28 or more. The largest number of entries for one session was 34. On the other hand, two searchers used very few terms over an entire search session; only 5 and 3 total entries respectively.

Entries per cycle likewise were high. At least half the group used 15 or more entries during a single cycle at some time in their search session. The highest number of entries per cycle was 29; other cycles consisted of just one or two entries.

5.2 Novice Searchers - Iterative Responses

In general the novice searchers' iterative choices in response to system results were largely appropriate or positive - short run success. Positive responses as percents of total responses per session ranged from 100% appropriate to 40% appropriate. Eight of the ten novice searchers had higher proportions of positive to negative responses. The two with higher negative responses than positive ones, were nonetheless able to demonstrate appropriate responses at least 40% and 44% of the time.

Looking at per cycle responses, cycles predominated in which there were more positive than negative interactive commands and of these, most were associated with positive results or interim success. Of twenty-nine total cycles, twenty-one or 72% showed more positive or appropriate interactions than negative ones. Six in which the negative commands predominated make up 21%. The remaining 2 are those in which positive and negative responses were equal. Cycles in which appropriate or positive commands predominate, produced retrieval " success" in some, but not all cases. A positive cycle balance produced useful results 16 times or 57% of the cycles counted. A positive cycle balance produced negative or not useful results 5 times or 18% of the time. A negative cycle balance resulted in positive retrieval hits 4 times or 14% of the time, and finally a negative cycle balance produced results that were also negative in three cases or 11%. This relationship is shown in

Table 1: Cycles and Results- Novice Searchers		
	Positive Results	**Negative Results**
Positive Cycles*	57%	18%
Negative Cycles	14%	11%

*Positive Cycles are those in which the percent of appropriate iterative responses outweighs the percent of inappropriate ones.

5.3 Novice Searchers - Short Run and Long Run Success

How these searchers' interactive responses relate to short term, interim, and long term successes varies considerably. In the short run, as has been indicated, iterative commands for most searchers were " successful" to the extent they were appropriate to the immediate searching situation and served to advance the search process. Eight of ten novice searchers were able to achieve positive entries at least 61% of the time. Positive Interim success, defined as locating useful records by pursuing additional search cycles *at any time* in the search process, and long run success defined as terminating the search process at the point where iteration had achieved successful identification of useful information, was demonstrated by fewer novice searchers. Although all novice searchers did locate useful information after the first cycle, despite high percentages of positive iterative responses during a search session, the interim and long run results were less satisfactory. Of the novice searchers, six demonstrated no long run success from iterative cycles; three were able to achieve some interim success through iterative cycles, but only one was able to pursue the iterative process successfully through to search session termination. Note here, however, that one searcher having found useful information, attempted no iteration after the first cycle.

A summary of novice searcher iterative responses is shown in Table 2.

Table 2: Summary of Novice Searcher Iterative Responses

Searcher	Cycles	Entries	Success: Short % positive entries	Success: Interim # successful cycles	Success: LongRun Final Improvement
One	2	28	57%	1	no
Two	4	34	44%	1,2,3*	no
Three	2	3	100%	1	no
Four	3	31	68%	1,2*	no
Five	3	11	55%	1	no
Six	5	13	69%	1,2,3*	no
Seven	4	5	40%	1	no
Eight	1	15	60%	1	no iteration
Nine	4	8	75%	1,2,3,4**	yes
Ten	2	29	61%	1	no

*Interim iterative improvement only; **Final iterative improvement

5.4 Skilled Searchers - Cycles and Entries

The group of skilled searchers consists of five undergraduates and five graduate students. The undergraduates used from 2 to 6 total cycles whereas the graduates used from 1 to 5. Of the combined group of skilled searchers, 6 of the ten used either 2 or 3 cycles.

Of this group only two skilled searchers used 19 or more entries per session. Here the high was 28 entries, and the lowest 5. Entries per cycle were a high of 21 and a low of 4. Looking at the two subgroups of skilled searchers, the undergraduate group used fewer entries per cycle than the graduate group. In fact on the entry continuum, 4,5,5,7,8,8,9,12,12,21 for the group, all undergraduate entry totals fall in the lower half - 4 through 8.

5.5 Skilled Searchers - Iterative Responses

Of this group, the iterative choices made in response to system responses were overwhelmingly positive. Positive responses as percents of total responses per session ranged from 92% to 42%. Nine of the ten and all of the skilled undergraduates had higher positive ratios than negative. With the exception of one graduate who had 42%, all skilled searchers were 60% or higher. Here as in the novice searcher group, defining short run success as the ability to move the search process forward in a positive manner through system interaction, we see that 9 of 10 skilled searchers were able to demonstrate short run success.

Looking at per cycle responses of skilled searchers, cycles predominate in which there were more positive than negative demands. Graduate students had 13 of 14 positive cycles or 93%; undergraduates had 10 of 17 positive cycles or 59%. There were no negative cycles for graduate students, but one cycle had equal numbers of positive and negative entries. For the undergraduate skilled searchers, there were 6 negative cycles (35%) and one with an equal number of positive and negative entries. For the total of 31 cycles for all skilled searchers, 74% or 23 were positive and 19% or 6 were negative and 2 or 6% had equal numbers of positive and negative entries.

Cycles for which appropriate or positive commands outweigh the negative ones produce retrieval success or useful results in some, but not all cases for the skilled searcher group. A positive cycle balance produced useful results 16 times in 31 cycles or 52%. Graduate students' tally was 7 of 14 for 50%; undergraduate had 9 or 17 for 51%. A positive cycle balance resulted in negative or non-useful information 28.6% of the time for graduates and 17.6% for undergraduates. Combining these gives a percent for the entire group of 7 or 22.6%. A negative cycle balance produce useful results 3 times or 17.6% for undergraduate, but no instances for graduate students. Negative cycle balances which produced negative results occurred in one case each for graduates and undergraduate, (6.5%). For the skilled searcher group, the relationship is shown in Table 3.

Table 3: Cycles and Results- Skilled Searchers

	Positive Results	Negative Results
Positive Cycles*	52%	22.6%
Negative Cycles	9.7%	6.5%

*Positive Cycles are those in which the percent of appropriate iterative responses outweighs the percent of inappropriate ones.

5.6 Skilled Searchers - Short Run and Long Run Success

How the skilled searchers' interactive responses relate to both short term, interim, and long term success varies as do those of the novice searchers. In the short run, iterative commands for most searchers were successful to the extent they were able to move the search process forward. All but one of the searchers were able to achieve short run success of 60% or greater. For interim success, or locating useful information at any time during a search session by pursuing additional search cycles, skilled searchers were less successful, although demonstrating a higher rate of success than the novices. Six of the ten (60%) were able to realize interim success, and of those, three were also able to achieve long run success. On the other hand, of this group, two searchers - both graduate students - were unable to locate any useful records. A summary of skilled searchers' iterative responses is found in Table 4.

Table 4: Summary of Skilled Searcher Iterative Responses

Searcher	Cycles	Entries	Success: Short % positive entries	Success: Interim # successful cycles	Success: LongRun Final Improvement
One	3	9	78%	1,2*	no
Two	3	10	70%	1,2,3**	yes
Three	3	13	77%	1,2,3**	yes
Four	2	5	60%	1	no

Five	6	15	67%	1,4,5*	no
Six	5	19	63%	1,2,3,4*	no
Seven	4	28	79%	0	no
Eight	1	12	42%	0	no
Nine	2	11	82%	1,2**	yes
Ten	2	13	92%	1	no

*Interim iterative improvement only; **Final iterative improvement

5.7 Comparing Skilled and Novice Searchers

Comparing the performances of skilled and novice searchers, very little difference is seen between the two groups in the number of cycles, number of entries, or even in the percent of positive iterative responses or short run success. The mean number of cycles for both groups differs by only .1, the mean number of entries differs by 4.2, and the mean percent of positive iterative responses differs by 8.1. Table 5 shows the comparisons for these two groups.

Table 5: Skilled and Novice Searchers Compared

	Skilled Searchers			**Novice Searchers**		
	Cycles	**Entries**	**% Positive**	**Cycles**	**Entries**	**% Positive**
Mean	3.1	13.5	71	3	17.7	62.9
Std. dev	1.52	6.29	13.9	1.24	11.65	16.9
t-score	.0152*	1.004*	1.170*			

* Not statistically significant at .05

The greatest difference between these two groups exists in the number of positive responses during a search session. The t-score is still not statistically significant, however.

5.7 Comparing Skilled and Novice Undergraduates

Looking at the scores for skilled and novice undergraduates for the same measures, cycles, entries, and % positive entries, shows no significant difference for these two groups although the greatest difference here is also in the % positive entries. Evidently for these measures, training makes not much of an impact. Table 6 gives the comparisons for this group.

Table 6: Skilled and Novice Undergraduates Compared

	Skilled Undergraduates			Novice Undergraduates		
	Cycles	**Entries**	**% Positive**	**Cycles**	**Entries**	**% Positive**
Mean	3.4	10.4	69.4	3	17.7	62.9
Std. dev	1.51	3.84	8.29	1.24	11.65	16.9
t-score	.5126*	.0179*	1.000*			

* Not statistically significant at .05

Here where we might have expected to see some significant difference between undergraduates with training and those who were true novices, the differences in these measures are not statistically significant at .05

5.8 Comparing Skilled Graduates and Undergraduates

Looking at scores for skilled graduates and skilled undergraduates for the same measures as the previous groups, we see the greatest difference here is in the number of entries per search session. Novices use more entries per session, however, the difference is still not statistically different at .05. These comparisons are given in Table 7.

Table 7: Skilled Undergraduates and Graduates Compared

	Skilled Undergraduates			Skilled Graduates		
	Cycles	**Entries**	**% Positive**	**Cycles**	**Entries**	**% Positive**
Mean	3.4	10.4	69.4	2.8	16.6	71.6

Std. dev	1.51	3.84	8.29	1.64	7.09	19.55
t-score	.6012*	1.719*	.2316*			

* Not statistically significant at .05

6. Conclusions

The most dramatic conclusion that can be drawn from the results of this investigation appears to be the very small amount of difference between groups of skilled and novice searchers. Not only are there few differences between the two large groups, this also extends to comparisons between skilled and novice undergraduates and between skilled undergraduates and skilled graduates. Tables 5, 6, and 7 show that for the number of cycles, number of entries, and percent positive responses within a search session, there is no statistical significance among these measures. However, there are some interesting observations to be made. First, the mean number of cycles is close to 3 for all groups. This is quite surprising given the possibility of unlimited search time and no emphasis in the instructions on search efficiency. Also interesting is that the number of cycles by skilled undergraduates falls at the lower end of the cycle continuum for all skilled searchers. Second, skilled searchers tend to use fewer entries per session, although skilled undergraduates used fewer than skilled graduates. Perhaps training does mean more efficient searching. Third, for the percent of positive entries for short-term success there is no significant difference, however skilled searches do have a higher percent of positive responses than do novices (71%, 62.9%), as do skilled to novice undergraduates (69.4%,62.9%), and skilled graduates to skilled undergraduates (71.6%, 69.4%). It appears that skill and experience may have an impact upon the short run success of iterative responses.

Tables 1 and 2 provide another interesting means of comparison. Although skilled searchers had slightly more positive cycles than novice searchers (74%, 72%), novice searchers were able to achieve a higher percent of positive results.

On the other hand, they also had higher positive results from negative cycles and a higher percent of negative cycles and negative results. This may indicate a certain degree of randomness where positive cycles lead to positive results, but these are accompanied by more mis-matches as well. Note that novices also tend to use more total entries as well. For the skilled searchers, although realizing a lower percent of success for their positive entries, they also had lower positive results from negative cycles and fewer negative to negative results. Again this may indicated a certain degree of purposive behavior and fewer random entries for skilled searcher.

Looking next to measures of success, it should be noted that in all cases but two, the recognition of useful information by novice and skilled searchers could be verified by the investigators. In one case, the searcher terminated the search session indicating he had found useful information when in fact he had not. In the second case, we observed the reverse situation. The searcher kept searching thinking she had been unable to locate useful information when in fact it had been found earlier but not examined carefully enough to identify the useful sections. Surprisingly, in all other cases, investigators were able to verify both successful and unsuccessful judgements by the searchers. Short run success as indicated by the percent of positive iterative entries to system responses is discussed above. However, a large number of these entries were positive or useful throughout a search session. The lowest percent of success was 40% for novices and 42% for skilled searchers. With generally high rates of short run success, it is surprising that the interim success rate is relatively low. Only four of the novice searchers were able to improve their positive results throughout 8 iterative cycles, although all retrieved information on their first search. Of the skilled searchers, 6 searchers were able to improve on their result through 11 cycles. Given the degree of short run success, it is surprising that iterative cycles were not more productive. It seems that successful entries do not necessarily lead to iterative success. This raises the question of whether iteration really is helpful. Well, for novices, at least in this environment, the answers is “ not

very” , and for skilled searchers, the answer is “ some of the time” ! This may, however, reflect the structure of a menu-based system rather than interactive searching in general. Long run success rates are even more problematic for novice and skilled searchers alike. Only one of the novice searchers achieved long run success by pursuing a series of cycles that led to ultimate success. Of the skilled searchers, just three were able to pursue iteration to achieve long run success. This may be due to a “ just one more try” mind set among some searchers who find useful information but make just one more try in case there may be more. This may have been true of at least 3 of the 4 novice searchers who used iteration to find information successfully up until the final iteration. Skilled searchers may be more directed and able to recognize and pursue success through a series of successful iterations. Of six who used iterative cycles successfully throughout their search session, for 11 cycles, three were able to continue through to search termination and two appeared to make “ just one more try” .

7. Discussion

From this study, it is clear that in this particular menu-based environment, short-run appropriate entries were not necessarily parlayed into successful iterative cycles leading to the location of additional useful information. Only half of all searchers were able to realize increased search success through iterative searching. This may be due in part to the structure of menu-based systems since most skilled and novice searchers were able to find useful information from their first cycle. If the first cycle had not been successful, perhaps additional cycles would have identified more information. However, for the skilled searchers who were unable to locate any information on the first try, successive iterations did not lead to any improvement. If this menu-based format is structured to produce maximum benefit for minimum effort and to offer equal access to novice and skilled searchers alike, it appears to have been successful at least in this investigation. There are no significant differences between novice and skilled responses to the system, skilled searchers were able to use iteration to produce

relevant information to a somewhat greater degree than were the novices, but nonetheless, novices were able to find some useful information in every case. Blurring the differences between novice and skilled searching may be good news for the designers of menu-based systems, but it may raise real questions about the rewards of highly skilled searching for the information professional. In any event, the interesting results of this investigation suggest the need for further verification - both in other menu-based formats as well as in other searching environments.

References

[1]Bates, M.J. " Where should the person stop and the information search interface start?" , *Information Processing and Management*, 26(5):575-591, 1990.

[2]Belkin, N.J, Marchetti, P.G., Albrecht, M., Fusco, L., Skogvold, S., Stokke, H. and, Troina, G. " User interfaces for information systems" , *Journal of Information Science*, 17:327-344, 1991.

[3]Saracevic, Tefko, " Selected results from an inquiry into testing of information retrieval systems" , *Journal of the American Society for Information Science*, 22(2):126-139, 1975.

[4]Savage-Knepshield, Pamela A. and Belkin, Nicholas J. " Interaction in information retrieval: trends over time." *Journal of the American Society for Information Science* 50(12):1069, 1999.

[5]Saracevic, Tefko, " Information Science" , *Journal of the American Society for Information Science* 50(12):1062, 1999.

[6]Marchionini, G. Analytical Search Strategies . In *Information Seeking in Electronic Environments*: Cambridge Series on Human-Computer Interaction 9. London: Cambridge University Press. 76-99,1995.

[7]Fidel, Raya, " Moves in Online Searching" , *Online Review* 9(1):61,1985.

[8]Xie, Hong and Cool, Colleen, " The Importance of Teaching ' Interaction' in Library and Information Science Education" , *Journal of Education for Library and Information Science*, 39(4):323, 1998.

[9]Shaw, W.M. Jr, Burgin, R, and Howell, P. " Performance Standards and Evaluation in IR Tests Collections: cluster-based retrieval models" . *Information Processing and Management*, 33(1):1, 1997.

Shaw, W.M. Jr, Burgin, R, and Howell, P. " Performance Standards and Evaluation in IR Tests Collections: vector-space and other retrieval models" . *Information Processing and Management*, 33(1):15, 1997.

[10]Hsieh-Yee, Ingrid, " Search Tactics of Web Users in Searching for Texts, Graphics, Known Items and Subjects : A Search Simulation study" , *The Reference Librarian*, 1998.

[11]Saracevic, Tefko, " Modeling Interaction in Information Retrieval (IR): a Review and Proposal," *Proceedings of the American Society for Information Science*, 33:1-15, 1996.

[12]Wolfram, Dietmar and Dimitroff, Alexandra, " Hypertext vs. Boolean-based searching in a bibliographic database environment: a direct comparison of searcher performance." *Information Processing and Management* 34(6):669-679, 1998.