# A Preliminary Investigation of Search Self-Efficacy

Diane Kelly
School of Information and Library Science
University of North Carolina at Chapel Hill, USA
dianek@email.unc.edu

May 21, 2010

In many studies of online search behavior, it is common for researchers to ask searchers to provide an indication of their search expertise. This is typically done with one or more items that ask searchers to indicate the frequency with which they perform searches, the length of time they have used search engines, or their general expertise (from novice to expert). The basic idea is that differences in search expertise may cause differences in search behaviors and success.

Measuring search expertise was very important in early studies of interactive searching as online information systems often required searchers to have extensive knowledge of search syntax and query languages, and were designed primarily for search experts such as librarians. After search went 'public' in the form of Web search engines, more and more people had an opportunity to search, and to develop search skills. Although it is unlikely that today's average Web searcher's skills are equivalent to the skills of a trained searcher, it is reasonable to assume that the public's general search abilities have increased over the last 10 years and that a greater range of expertise exists.

In today's studies of online search behavior, search expertise is typically used to characterize searchers rather than as a variable that might explain differences in behaviors. The measures that are used to characterize search expertise are quite coarse and general, and are not good at making distinctions among searchers. In today's search environment, it is likely that there are some important variations in search expertise that might be used to understand differences in behaviors. However, there are no contemporary measures of search expertise which are valid and reliable (at least not demonstratively); typically, researchers create reasonably sounding items on an ad-hoc basis. The development of a valid and reliable measure of search expertise would facilitate more in-depth investigations of the relationship between search expertise and search behaviors. This has the potential to impact the responsiveness of the system; for example, if search expertise can be inferred from search behavior then the system can use this information to provide more diverse types of support.

The goal of this work is to identify a measure of search expertise that can be used by researchers to better understand searchers' information seeking behaviors and interactions with information systems. This technical report describes initial investigations into one approach to measuring search expertise based on Bandura's concept of self-efficacy (1977), Compeau and Higgins' (1995) use of this concept in their Computer Self-Efficacy scale, and Debowski, Wood and Bandura's (2001) use of it in their Search Self-Efficacy scale. This report also describes a modified version of Debowski, et al.'s Search Self-Efficacy scale and results of a preliminary investigation of how 23 undergraduates used the scale to characterize their search self-efficacy.

**Origins and Evolution of Self-Efficacy**

Self-efficacy is from Bandura's Social Cognitive Theory (1977). The basic premise of the theory is that environmental influences (e.g., social pressures or unique situational characteristics), cognitive and other personal characteristics (e.g., personality, demographic) and behavior are reciprocally determined. Bandura presents these three components as a constantly interacting triangle. This is called 'triadic reciprocality' or 'reciprocal determinism.'

Bandura posited that the main cognitive drivers of behavior are: (1) beliefs about behavioral outcomes leading to favorable consequences and (2) beliefs about one's ability to perform particular behaviors (self-efficacy) (note that environmental influences are absent in this statement). Bandura (1986) provides a succinct definition of self-efficacy as "People's judgments of their capabilities to organize and execute courses of action required to attain designated types of performances. It is not concerned with the skills one has but with one's judgments of what one can do with whatever skills one possesses" (p. 391). Notice that the focus is on a person's *beliefs* and *perceptions* not their *actual* skill.

Bandura (1977) proposed that self-efficacy is one of the most powerful determinants of behavioral change because it affects a person's initial decision to perform a behavior, the effort expended and persistence exhibited in the face of adversity.

**Computer Self-Efficacy**

Compeau and Higgins (1995) used Bandura's ideas about self-efficacy to create a Computer Self-Efficacy Scale. This work was quite influential and has been cited over 500 times (according to ISI Citation Index). Compeau and Higgins describe previous efforts to create various instruments to measure people's computer abilities and indicate that self-efficacy is a better measure since it goes beyond a simple check-list of component skills, "computer self-efficacy represents an individual's perceptions of his or her ability to use computers in the accomplishment of a task, rather than reflecting simple component skills" (p. 191).

Compeau and Higgins divided self-efficacy into three dimensions: magnitude, strength, and generalizability. Magnitude refers to the level of task difficulty one believes is attainable. Strength refers to the level of conviction about the judgment, and generalizability refers to the extent to which perceptions of self-efficacy are limited to particular situations.

Compeau and Higgins developed and tested a 10-item measure of computer self-efficacy. The measure did not reflect component skills; instead it presented tasks of varying levels of difficulty that a person might execute with a computer. In response to each of the 10 items, individuals first indicated a binary response about whether they could perform the described action (yes or no) and then how confident they were that they could perform the task on a 10-point scale (1=totally unconfident; 5/6=reasonably confident; 10=totally confident). The scale incorporated elements of task difficulty by capturing differences in self-efficacy magnitude (how many items individuals could perform). Self-efficacy strength was captured in the confidence rating. Self-efficacy generalizability was not studied directly.

**Search Self-Efficacy**

Following the publication of Compeau and Higgins' (1995) scale, many specialized computer self-efficacy instruments were created which were tailored to specific software packages (e.g., Excel). There were also a few instruments for measuring search self-efficacy that were created, although most were not created using formal measurement development methods or subjected to rigorous testing and evaluation. One exception is Debowski, Wood and Bandura (2001). Debowski, et al. used classic library science research to identify items for their search self-efficacy scale and performed some validity and reliability testing.

Debowski, et al.'s instrument was designed similarly to Compeau and Higgin's Computer Self-Efficacy instrument, in that a list of items was presented and individuals indicated whether they could perform the task if they wished and their confidence in their abilities to perform the task. Major differences were that Debowski, et al.'s instrument listed items that were related to searching instead of general computer use, and it consisted of 21 items instead of 10.

In the fall of 2008, members of a small research seminar in Interactive Information Retrieval[1] at SILS reviewed the items from Debowski, et al.'s scale and made several modifications to the scale. This included deleting some items and modifying others to make them more contemporary. One major difficulty of creating an instrument related to technology is that since technology changes quickly, items often need to be reviewed and modified (for example, it is unclear how useful Compeau and Higgins' Computer Self-Efficacy scale is today). This, of course, perturbs any previously established validity and reliability.

Ultimately, the scale tested in this report had 14 items and the two-step response was eliminated. Instead, individuals were only asked to indicate how confident they were that they could execute the tasks described by the items. The assumption was that if an individual did not feel they could execute a particular task, then they could express this through the confidence scale. The modified version of the Search Self-Efficacy scale can be viewed in Appendix A.

**Testing the New Search Self-Efficacy Scale**

The subjects who completed the Search Self-Efficacy instrument were enrolled in a study whose primary goal was to investigate query suggestions (Kelly, et al., 2010). In this study, subjects used an experimental IR system to complete four assigned search topics. Subjects completed the Search Self-Efficacy instrument prior to searching. Sixteen of the subjects were females and 7 were males. Subjects' mean age was 21 years (SD=1.9). One subject was a graduate student. Thirteen percent of the subjects were humanities majors, 30% were social science majors, 22% were science majors and 35% were in a professional school.

Table 1 presents statistics describing subjects' responses to each item on the Search Self-Efficacy Scale. The mean values for most items (n=9) were between 7.50 and 7.87, the mean values for three items

[1] Amber Cushing (SILS Ph.D. student), Maureen Dostert (SILS MSIS student) and Xi Niu (SILS Ph.D. student) (Kelly was the instructor and also reviewed items).

were between 6.57 and 6.91 and the mean value for one item was slightly less than 6. Overall, subjects characterized themselves as at least being *reasonably confident* (represented by 5 and 6 on the scale) in their abilities to execute the various tasks. In looking at the minimum and maximum scores, it can be observed that some subjects indicated that they were very confident in their abilities to perform all the tasks (most notably, one subject indicated that they felt confident that they could find articles similar in quality to a professional searcher).

The items that received the lowest mean responses ["Use special syntax in advanced searching (e.g., AND, OR, NOT)," "Find articles similar in quality to those obtained by a professional searcher," and "Devise a query which will result in a very small percentage of irrelevant items on my list"] are unsurprising as they represent more advanced search skills, although overall, subjects were still reasonably confident they could execute these tasks. These were also the items where the greatest variability was observed in subjects' responses. In addition to these items, subjects used a greater range of values in response to several other items ("Efficiently structure my time to complete the task," "Develop a focused search query that will retrieve a small number of appropriate articles," and "Structure my time effectively so that I will finish the search in the allocated time").

**Table 1. Descriptive statistics for items on the Search Self-Efficacy Scale (SD=standard deviation).**

| Item | Mean | Median | Mode | SD | Min. | Max. |
|---|---|---|---|---|---|---|
| Identify the major requirements of the search from the initial statement of the topic. | 7.87 | 8.00 | 9.00 | 1.58 | 5 | 10 |
| Correctly develop search queries to reflect my requirements. | 7.78 | 8.00 | 8.00 | 1.51 | 5 | 10 |
| Use special syntax in advanced searching (e.g., AND, OR, NOT). | 6.57 | 6.00 | 5.00<br>6.00 | 2.00 | 3 | 10 |
| Evaluate the resulting list to monitor the success of my approach. | 7.70 | 8.00 | 8.00 | 1.58 | 5 | 10 |
| Develop a search query which will retrieve a large number of appropriate articles. | 7.61 | 8.00 | 7.00<br>8.00 | 1.53 | 5 | 10 |
| Find an adequate number of articles. | 7.78 | 8.00 | 9.00 | 1.70 | 5 | 10 |
| Find articles similar in quality to those obtained by a professional searcher. | 6.57 | 6.00 | 5.00 | 1.75 | 4 | 10 |
| Devise a query which will result in a very small percentage of irrelevant items on my list. | 5.96 | 6.00 | 6.00 | 1.77 | 1 | 10 |
| Efficiently structure my time to complete the task. | 6.91 | 7.00 | 5.00 | 2.04 | 3 | 10 |
| Develop a focused search query that will retrieve a small number of appropriate articles. | 6.70 | 6.00 | 6.00 | 1.74 | 4 | 10 |
| Distinguish between relevant and irrelevant articles. | 7.87 | 8.00 | 8.00 | 1.60 | 5 | 10 |
| Complete the search competently and effectively. | 7.74 | 8.00 | 8.00 | 1.54 | 5 | 10 |
| Complete the individual steps of the search with little difficulty. | 7.78 | 8.00 | 9.00 | 1.65 | 5 | 10 |
| Structure my time effectively so that I will finish the search in the allocated time. | 7.65 | 8.00 | 9.00 | 1.67 | 4 | 10 |

The Search Self-Efficacy scale is designed to produce a single composite value representing subjects' overall search self-efficacy. A composite value could be produced by, for instance, averaging or summing subjects' responses to each individual item. In order to produce a composite value it must be demonstrated that there is good evidence that all of the items measure the same underlying latent concept (i.e., search self-efficacy), or that responses to items are at least highly correlated.

A principal components factor analysis was conducted to determine if all items loaded on a single factor (concept) and could subsequently be combined into a composite measure. Bartlett's Test of Sphericity and the Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO) were computed initially to determine if the data were good candidates for factor analysis (that is, that the data met the initial assumptions required to conduct a factor analysis). Bartlett's Test of Sphericity showed that the relationship among the questionnaire items was linear [$\chi^2(91)=328.05$, $p<0.001$]. The KMO value was 0.783 which is adequate according to standard heuristics (Mulaik, 1972). A single component was extracted which explained 68% of the variance. The factor loading scores for each item ranged from 0.668 to 0.926, which signals fairly strongly that the items measure the same latent concept (in this case, search self-efficacy). Items were then averaged for each individual to arrive at a composite measure of search self-efficacy. Statistics describing these values are presented in Table 2.

**Table 2. Descriptive statistics of subjects' overall search self-efficacy (composite score).**

| | | | | | | Percentiles | | |
|---|---|---|---|---|---|---|---|---|
| **Mean** | **Median** | **Mode** | **SD** | **Min.** | **Max.** | **25** | **50** | **75** |
| 7.319 | 7.143 | 8.79 | 1.386 | 5.14 | 9.79 | 6.214 | 7.143 | 8.357 |

In general, subjects' average search self-efficacy scores are slightly above the mid-points of 5/6, and ranged from a minimum of 5.14 to a maximum of 9.79. The median shows that the middle score was slightly less than the average and the mode was actually greater than the score at the 75$^{th}$ percentile, indicating that the scores at the high end are outliers to a certain extent. In particular, there were 3 scores of 8.79 and then a single score of 9.71 and 9.79 each.

**Conclusions and Future Work**

The paper presented results of an initial investigation of how search self-efficacy can be measured. Several minor modifications were made to a search self-efficacy scale created by Debowski, et al. (2001) and given to 23 undergraduates who were enrolled in an interactive information retrieval system study. The undergraduates indicated an average search self-efficacy of 7.319 (on a scale of 1-10), indicating reasonably strong confidence in their abilities to perform several tasks related to search. The standard deviation was 1.386 indicating slight variance, but scores were clustered around the larger values of the scale.

While there is likely some variability in this population (undergraduates), the scale may not have been sensitive enough to make distinctions. This scale might, however, be more sensitive if administered to a larger, more diverse set of subjects. It might be that undergraduates (or the age group represented by

undergraduates) are too much alike, and are in general, fairly good searchers, or given that subjects self-selected to participant in an IIR study, they may have more interest in, and knowledge of, searching. If more members of the general public were studied, a larger range of search self-efficacy scores might be observed. It is also the case that this was a preliminary investigation and the sample size is too small to establish solid findings in the context of measurement development. Finally, it may be the case that the items do not reflect a great enough range of search skills (i.e., from very easy to very difficult). For some of the more difficult items, a greater range of responses was observed, but it might be that easy to midlevel skills were overly represented in the item set.

One potential next step for this research is to administer the current version of the search self-efficacy scale to a larger number of more diverse subjects to see if it is better at distinguishing among people. Another potential direction is to engage in activities related to the creation of new search self-efficacy scale. Such activities would include the collection and testing of new items that describe search skills possessed by people who have various expertise, as well as validity and reliability testing of the resulting instrument.

**References**

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191-215.

Bandura, A. (1986). *Social foundations of thought and action*. Prentice Hall: Englewood Cliffs, NJ.

Compeau, D. R. & Higgins, C. A. (1995). Computer self-efficacy: Development of a measure and initial test. *MIS Quarterly (June),* 189-211.

Debowski, S., Wood, R. & Bandura, A. (2001). The impact of guided exploration and enactive exploration on self-regulatory mechanisms and information acquisition through electronic enquiry. *Journal of Applied Psychology, 86*(6), 1129-1141.

Kelly, D., Cushing, A., Dostert, M., Niu, X., & Gyllstom, K. (2010). Effects of popularity and quality on the usage of query suggestions during information search. *Proceeding of the ACM Conference on Human Factors in Computing Systems (CHI)*, Atlanta, GA.

Mulaik, S. A. (1972). *Foundations of factor analysis*. Boca Raton, FL: CRC Press.

**Appendix A: Search Self-Efficacy Scale**

| LEVEL OF CONFIDENCE: | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** |
| **Totally Unconfident** | | | | **Reasonably Confident** | | | | | **Totally Confident** |

| *I can:* | *Confidence (1 – 10)* |
|---|---|
| **1. Identify the major requirements of the search from the initial statement of the topic.** | |
| **2. Correctly develop search queries to reflect my requirements.** | |
| **3. Use special syntax in advanced searching (e.g., AND, OR, NOT).** | |
| **4. Evaluate the resulting list to monitor the success of my approach.** | |
| **5. Develop a search query which will retrieve a large number of appropriate articles.** | |
| **6. Find an adequate number of articles.** | |
| **7. Find articles similar in quality to those obtained by a professional searcher.** | |
| **8. Devise a query which will result in a very small percentage of irrelevant items on my list.** | |
| **9. Efficiently structure my time to complete the task.** | |
| **10. Develop a focused search query that will retrieve a small number of appropriate articles.** | |
| **11. Distinguish between relevant and irrelevant articles.** | |
| **12. Complete the search competently and effectively.** | |
| **13. Complete the individual steps of the search with little difficulty.** | |
| **14. Structure my time effectively so that I will finish the search in the allocated time.** | |