Journal of Librarianship and Information Science

Scaffolding in Information Search: Understanding Novice-Expert Differences

Journal:	*Journal of Librarianship & Information Science*
Manuscript ID:	Draft
Manuscript Type:	Original Manuscript
Keywords:	scaffolding, expert-novice comparison, information search, satisfaction, postgraduate

Abstract:	This study aims to investigate how expert scaffolded training could help, from novice postgraduate students point of view, and foster development of information search ability among postgraduate students. Using a quasi-experimental design over a year and a half, eight doctoral students (novice searchers) participated in a series of five sessions with an expert searcher who was an information professional. A novice-expert comparison examined the differences between novices and experts in information searching; and the effect of scaffolding sessions in which the expert information searcher helped novice information searchers was examined. Findings showed differences existed between the novice and the expert searchers in use of complex formulation of query statements, choice of keywords, and operators. Scaffolding sessions with the expert searcher resulted in self-reported and observable improvement in information searching among the novice searchers. The paper concludes with a discussion of the design of information retrieval systems and recommendations for library programs to support the continued development of research students' information literacy skills.

Introduction

Consulting an academic library's catalog, searching for scholarly information online, and conducting rigorous database searches to support their studies are an indispensable part of the research students' academic life. However, several studies have found that the information literacy level of university students, including postgraduate students, is still far from satisfactory (e.g.; Chu & Law, 2008; Graham and Metaxas, 2003; Laverty, Reed, and Lee, 2008). Scholars have tried to address the poor information searching skills by analysing the searching behaviours of students so that remedial programmes or interventions could be developed to improve the situation. Some made comparisons between the searches of novices and that of experts (e.g. Hölscher & Strube, 2000). Others explored the effects of domain or subject knowledge on information seeking results (e.g. Duggan & Payne, 2008). Despite the existence of these studies, few have been targeted at examining the influences of expert scaffolding support on the development of information literacy among novices. This paper explores how novice searchers interpret an information search task, what search strategies they used, and the formulation and revision of a search statement over a year and a half period. The novice searchers were first-year doctoral students at the Faculty of Education of Hong Kong University (HKU), who towards the end of the our study, were just a few months away from proceeding to their third year of their Ph.D. program. Comparisons were then made with how an expert searcher would carry out the same task, thus, identifying potential areas for improvement in the research students' searching skills. Further, novice searchers' perceptions on helpfulness of the expert scaffolding training were also examined.

Literature Review

Information Search Strategies and Query Formulations

A typical search scenario that the doctoral students encounter is sitting in front of their computers and pondering upon what types of research materials they would like to locate. According to Chu and Law (2007a), research-oriented sources include refereed journals, review articles, books, free Web resources, bibliographies, conference papers, and theses. There are numerous databases to choose from and their usefulness is highly dependent upon whether the contents of the databases

match the students' fields of interest. ERIC[1], ProQuest[2], and Academic Search Fulltext Elite[3] are some academic subscription databases considered useful to students in the education discipline; while IEEE Xplore[4], ScienceDirect[5], Springer-LINK[6] are some well-known databases for engineering students (Chu and Law, 2005). Meanwhile, some databases are equally important among students of all disciplines, such as, the university library online catalogues and the highly popular Google and Google Scholar[7]. The databases that research students use may, in turn, fall into one of the three main Information Retrieval Systems (IRS) categories: 1) online public access catalog (OPAC); 2) free Web search engines; or 3) online bibliographic or full-text databases subscribed to by academic libraries that may be either multi-disciplinary or discipline-specific (Chowdhury, 2010). In this study, a series of actions and decision-making in the selection of databases and information retrieval systems for various types of information constituted what we term as "information seeking strategies" (Chowdhury, 2010). The effectiveness of the information search depends not only on whether the correct database is being used, but also on how effectively a search query is formulated by students. There have been a few studies on query statement construction and various tactics have been formulated and developed over different time periods (Bates, 1979; Fidel, 1985; Hembrooke, Granka, Gay, and Liddy, 2005; Hölscher and Strube, 2000). Most recently, nine search tactics being identified, Hembrooke et al. (2005) investigated the effects of domain knowledge on search term selection and query formulation and reformulation.

Novice-Expert Studies on Information Search Behavior

Expertise or novice-expert studies are important because these researchers believe that a good understanding on how people become experts or how experts actually perform the required tasks can help shorten the novices' learning curve in becoming experts themselves (Larkin et al., 1980). Brand-Gruwel, Wopereis, and Vermetten (2005) attempted to decompose information problem-solving into

[1] www.eric.ed.gov/
[2] www.proquest.com/
[3] www.ebscohost.com/academic/academic-search-elite
[4] ieeexplore.ieee.org/
[5] www.sciencedirect.com/
[6] link.springer.com/
[7] scholar.google.com/

different skills and sub-skills in order to design instructions. They hoped that it could foster the development of information problem-solving skills and that their "expert-novice analysis gave more insight in which skills need more attention and need to be further analyzed" (p. 502). Hölscher and Strube (2000) compared participants rates of solving information tasks, and query formatting behaviors. The findings indicated that expert searchers made use of Boolean operators, modifiers, phrase searches and other IR system features more often than novice searchers did.

Examining the interactive effects of both cognitive style and on-line database search experience on the search performance measured by both the time and the average number of nodes visited for an information task, Palmquiest and Kim (2000) found that cognitive styles significantly affected the search performance of novice searchers only, but not that of experienced searchers. In exploring the way search terms were derived, Hsieh-Yee (1993) found that when searching in an unfamiliar domain, expert searchers used the thesaurus more often for term suggestions, used more synonyms, and used more term combinations than they did when searching in a familiar domain. The study concluded that users' search experience (information search expertise) affected their use of search strategies more than their domain knowledge, and hence, played a more important role in information searching.

Scaffolding support and information literacy acquisition

Vygotsky (1978) proposed the idea of the "zone of proximal development", suggesting that students learn through the "guidance and collaboration of more capable others". More recently, Halttunen & Järvelin (2005) restated that scaffolding refers to the assistance offered to students that enables them to successfully complete a task. In terms of information search skills acquisition, some scholars consider studies on how experts search as one possibility in providing help to facilitate the development of novice to expert (Larkin, McDermott, Simon, & Simon, 1980). Specifically, Larkin et al., (1980) in their article claimed, "Our growing understanding of an expert's knowledge and the kinds of processes an expert uses when solving problems enables us to begin to explore the learning processes needed to acquire suitable knowledge and problem-solving processes." (p. 1,342). Although

the idea of scaffolding is popular in the general education arena, it is less common in the field of information literacy education. Existing studies generally compared the similarities and differences between novices and experts with respect to how they formulate their searches and make use of various search tactics. However, little studies have been conducted to specifically examine the influences of expert scaffolding support on the development of information literacy skills among novice searchers.

Our Study

This research attempts to address the gap through a series of expert scaffolded training sessions to help novice searchers develop knowledge of IRS and search skills. The issue of how satisfied students would be in the usage of different databases when they intended to look for a certain information type has not been studied before, let alone examined with an expert-novice comparison approach. Regarding search query formulation, although there has been some previous research conducted using the novice-expert study approach, they were not focused on studying or directly addressing the academic searching behaviours of doctoral students. Further, this study looks into details how expert scaffolding support influences the development of information literacy among novices from the novice's perspective.

Research Questions

This study aims to investigate how expert scaffolded training could help, from the novice postgraduate students point of view, foster the development of information search ability among postgraduate students. In order to advance this research aim, two main research questions have been devised:

- *RQ1*: What are the effects and influence of the scaffolded search sessions, from the novice searchers' perspective, on their development of search skills?
- *RQ2*: Are there any differences observed between novice and expert searchers in terms of IRS usage and query formations?

METHODOLOGY

Participants

Eight research postgraduate students in their doctoral program in the Faculty of Education of the University of Hong Kong were recuited employing purposive sampling. The eight novice searcher participants were given ID numbers from P1 to P8, see Table 1. The expert searcher is referred to as ES. He has a Master in Library Science and a PhD in Education focusing on information literacy. He also has 12 years of experience as an instructor of information literacy training workshops and has published in the area of information literacy and training for over 10 years.

[Insert Table 1 here]

Procedure

This study tracked the changes of the search behaviors of eight novice searchers and the ES over a year and a half period. During the entire research period, each novice searcher attended five research meetings. In the first meeting, each novice searcher was required to come with a self-designated research topic which they hoped to work on for the rest of the research meetings. The research topics included, for example, professional development models, knowledge management and open-source software, activity theory and blogging, etc. In each research meeting, the novice searcher was given 15 minutes to search for information relevant to their research topic independently. This was followed by a 15-minute search session conducted by the ES on the same topic, as a way of demonstrating to the novice searcher how their search could be enhanced by additional search skills. The ES also responded to the inquiries from the novice searchers and helped them look for relevant information by explicitly showing them how search queries could be formulated in various databases or search engines to advance desirable search results. The interactions between the novice searchers and the ES and the computer were videotaped and the searchers were asked to think-aloud their entire search process.

At the end of each research meeting, the eight novice searchers were asked to rate their satisfaction levels, on a 5-point Likert scale, for both the unaided and scaffolded sessions based on

their perceptions about the search quality and results relevancy. A self-report questionnaire was administered to the eight novice searchers, at the end of three of the five meetings -- the first, the third and the fifth meetings. On each questionnaire, the participants were asked to rate their perceived familiarity with and the importance of the various databases, along with search knowledge and skills based on a 4-point Likert type scale so no neutral option was available.

In addition, a final interview was conducted at the end of the fifth meeting to probe students' perceptions of the development of their search strategies over time, as well as their comments and opinions on the helpfulness of the scaffolding support provided by the ES in facilitating their information literacy development.

To answer the two research questions, five kinds of data were collected and analyzed:

RQ1: What are the effects and influence of the scaffolded search sessions, from the novice searchers' perspective, on their development of search skills?

a) novice searchers' rating on satisfaction levels towards the unaided and scaffolded sessions;

b) novice searchers' responses to the questionnaires; and

c) transcriptions of the final interviews conducted at the end of the fifth meeting.

RQ2: Are there any differences observed between novice and expert searchers in terms of IRS usage and query formations?

a) search statements used by students when searching various databases; and

b) transcriptions of students' think-aloud protocols as they verbalized their thoughts and actions when performing database searches.

Analysis Tool for Information Search Tactics

In order to make an informative comparison of the differences between the search skills of novice and expert searchers, we borrowed Hembrooke et al.'s (2005) framework to analyze the search statements harvested from all search sessions by categorizing the search tactics used into nine major categories. Definitions and examples of these tactics are provided as in Table 2. All the search statements used by students and the ES were categorized into the nine tactics accordingly.

[Insert Table 2 here]

Findings and Discussion

The following discussions are divided into two major sections to answer to RQ1 and RQ2 respectively: (A) Influence of the scaffolded sessions on novice searchers; and (B) Differences observed in novices and experts when searching for information.

(A) Influence of the scaffolded sessions on novice searchers

Comparison of Satisfaction Level between Unaided and Scaffold Sessions

Satisfaction levels with each of the 36 research meetings was measured by a scale from 1 to 5 (1 as "totally dissatisfied" and 5 as "very satisfied"), based on the participants' perceptions about the search quality and relevancy of results in the session. The mean rating of satisfaction level was calculated and is presented in Figure 1. The satisfaction levels for the scaffolded search sessions were significantly higher than those for the unaided search sessions for all participants ($t = -9.36$, $p < 0.001$).

Among those 36 meetings, there was an increase in satisfaction level between the unaided and scaffold sessions was found for the majority (n=32) of them. The overall average scores of the unaided sessions and scaffolded sessions in all 36 research meetings were 3 and 4.35 respectively. The current study was based on the premise that it was the differences in the searching strategies (choices of databases used) and the tactics used for query formulation between novice searchers and the ES that contributed to the improvement in the satisfaction level.

[Insert Figure 1 here]

Participants' Developments in Information Search

Knowledge of Information Sources

The participants were asked to rate their familiarity with various resources and databases at the end of the first, third and the fifth meetings. Table 3 presents a comparison of the average ratings

by the novice searchers of their perceived familiarty with various databases at the first and the fifth sessions. The mean of the total ratings in the final survey is 3.03 (increased from 2.70 at the 1st session), indicating that students has improved from "a little familiar" to "familiar" with the listed electronic databases and internet resources upon the completion of all five research meetings. As shown in Table 3, the result of Wilcoxon signed rank test illustrated that the overall rating of perceived familiarity with various databases increased significantly after the five scaffolded sessions ($z = -2.366$, $p = 0.018$), implying a positive influence of the scaffolded training sessions on students' familiarty with various databases.

[Insert Table 3 here]

To further substantiate the positive quantitative findings, the interviews revealed that, 7 out of 8 of the participating students commented that their knowledge of different source databases had been enhanced due to the scaffolding sessions. Many acknowledged that the ES had helped them become aware of many advanced level databases which they were not familiar with before they had the scaffolding sessions. For example, P1 mentioned in the interview:

> *"I think after the five research sessions, I do search differently. It's because the five sessions have introduced me more resources that I previously neglected."*

Some other similar comments are as follows:

> *P2: "I learn more about database system I can access."*
>
> *P8: "If I did not attend these scaffolding sessions, I think maybe I will miss some important references/source. I may not be able to find some useful information because I will miss some databases."*

One of the values of involving an expert searcher in the training seems to lie in his resourcefulness in terms of knowledge of a great variety of information databases. By guiding the novice searchers to the less commonly known databases, the resource repertoire of novice searchers was broadened.

Knowledge and Skills Acquisition: the novice searchers' perceptions

Apart from knowledge about various databases and information sources, students substantially developed their information search knowledge and skills. At the end of the study, the mean rating of the students' perceived familiarity with various information search skills was 3.16 (increased from 2.86 at the 1st session), as shown in Table 4, indicating that the students has improved from "a little familiar" to "familiar" with the skills that helped them locate and retrieve information upon the project completion. This survey result is well-supported by the interview findings. Five out of eight students commented that they acquired more knowledge of different search skills and techniques during the scaffolding sessions. In particular, students found that the seldomly used Boolean operators offered them essential help in generating more relevant search results. P4 made a comment on this aspect in her interview session:

> *"I learnt more about different kinds of electronic databases and how to use the truncation, such as 'and, or, not' which I seldom used before."*

[Insert Table 4 here]

While the participants acknowledged their increased familiarity with various search knowledge and skills after the completion of the five scaffolded sessions, a statistical analysis was not significant ($p<.05$), except for the item "keyword search". Such discrepancy could be explained by three aspects of the study: (i) different levels of self-evalution evaluation on the pre- and post-tests; (ii) design of the research meetings:

The design of the pre- and post-tests relies heavily on participants' self evaluation of their familiarty with various information search knowledge and skills. It requires the participants to rate their familiarity with the given skills and knowledge according to a scale from 1 to 4 (from "not familiar" to "very familiar"). It included no questions which objectively assessed participants' attainment of skills or knowledge. Also, the post-test was administered without the completed pre-test being presented to the participants. Hence, the participants did not rate their perceived familiarity according to a comparison to their ratings in the pre-test. As participants, through the five research sessions, got to know more about various search skills and knowledge, it is possible that they rated

their perceived familiarity even lower to some items as they realized how little they actually knew about the topic.

Further, it is also possible that the design of the research meetings contributed to the rather insignificant improvement, though the effectiveness of such meetings in search skills development was recognized by most of the participants. One possible explanation is that participants were not given enough practice opportunities in the sessions, as suggested by P5, which hindered her from becoming familiar with the Boolean operators:

> *"I think I am not searching very differently when compared to the way before. I still have very little knowledge of the operators. I think the reason for this is that I don't have a lot of opportunities to practice."*

This comment made by P5 offered an important insight that the effectiveness of such scaffolding sessions in fostering the development of search ability among novices will be maximized when sufficient opportunities are given to the novice searchers to practice the knowledge acquired from the demonstration conducted by the ES.

Information Search Ability

After the 5 scaffolded sessions, most students indicated an improved information search ability in the sense that they are more able to search for results that are relevant to their topics. P2 gave a more concrete interpretation of how the scaffolding sessions helped him narrow down his search results:

> *"After the five sessions, I know how to use more Boolean search function. Before I just search in simple ways so I just get a lot of output (maybe a thousands) and after the sessions, now I tried to make it more specific by using Boolean search like "and", "or", "not" and the truncation."*

P6 acknowledged the help of the scaffolding sessions which shortened her learning curve to become a more competent searcher:

> *"I think if I didn't have those training sessions, I probably will still figure it [the search techniques] out, but it will probably take longer. So, by having those sessions, I have learnt*

how to come up with better search phrases, or how to combine search terms. The learning curve sort of get shorter."

Affective and Cognitive Development

Apart from the hard skills, students also indicated some cognitive benefits from the scaffolded sessions as they became a more sophisticated and confident searcher. For example, P3 indicated that the search skills which she learnt from the scaffolding sessions helped her become a more confident and capable searcher:

"I have equipped many skills for researching, for example some Boolean operators…I think training can help me to have more confident. Because searching is like swimming in the big ocean and we need to find a tool, and if we do not have adequate skills or enough skills, we will easily be very hopeless and despair, you don't know how to move further."

(B) Differences observed in novices and experts when searching for information

To further inform the design and implementation of an expert scaffolded training on information search skills, a comparison was made to understand the differences in search performance between the novice and expert searchers.

Information Retrieval System Usage

Figures 2a and 2b show the distribution of the 3 main IRS usages for seeking each of the four different types of research materials by novice searchers and the ES respectively. In those 36 research sessions (each participant had five meetings; however, 4 out of the total of 40 meetings were lost and this paper reports the data from 36 meetings), the participants spent approximately 35% of those sessions searching for journal articles, 25% for theses, 23% for books and 17% free Web resources. As the decisions on sources types for the information task in each research session were made solely on the basis of the participants' own preferences, this supports the previous findings of Chu and Law (2007a) that journal articles were ranked highest in terms of perceived importance by doctoral

students. This also supports the common perception that refereed journals are more scholarly and respected by researchers (Morner, 1993), and thus are better reference sources for academic research.

[Insert Figure 2a here]

[Insert Figure 2b here]

Overall, the main IRS used for various source types in both novice and expert searchers shared a high degree of similarities -- online academic databases used mainly for searching journal articles and theses, OPACs mainly used for books, and Web search engines for free Web resources. One notable phenomenon that occurred only in the novice searchers' searching approach was that Web search engines were found to be used for searching all four source types. One possible reason is that "using the Internet to find information has become an integral part of everyday life" (Hembrooke et al., 2005, p. 861). As this particular aspect of the findings might suggest, even doctoral students looking for academic information are not an exception. It also reflects the idea that the Web has gradually evolved into an educational tool for exploring and learning purposes (Kang and Fu, 2010; Lazonder, Biemans, and Wopereis, 2000).

In contrast to novice searchers, the ES adopted a straightforward approach to the information task of locating journal articles and theses, resorting to a single resource - online academic databases. In the interview, the ES commented that he focused on online academic databases for journal articles and theses because:

> *"when searching through Web search engines, you may not be able to find the articles that are included in academic databases; even if you are able to find them, you may not be able to access the full version of the article because, usually, you need to pay for the authorized access."*

This points to one of the major issues that information specialists highlight about Web search engines, their inability to retrieve "accurate", "truthful", "authoritative", extensive, and simply richer information from the "Deep Web" (Devine and Egger-Sider, 2009, p. 11). The deep Web is where Web search engines fail to gain access to those Web information resources that can only be accessible either through "authorized access" or by "activating an appropriate program" (Chowdhury, 2010),

and such failure may be attributed to the indexing limitations of the Web-searching programs called Web spiders, Web crawlers or robots (Devine and Egger-Sider, 2009; Lewandowski, 2005; Segev, 2010), as well as the fact that some IRS require a subscription fee for users to access full information. Therefore, being aware of the deep Web issue, expert searchers when searching for scholarly journal articles, theses and dissertations mainly rely upon the high-quality academic databases that provide accurate research-oriented papers.

Database Usage

Table 5 displays the databases used by P1 to P8 and the ES when searching for different source types. While the frequency of database usage varied from the beginning search sessions to the latter ones, the usage was calculated as total for easy comparison between novice and expert searchers. At first glance, the databases used by novice searchers to search for materials of each particular source type were more diversified than those used by the ES. The number of databases used by the ES were limited to three or at most four while the number of databases used by the novice searchers could go up to as many as 10. All databases used by the ES were also used by novice searchers except WorldCat[8] and Networked Digital Library of Theses and Dissertation[9] (NDLTD).

[Insert Table 5 here]

Databases for Journal Articles

EBSCOhost, ProQuest, and Web of Science were the top three most frequently used databases for a journal article search in the research sessions of both novice and expert searchers. This finding is similar to what has been found in several previous studies in which Academic Search Fulltext Elite (a subset of EBSCOhost), ERIC (a subset of EBSCOhost), and Web of Science, and ProQuest were all rated within the top ten databases by the doctoral students in Education field (Bar-Ilan, Peritz, and Wolman, 2003; Talja and Maula, 2003; Chu and Law, 2005).

[8] www.worldcat.org/
[9] www.ndltd.org/

Databases for Books

Overall, the top three databases for searching for books for both P1 to P8 and the ES combined in descending order of usage frequency, were: 1) HKUL Catalogue, Dragon; 2) WorldCat; and 3) HKALL and Google Books (both having an equal frequency of usage). Two issues are worth noting here. First, the usage of HKUL Catalogue (Dragon) by the novices was considerably higher than that of the ES. When being interviewed, the ES explained,

> *"If I found that the novice searchers were able to get satisfactory results from the HKUL Catalog, I would not repeat using it in the scaffolding support session. I wanted to show them an alternative way of finding books – using WorldCat."*

Therefore, the number of visits to HKUL Catalogue by the ES would have been conceivably higher if he had performed the search on his own without the presence of the unaided search session. Secondly, no participants used the WorldCat for a book search. Out of the 8 participants, half of them had explicitly stated that either they did not know WorldCat at all or they just seldom used it. The ES had once explained the advantages of using WorldCat to one of the participants (P3):

> *".....HKALL and HKU libraries have been very useful and so the importance of WorldCat has decreased relatively."*

Besides the "diminishing importance" of WorldCat, P1 showed a possible hint of reluctance to use WorldCat:

> *"WorldCat, I rarely go to that because I discovered many of the books I need can be accessed through the [local] library already. Also, I just wonder, even if I manage to find a book I need on WorldCat, how long would it take for me to request the book? It might not be as convenient."*

As inferred from the statement above, the time factor in borrowing books from overseas is considered by novice searchers in the internet era where one of the major characteristics is the need for the instant gratification of the information needs (Agee, 2007; Prensky, 2001).

Databases for Theses

ProQuest proved to be the most popular database for dissertation and theses searches by both novice searchers and the ES. While the second most popular database for novices to search for dissertation and theses was the Digital Dissertation Consortium, the ES's next favorite one after ProQuest was WorldCat. Similar to what was discussed in the previous section, no single participant used WorldCat for theses search.

Databases for Free Web Resources

As indicated in Table 5, an apparent contrast in the preferences of databases used to search free Web resources was observed. The novice searchers tended to favor Google Scholar more than Google. On the contrary, it was the opposite for the ES, who tended to use Google most of the time. Such discrepancy was attributed to their different interpretations of what free Web resources actually meant. For novice searchers, most of the time when they searched for academic information on the Web, they were actually looking for and paying attention to journal articles, theses, or dissertations that were accessible through the Web. It was exactly what Google Scholar was intended to do when launched in late 2004 as a subset of Google. However, to the ES, his interpretation of "free Web resources" meant something entirely different in nature as he explained:

> "*Beside Google Scholar, I think I'd like to suggest another way of searching for information. How about we will try to search by some either top scholars or top research centers in the area? How about let's try for some research centers for the area?*"

Query Formulation

The eight novice searchers conducted a total of 192 searches in the unaided search sessions while the ES conducted 127 searches in the scaffold search sessions. The distributions of different query tactics used by the novice searchers and the ES are summarized in Figure 3. Since it is possible for a search query to contain more than one tactic, this explains why the percentages of all the tactics used do not add up to 100. The top three tactics used by novice searchers were topic terms, backtracking, and elaboration while the ES used, in descending order, topic terms, elaboration, and

system modifiers. The following subsections includes a more extensive comparision of the query formulation tactics used by the novices and the expert.

[Insert Figure 3 here]

Elaboration

The single largest difference between the novices and the ES was the use of Elaboration in their construction of a search query. The ES used it in 71% of his search queries and novices only 52%. Throughout the search processes in all the 36 scaffolded search sessions, the ES drew upon two main resources to construct search queries: 1) the students' articulation of their information needs; and 2) terms that the students found useful and relevant in the sources that the ES or the students themselves managed to retrieve during the session.

Information needs at the doctoral level usually tends to be relatively specific even at the early stage of their study, and the search topic may contain more than one or two concepts. Some topic examples articulated by the participants under investigation were "collaborative writing using Wiki technology in upper primary English language classroom" (P7), "action research/teaching method/pedagogy enhancement" (P3), "use of think-aloud protocols in SLA research, protocol analysis, validity and reliability of think-aloud protocols" (P5). This might require a very heavy cognitive load when trying to manage and grapple with multiple search elements and concepts (Debowski, 2001). When articulating his justification of the finding, the ES commented,

> *"Regarding the fact that I used elaboration considerably more than the novice searchers did, I believe this is to do with my longer search experience – I am more able and experienced in managing different search terms in a logical way which aids a better search."*

One example can be found in a search session with P2, the ES addressed this particular issue and provided an insight on how to deal with it:

> *".....then let's try to construct a search that would use all or most of these words.....first of all, put them into different concepts -- two or three concepts. Put the words that belong to the same concept together....."*

When doing the scaffolded search session with P3, the ES also suggested not only to draw the search terms from one's own knowledge store, but the others' as well:

".....we included the phrases that we can think of and also included the phrases used by others....."

Listed below was one of the search statements the ES formulated with regard to P2's information need on topic "well-being in school" versus another search statement by P2 in the unaided session :

ES: (well-being or vitality or life satisfaction or positive affect or negative affect or "locus of control" or burnout or self-esteem) w/5 (goal or engagement) w/5 (school or student or pupil)

P2: student* or school* AND well-being AND goal*

Topic Terms and Backtracking

Both novice searchers and the ES shared similar percentages in the number of times they used Topic Terms and Backtracking. In about 90% of the search statements both of them used Topic Terms and about 50% of the time Backtracking was used. This finding appears to be contrary to Hembrooke et al.'s (2005) hypothesis that both Topic Terms and Backtracking (a subset of Redundancy) are considered as the search tactics that usually only novice searchers resort to due to their inadequate resourcefulness in generating search terms. The ES believed that:

"It is natural that topic terms was the most used tactics for both novice searchers and me because it is the most basic way of constructing a search query – you look at your search topic, extract some related topic terms, and do a search on it."

What this implies is that formulating a clear topic and focus is a pivotal step before actually engaging in the searching action. The clarity of the topic influences on the quality and relevance of the information retrieved. Perhaps, this also provides an insight into why, in Kuhlthau's Information Seeking Process (ISP) model, the Formulation stage, the fourth stage of the six-stage model, is considered a turning point in the whole information seeking process in which the information seeker

finally formulate his or her own information topic and experiences feelings of decreased uncertainty, increased confidence and a sense of clarity (Kuhlthau, 2004).

Broadening and Refining

Broadening and refining the search terms semantically was seldom employed by either novice searchers or the ES. Novice searchers were found to use these two tactics slightly more often than the ES. One possible reason was that the information need, after all, was internally determined, defined, and generated by the student searchers. As a result, they were more able to draw on ideas to modify the search terms.

Kitchen Sink, Poke-and-hope, and Plural Making/Taking

These three tactics are characterized by trial-and-error in which the searchers tried out their luck by entering some uncertain terms. This explains the fact that novice searchers used these tactics on some occasions while the ES did not use any Kitchen Sink or Poke-and-hope tactics at all. The ES explained that he did not make use of these tactics because

> *"Kitchen Sink and Poke-and-hope, as the names suggested, do not involve a logical arrangement of search terms. These tactics rely heavily on the luck of the searcher – you change a little bit of your search query, hoping to get a better result. Novices may resort to these tactics because of their lack of knowledge on other more useful search skills."*

Operational Tactics

In order to construct an effective search query, it is not sufficient for the information searcher to only select appropriate search terms. Applying different search operators in an effective manner is an essential skill. Besides Boolean operators, users also need to know how to use other system special modifiers, such as parentheses, quotations, asterisk, or other modifiers for truncation, proximity (Chu and Law, 2007a; Fidel, 1985; Hölscher and Strube, 2000). An additional operational tactic, System Modifier Usage, was therefore examined.

The common system modifiers that help a searcher to build an effective search statement usually include parentheses ( ), quotation marks " ", question marks ?, asterisk *, and w/. Novice searchers and the ES demonstrated a significant discrepancy in the use of the various system modifiers, 41% of the search queries of the novices included system modifiers versus 70% for the ES. Even among the 41% of the search statements in which the novices used system modifiers, these were limited to parentheses, quotation marks, and asterisk most of the time. Such discrepancy can be best explained by what P5 once said:

".....I noticed the ES knew how to use different symbols to find the results. But for me, I guess I learned it before, maybe in the library workshop, or somewhere else, but I did not remember. Actually I cannot remember the meaning of symbols and how to use them. So I did not use the symbols in my search."

All novices in the study raised concerns about and demonstrated difficulties in using system modifiers in one way or another. Some would specifically ask what a particular symbol meant when they observed that the ES used it in the scaffold sessions. What made the issue more complicated was that the symbols, operators or modifiers used for the same function might vary from system to system. P7 addressed the issue without knowing appropriate terms to describe the problem:

"The problem with searching on my own is that I don't know what to put in the search engine. Also, I don't know how to cope with different databases -- it turns out that different search engines -- or databases -- have different 'requirement'.....so I just input what comes to my mind."

Another example mentioned by P5 who has just learnt how to search with a phrase in different databases:

".....I think for different databases they use different symbols to mark it as a phrase. For example, in EBSCOhost maybe we can use parentheses to indicate that it is a phrase. But in Web of Science, we cannot use parentheses. We need to use quotation mark to indicate it as a whole phrase, not separated words....."

Concluding Remarks

The findings of this study suggest that expert scaffolded sessions were well-received by novice searchers as a means to foster information search skills development. Therefore, it may be valuable to incorporate expert scaffolding sessions into the teaching of information search skills. One-to-one or one-to-many expert-novice searching sessions could be arranged within formal curriculum or as an optional library workshop to help facilitate the development of postgraduate students information search abilities. This approach has been found beneficial and welcomed by staff and students in a variety of research fields (e.g., Washington-Hoagland & Clougherty, 2002; Bellard, 2007). If time and human resources do not permit this personal training, scaffolding could be provided through taped videos uploaded onto a library's webpage where expert searchers perform examples of searches to illustrate various skills and techniques in performing effective searches. This online approach was found favorable by students in the study conducted by Kuruppu and Gruber (2006).

The findings derived from the novice-expert comparison provided a well-informed insight into how information retrieval systems and databases could be designed to facilitate a search conducted by novice searchers. Many of the novice searchers that participated in this study indicated that they were not familiar with the search operators, did not know the functions of them, and hence they were unable to apply them when they contructed a search query. Other participants reflected that the issue became more complicated because the operators used for the same function might vary from system to system. This suggests that either future systems should be constructed based on a common guideline, which may take a long time, or a clear user guide should be made easily accessible to users on IRS. Furthermore, it was found that novice searchers tended to resort to the Web search engines. This suggested that the Web has taken an increasingly important role for information seekers in education. As it is inevitable for information seekers to find resources from the Web, it may be helpful if Web search engines include more customized options to help novice searchers to refine their search results.

The findings suggest a variety of areas that could be developed for information literacy instruction for postgraduate research students. For example, many participants demonstrated a lack of knowledge of operational tactics, either they were not knowledgeable of the functions of the

operational symbols, or they were not able to identify the proper symbols to be used in a specific database. As these operators are important in helping the searchers to search effectively, it is important that sufficient training is provided to postgraduate students who need to conduct many information searches over their course of studies. To address this issue, students should be provided with the opportunity to develop information search skills either through training incorporated into formal curriculum, or through library self-registered consultation sessions. At the very least, postgraduate students should have the opportunity to attend an introductory course which covers various major IRS that might be useful for their research in their subject area. As students were found to rely heavily on Web search engines, more effort should also be devoted to the teaching of online academic databases, such as WorldCat and NDLTD, which are of enormous help to postgraduate students in finding credible and useful resources.

To conclude, this study aimed to examine the effects of scaffolded search sessions on novice searchers' development of search ability, as well as to provide empiricial evidence to inform future implementation of expert scaffolded traning through looking at the differences between novice and expert searchers. This study deliberately included only eight participants to allow an in-depth exploratory novice-expert comparison. While the relatively small sample size may preclude a generalizable statement on the search behavior of postgraduate students, it enabled us to conduct an intensive analysis and hence understand the search behavior of individual postgraduate students. Follow-up studies which include a larger data set will help us understand whether the findings of this study are generalizable.

References

Agee, J. (2007), *Acquisitions go global : an introduction to library collection management in the 21st century*, Oxford, Chandos.

Bar-Ilan, J., Peritz, B. C., and Wolman, Y. (2003), "A survey on the use of electronic databases and electronic journals accessed through the Web by the academic staff of Israeli universities", *The Journal of Academic Librarianship, 29(6)*, pp. 346-361.

Bellard, E. M. (2007), "Information literacy needs of nontraditional graduate students in social work", *Research Strategies,* 20(4), 494-505.

Brand-Gruwel, S., Wopereis, I., and Vermetten, Y. (2005), „Information problem solving by experts and novices: analysis of a complex cognitive skill", *Computers in Human Behavior, 21(3)*, pp. 487-508.

Chowdhury, G. G. (2010), *Introduction to modern information retrieval* (3rd. ed.), London, Facet.

Chu, S., and Law, N. (2005), "Development of information search expertise: research students' knowledge of databases", *Online Information Review, 29(6)*, pp. 621-642.
Chu, S., and Law, N. (2007a), "Development of information search expertise: Research students' knowledge of source types", *Journal of Librarianship and Information Science, 39(1)*, pp. 27-40.
Chu, S., and Law, N. (2007b), "Development of information search expertise: Postgraduates' knowledge of searching skills", *Portal: Libraries and the Academy, 7(3)*, pp. 295-316.
Chu, S., and Law, N. (2008), "The development of information search expertise of research students", *Journal of Librarianship and Information Science, 40(3)*, pp. 165-177.
Debowski, S. (2001), "Wrong way: go back! An exploration of novice search behaviours while conducting an information search", *The Electronic Library, 19(6)*, pp. 371-382.
Devine, J., and Egger-Sider, F. (2009), *Going beyond Google : the Invisible Web in learning and teaching*, New York, Neal-Schuman Publishers, Inc.
Duggan, G.B. and S.J. Payne. (2008), "Knowledge in the head and on the Web: Using topic expertise to aid search", *Proc. CHI 2008*, ACM, pp. 39-48.
Fidel, R. (1985), "Moves in online searching", *Online Review, 9(1)*, pp. 61–74.
Graham, L., and Metaxas, P. (2003), ""Of course it's true; I saw it on the internet!": critical thinking in the internet era", *Communications of the ACM, 46*, pp. 71-75.
Halttunen, K., & Järvelin, K. (2005), "Assessing learning outcomes in two information retrieval learning environments", *Information Processing Management*, 41(4), 949-972. doi: 10.1016/j.ipm.2004.02.004
Hembrooke, H. A., Granka, L. A., Gay, G. K., and Liddy, E. D. (2005), "The effects of expertise and feedback on search term selection and subsequent learning: Research Articles", *Journal of the American Society for Information Science and Technology, 56(8)*, pp. 861-871.
Hölscher, C., and Strube, G. (2000), "Web search behavior of Internet experts and newbies", *Computer Networks, 33(1-6)*, pp. 337-346.
Hsieh-yee, I. (1993), "Effects of search experience and subject knowledge on the search tactics of novice and experienced searchers", *Journal of the American Society for Information Science*, *44(3)*, pp. 161-174.
Kang, R., and Fu, W.T. (2010), "Exploratory information search by domain experts and novices", Paper presented at the *IUI '10 Proceedings of the 15th international conference on Intelligent user interfaces*.
Kuhlthau, C. (1993), "Implementing a Process Approach to Information Skills: A Study Identifying Indicators of Success in Library Media Programs", *School Library Media Quarterly, 22(1),* 11-18.
Kuhlthau, C. (2004), *Seeking meaning : a process approach to library and information services* (2nd ed.). Westport, Conn.: Libraries Unlimited. Retrieved from http://www.ala.org/aasl/aaslpubsandjournals/slmrb/editorschoiceb/infopower/slctkuhlthau1
Kuruppu, P.U. & Gruber, A. M. (2006), "Understanding the information needs of academic scholars in agricultural and biological sciences", *The Journal of Academic Librarianship,* 32(6), 609-623.
Larkin, J., McDermott, J., Simon, D. P., & Simon, H. A. (1980), "Expert and novice performance in solving physics problems", *Science*, 208(4450), 1335-1342.
Larkin, K., McDermott, J., Simon, D., and Simon, H. (1980), "Expert and novice performance in solving physics problems", *Science, 208*, pp. 1335-1342.
Laverty, C., Reed, B., and Lee, E. (2008), "The "I'm feeling lucky syndrome": Teacher-candidates' knowledge of Web searching strategies", *The Canadian Journal of Library and Information Practice and Research*, *3(1)*. Retrieved from http://journal.lib.uoguelph.ca/index.php/perj/article/view/329/892
Lazonder, A. W., Biemans, H.J.A., and Wopereis, I. G. J. H. (2000), „Differences between novice and experienced users in searching information on the World Wide Web", *Journal of the American Society for Information Science and Technology, 51(6)*, pp. 576-581.
Lewandowski, D. (2005), "Web searching, search engines and Information Retrieval", *Information Services & Use, 25*, pp. 137 - 147.
Morner, C. J. (1993), *An instrument for measuring library skills of education graduate students* (PhD dissertation, Boston College, Boston.

Prensky, M. (2001), "Digital natives, digital immigrants Part 1", *On the Horizon, 9*(5), 1-6. Retrieved from http://www.marcprensky.com/writing/Prensky%20-%20Digital%20Natives,%20Digital%20Immigrants%20-%20Part1.pdf

Segev, E. (2010), *Google and the digital divide : the bias of online knowledge*, Oxford, Chandos Publishing.

Talja, S., and Maula, H. (2003), "Reasons for the use and non-use of electronic journals and databases", *Journal of Documentation, 59(6)*, pp. 673-691.

Vygotsky, L. S., & Cole, M. (1978), *Mind in society: The development of higher psychological processes*, Harvard university press.

Washington-Hoagland, C. & Clougherty, L. (2002), "Identifying the resource and service needs of graduate and professional students: the university of Iowa user needs of graduate professional series", *Libraries & the Academy* 2(1): 125-143.

Table 1. Profile of the eight participants.

Participant	P1	P2	P3	P4	P5	P6	P7	P8
Gender (M/F)	M	M	M	F	F	F	F	F
Age range*	(2)	(1)	(3)	(2)	(1)	(2)	(3)	(1)
Year in graduate program	2	1	2	1	2	1	2	1
Current program (mode)	EdD (PT)	PhD (FT)	EdD (PT)	PhD (PT)	PhD (FT)	PhD (FT)	EdD (FT)	PhD (FT)

Note: EdD stands for Doctor of Education; PhD stands for Doctor of Philosophy
* (1) 21-30; (2) 31-40 and (3) 41-50 years old

Table 2. Definitions of nine information search tactics proposed by Hembrooke et al. (2005).

Tactics	Definition
Elaboration	Expanding and extending search statements using details and searchers' intrinsic ideas.
Example: (P1 - M1, p.1)[1]	[knowledge management and open-source software]: knowledge management AND wiki, knowledge management OR KM OR wiki, knowledge management OR KM OR knowledge sharing[2]
Redundancy	An overall index of the extent to which search terms are used repeatedly on successive queries. Redundancy can be further divided to backtracking, topic terms, and plural making/taking.
Example: (P4 – M3, p.3)	[social tagging and digital library]: social tagging AND digital library, tagging AND digital library, social tagging AND digital
Backtracking	The frequency with which a searcher reuses prior search terms over successive trials.
Example: (P2 – M4, p.4)	[Personal epistemology]: personal epistemology, epistemology
Topic Terms	The extent to which the given topic terms are used as main querying terms.
Example: (P2 – M3, p.3)	[achievement motivation of student]: motivation AND academic achievement AND students, motivation AND engagement AND academic achievement
Plural Making/Taking	Reflects instances when a user repeatedly incorporates similar nouns into their search attempt, with the slight modification of making the word plural or singular.
Example: (P1 - M2, p.2)	[Activity theory and COP]: activity theory, activity theories, community of practice, communities of practice
Broadening	The extent to which a user begins with a specific query and expands the scope of the search phrase over successive trials.
Example: (P5 – M1, p.1)	[Role of feedback in promoting students' learning]: effect of feedback AND English language learning, feedback AND second language acquisition
Refining	The extent to which a subject begins broadly and narrows the search with increasing specificity.
Example: (P5 – M5, p.4)	[ecological perspective on learning and teaching]: ecological perspective AND language learning, ecological perspective AND English writing
Kitchen Sink	The extent to which a searcher incorporates search terms related to the subject, but not specific to the query task
Example: (P2 - M5, p.5)	[Well-being in school]: well-being AND student OR school, well-being AND student OR school AND goal, well-being AND achievement goal
Poke-N-Hope	The extent to which a searcher retains the same basic structure throughout all search queries, changing only a single word within each trial
Example: (P4 – M2, p.2)	[digital preservation]: networking for digital preservation, nobles networking for digital preservation, barnes networking for digital preservation, barnes and nobles networking for digital preservation

[1] Reference to page 1 of the transcript of participant 1 (P1), meeting 1 (M1)

[2] [The words inside the brackets] means the search topic decided by the novice searchers; and the words following are the search terms developed by the participants or experts. Each search query is separated by comma.

Table 3. Students' perceived familiarity with various databases.

Databases	Average ratings at the 1st session (Pre-test) Mean (SD)	Average ratings at the 5th session (Post-test) Mean (SD)	Comparing the first and final surveys Z value	Asymp. Sig.
Journal Articles				
- EBSCOhost research databases	3.57 (0.79)	3.63 (0.52)	0.000	1.000
- ProQuest	3.57 (0.79)	3.50 (0.76)	0.000	1.000
- Web of Science	3.14 (0.90)	3.25 (0.46)	-0.577	0.564
- Scopus	3.00 (0.71)	2.71 (0.49)	-0.577	0.564
Books				
- HKU Libraries Dragon Catalogue	3.29 (0.49)	3.75 (0.46)	-1.732	0.083
- HKALL	3.00 (0.63)	3.38 (0.52)	-1.000	0.317
Theses				
- Local theses	2.67 (0.82)	3.00 (0.58)	-1.000	0.317
- US theses	2.14 (0.69)	2.87 (0.64)	-1.890	0.059
- UK & Ireland theses	1.57 (0.53)	2.29 (0.76)	-1.890	0.059
- International theses	1.57 (0.53)	2.57 (0.53)	-1.890	0.059
General				
- Google Scholar	3.29 (0.49)	3.50 (0.76)	-0.577	0.564
- WorldCat	1.60 (0.55)	2.43 (0.98)	-1.633	0.102
- Emerald	2.50 (1.22)	2.71 (0.49)	-0.707	0.480
Overall	**2.70 (0.12)**	**3.03 (0.29)**	**-2.366**	**0.018***

Notes: * indicates p<.05

0=Don't know, 1=Not Familiar, 2=A Little familiar, 3=Familiar, 4=Very familiar; all rating of 0 has been excluded from calculation

Table 4. Students' perceived familiarity with information search knowledge and skills.

Item	Average ratings at the 1st session (Pre-test) Mean (SD)	Average ratings at the 5th session (Post-test) Mean (SD)	Comparing the first and final surveys Z value	Asymp. Sig.
Computer and IT skills	3.43 (0.53)	3.57 (0.53)	-0.577	0.564
The subject coverage of the databases you use	3.00 (0.00)	3.29 (0.49)	-1.414	0.157
Field Search by:				
- Author	3.00 (0.58)	3.57 (0.53)	-1.414	0.157
- Title	3.14 (0.38)	3.57 (0.53)	-1.342	0.180
- Date/Year/Time period	3.00 (0.58)	3.29 (0.76)	-0.707	0.480
- Material Type	3.14 (0.38)	3.43 (0.53)	-1.000	0.317
- Subject search	3.14 (0.38)	3.43 (0.53)	-1.000	0.317
Keyword search	2.86 (0.38)	3.43 (0.53)	-2.000	0.046*
- Step 1: Choose concepts	2.86 (0.38)	3.14 (0.69)	-1.000	0.317
- Step 2: Choose search terms	3.00 (0.00)	3.17 (0.75)	-0.577	0.564
- Step 3: Decide on appropriate search commands	3.00 (0.00)	3.17 (0.75)	-0.577	0.564
- Step 4: Construct a search statement.	2.86 (0.38)	2.86 (0.69)	0.000	1.000
Boolean Operator:				
- AND	3.00 (0.00)	3.43 (0.79)	-1.342	0.180
- OR	3.00 (0.00)	3.14 (0.69)	-0.577	0.564
- NOT	2.86 (0.38)	2.86 (0.69)	0.000	1.000
The use of parentheses	2.86 (0.38)	3.00 (0.57)	-0.577	0.564
Wildcard	2.43 (0.98)	3.14 (0.69)	-1.518	0.129
Truncation	3.00 (0.00)	3.29 (0.95)	-0.816	0.414
Proximity	2.29 (0.76)	3.14 (0.90)	-1.667	0.096
Non-database search skills:				
- Identify & browse core journals	3.00 (0.57)	2.86 (0.38)	-0.577	0.564
- Identify publications from key university/research centres for your research	2.28 (1.11)	2.86 (0.38)	-1.414	0.157
- Identify key researchers in your research area	2.71 (0.49)	2.86 (0.69)	-1.000	0.317
Overall	**2.86 (0.20)**	**3.16 (0.52)**	**-1.352**	**0.176**

Notes: * indicates p<.05

0=Don't know, 1=Not Familiar, 2=A Little familiar, 3=Familiar, 4=Very familiar; all rating of 0 has been excluded from calculation

Table 5. Database usage frequency comparison between novice and expert searchers.

SOURCE TYPES	DATABASES USED	NOVICES	EXPERT	TOTAL USAGE
Journal Articles	EBSCOhost	9	6	15
	ProQuest	6	6	12
	Web of Science	5	6	11
	HKUL Catalogue (Dragon)	2	1	3
	Scopus	1	-	1
	ScienceDirect	1	-	1
	Wiley Online Library	1	-	1
	HKALL	1	-	1
	HKIED Library Catalogue	1	-	1
	Google Scholar	1	-	1
	Google	1	-	1
Books	HKUL Catalogue (Dragon)	9	3	12
	WorldCat	-	7	7
	HKALL	3	3	6
	Google Book	3	3	6
	Google Scholar	2	-	2
	PolyU Library Catalogue	1	-	1
	Google	1	-	1
Theses	ProQuest	5	8	13
	WorldCat	-	6	6
	Digital Dissertation Consortium	4	1	5
	Emerald	1	-	1
	HKU Theses Online	1	-	1
	Networked Digital Library of Theses & Dissertation	-	1	1
	Google Scholar	1	-	1
Web Resources	Google	5	6	11
	Google Scholar	8	2	10
	HKUL Catalogue (Dragon)	1	1	2

Figure 1. Comparison of satisfaction level between unaided and scaffold search Sessions.

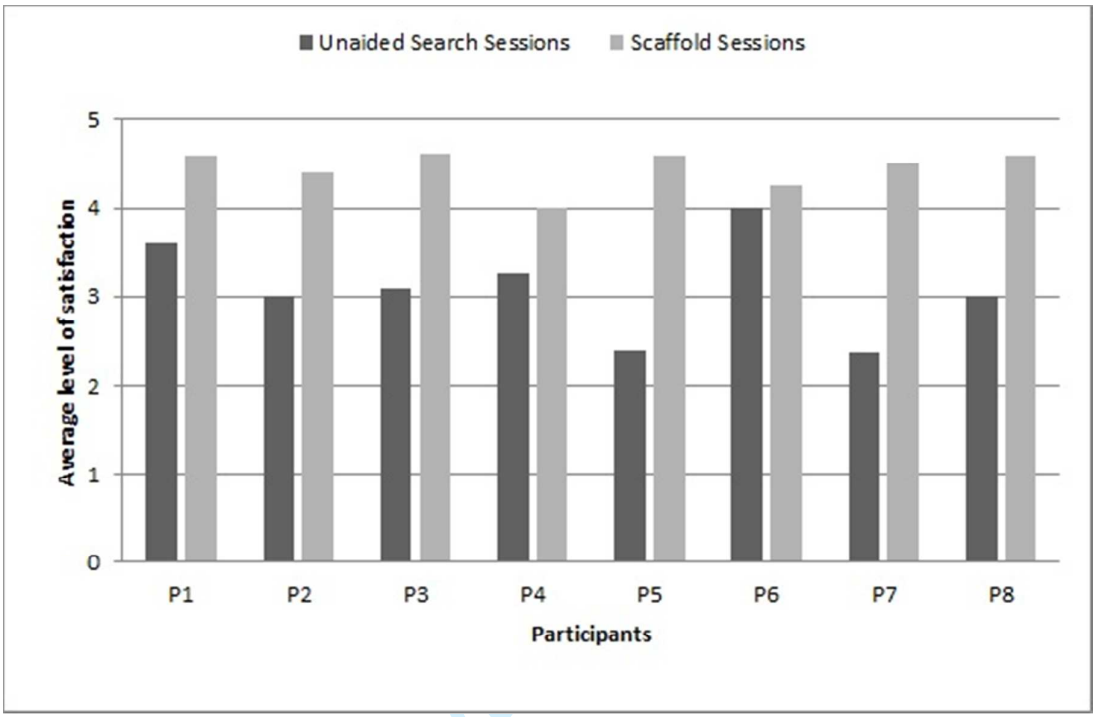

Figure 2a. IRS usage frequency breakdown on four source types by novice searchers.

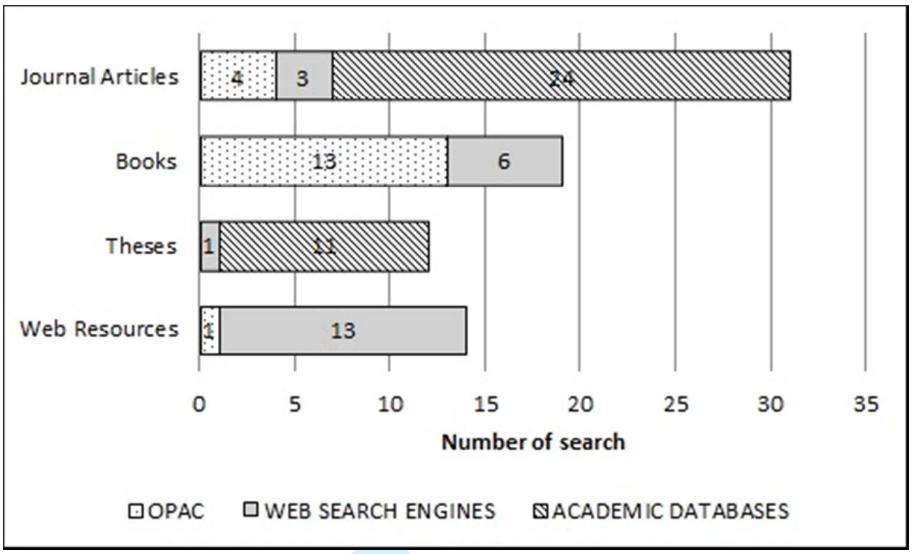

Figure 2b. IRS usage frequency breakdown on four source types by expert searcher.

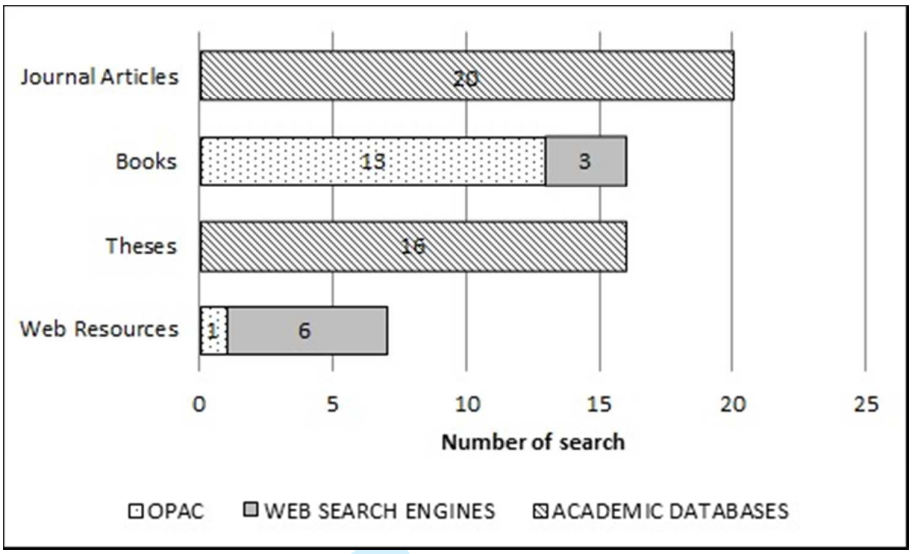

Figure 3. Query formulation tactics comparison between novice searchers and the expert searcher.