Scaffolding Performance in IR Instruction
Exploring Learning Experiences and Performance in Two Learning Environments

Kai Halttunen
University of Tampere, Finland

Kai Halttunen, Department of Information Studies, University of Tampere, FIN-33014 University of Tampere, Finland. E-mail:Kai.Halttunen@uta.fi

Abstract

In order to design information retrieval (IR) learning environments and instruction, it is important to explore available pedagogical solutions and their impact on learning experiences and performance in learning situations. The particular focus of this study is the demonstration how pedagogical solutions like scaffolding and anchored instruction can be implemented in IR instruction and what effects it has on learning experiences and performance compared to a more traditional learning environment and instruction. The study of these solutions is a novel approach in information science. The 65 participants of an introductory course on IR were selected for this study, and the analysis illustrates their learning experiences and performance. The method for empathy based stories (MEBS) as well as course feedback questionnaire and log files of search exercises were used to collect data. The results indicated that anchoring and scaffolding are promising strategies to make learning experiences meaningful and create learners' ownership to one's learning. Performance analysis indicates that scaffolding with an instructional tool, the IR Game, with performance feedback and scaffolding and anchoring by tutors enables students to construct more effective queries. Results of the study, in terms of the benefits of anchoring and scaffolding, are not categorical because the range of intervening variables and the difficulty of setting up a field experiment which tried to be naturalistic but at the same time tried to focus on a specific aspect.

1. Introduction

Web and CD-ROM mediated information systems and sources have made information retrieval (IR) a commonplace activity. These IR activities include selection of relevant information sources, construction of queries representing information needs and search requests, interaction with IR systems and evaluation of search results. Searching is done with computerized IR systems. Everyday users are afforded with the same opportunities and tools as information specialists by using emerging information systems such as the Web. IR know-how is needed in several task situations like education, business and everyday activities. The commonality of IR activities has risen interest in user behavior in information searching activities, but it has produced very little, if at all, research and development in the area of teaching and learning IR. Webber and Johnston [1] have argued that research into information seeking behavior does not seem to have had much influence on how information searching is taught. They hypothesise that users of information systems are taught in

(

much same sort of way as 20 years ago. The contribution of the research described in this paper is to provide new approaches and methods in to field of IR instruction. The present research introduces two pedagogical solutions, scaffolding and anchored instruction, and describes and evaluates their implementation in IR instruction along with participants' learning experiences and performance.

2. The field of IR instruction

IR instruction is routinely organized in different levels by schools, universities, libraries, online vendors, consultants etc. Besides the commonality of IR skills, they are a key area of expertise for information professionals. A wide variety of textbooks about the basics and principles of searching has been published [2-5]. The educational material covers four main areas focusing on presenting (1) the context of IR as a part of information seeking activities, (2) basic principles of IR systems, (3) general search strategies applicable in all ordinary retrieval settings, and (4) specific search strategies for particular retrieval settings and information sources. The main goal of instruction is to develop learners' practical capability to perform successfully any search task appearing in the professional work situation.

Research covering instruction of IR is disjointed, without solid background in either information studies nor educational research. Although different approaches to IR instruction have some common elements it is quite a fuzzy field of activities with different levels, approaches and stakeholders. The following paragraphs categorize these different approaches. Firstly, education of professionals in information and library science. This research has analyzed for example the amount of IR instruction in curriculum, presence and integration of IR instruction in different courses and differences between covered domain and instructional methods and materials. (See, e.g. [6, 7]).

Secondly, there is a large amount of literature on user education in libraries. Bibliographic instruction covers some aspects of IR know-how. Research has concentrated mainly on teaching methods and implementation of computer-assisted tutorials as well as coverage of courses. (See, e.g. [8-11]).

Thirdly, there are a few studies of instruction of IR in other domains than information studies. Examples of these are journalism [12] and education [13-17]. This rising trend of IR instruction especially in the educational field calls for solid research on IR instruction and learning environments.

Fourthly, research and development in the field of information skills covers partly the same elements as previous approaches, but concentrates more on research oriented tasks in educational settings. Information skills cover more widely the research and information seeking process than IR instruction. Approaches like the information search process [18], information problem-solving [19-20], study and information skills [21], library research process [22] are well known examples of different approaches to information skills instruction.

(

Fifthly, the concept of information literacy has gained increasing attention in recent years. Discussion about information literacy covers elements of information skills with some more emphasis on communicating, synthesis and creation of information. It also covers more economic, legal, and social issues surrounding the use of information. Discussion of information literacy has clearly two different starting points, namely a normative, top down approach which presents requirements for information literacy education and information literate person [23] and on the other hand a bottom up approach which tries to describe and evaluate different empirical conceptions of information literacy [24]. Webber & Johnston [1] and Bawden [25] provide exhaustive review of the topic.

3. Pedagogical solutions for IR instruction

Two applicable approaches to IR instruction, scaffolding and anchored instruction are described in following subsections.

3.1 Scaffolding

The term "scaffolding" in nowadays widely used in instructional settings. Scaffolding refers to different kinds of supports that learners receive in their interaction with teachers, tutors and different kinds of tools within a learning environment as they develop new skills, concepts or levels of understanding. The term scaffolding was originally put forward by Wood, Bruner and Ross [26] as a metaphor to describe the intervention of peer, adult or competent person in the learning of the other person. The notion of scaffolding relies heavily to Vygotsky's concept of "the zone of proximal development (ZPD)". The socio-cultural approach based on the work of Vygotsky has had a major influence on the development of scaffolded instruction and apprenticeship models of learning [27, 28]. Scaffolding should enable learners to perform activities they were unable to perform without this support. Along with support the other important element in scaffolding is fading. Fading represents gradual removal of support when learners can cope with the task independently. Winnips and McLoughlin [29] have distinguished between initial and ongoing support. Initial support is given in the beginning of the task, and faded so that the student can learn to self-regulate. Ongoing support is provided during the task completion and is based on student input. Scaffolding can be implemented by human tutors or it can be embedded in instructional software environments. Gutzdial [30] has made a distinction between macro and micro level scaffolding. The macro level is concerned with the stages or collections of activities which the student undertakes. The micro level identifies the individual activities which the students undertake. Examples of initial and ongoing support at macro and micro level, as well as software-based and face-to-face scaffolding, are given in Chapter 4.

3.2 Anchored instruction

In many cases the search tasks in IR instruction are well specified. This exactness and predefined linguistic expressions can be seen as an obstacle for learning because they are not related to any meaningful real-world like situation. The more complete the specification of values for each instructional component, the less inclined teachers

(

may be to map onto the unique features of particular students and communities. One solution to this problem could be anchored instruction, an instructional approach developed by the Cognition and Technology Group at Vanderbilt [31-33].They have tried to build an instructional design approach that provides some specific tools for instruction yet also leaves room for a great deal of flexibility. The basis of the approach is to build semantically rich "anchors" that illustrate important problem solving situations. These anchors create a "macrocontext" that provides a common ground for experts as well as teachers and students from diverse backgrounds, to communicate in ways that build collective understanding. Macrocontexts are semantically rich environments that can be used to integrate concepts across the curriculum and in which meaningful, authentic problems can be solved.

A number of theorists emphasize the importance of helping students to engage in generative rather than passive learning activities. Findings from a number of studies suggest that knowledge that is not acquired and used generatively tends to become "inert knowledge" – knowledge that is not used spontaneously even though it is relevant. Anchored instruction and apprenticeship learning stress the importance of anchoring or situating instruction to meaningful problem solving context that allows one to simulate some of the advantages of apprenticeship learning. A major goal of this approach is to create shared environments that permit sustained exploration by students and teachers and enable them to understand the kinds of problems and opportunities which experts in various areas encounter and the knowledge that these experts use as tools. [34].

Honebein, Duffy and Fishman [35] specify authentic activities and context as the elements of constructive learning environments. Authenticity of learning activity refers to the activity of the learner in the learning environment relative to the environment in which learning will be used. Authentic learning activity requires that learners have ownership of their learning and performance. The instruction must support them in assuming responsibility for establishing and monitoring their goals and strategies. The second critical feature in the design of an authentic task is a holistic view of the task complex. Authentic activities exist in both global and local entities. The global entities defines the entire task, while the local entities refers to sub-tasks. The global task environment, including the purpose for undertaking the global task, gives meaning to each of the local tasks. It is therefore essential that the global activity is very closely tied to the tasks and experiences the learner will face in the transfer environment. Examples of our implementation of anchored instruction are presented in the following chapter.

4. Implementation of scaffolding and anchored instruction

The following subsections describe the implementation of scaffolding and anchored instruction in an introductory course of information retrieval.

4.1 Introduction to information retrieval

The course called "Introduction to Information Retrieval" (6 ECTS credits) in the department of information studies at University of Tampere consists of four elements. First, lectures are given on basic concepts of information storage and retrieval.

(

Themes like production of databases, matching, metadata, query formulation, and evaluation are covered. Second, weekly web-exercises concentrate on putting the themes covered in lectures in to practice. Every participant is working on these exercises making use of web-based tools and resources. Exercises are reported on web-forms, which are posted to the teacher. Third, tutored exercises in the classroom cover different aspects of information retrieval systems and their efficient use. In these sessions students work in pairs in groups of 8-10 people. Different kind of search services are used (an OPAC, a union catalogue, article reference databases, full text databases, Internet search engines and directories). Fourth, a course feedback web-form is filled out at the end of the course. Evaluation covers three main areas: 1) course design and teaching methods, 2) the learner's self evaluation and role in the course and 3) the teachers role in the course. The course is intended for first year undergraduate students. The course provides an overview of information storage and retrieval as practice and research area. After the course students are able to construct and evaluate queries in various IR environments (references, full-text, images, best-match, exact-match) and they have basic knowledge of production and structures of different database types. They have also basic knowledge of information organization (e.g. ISBD, AACR, Dublin Core, classification, abstracting, and indexing). About 120 students attended this course, which was run by a lecturer and two tutors taking care of part of the tutored sessions. Assessment of learning outcomes was based on weekly web-exercises and tutored exercises. The dichotomy pass/fail was used in course evaluation as such without grading to indicate that student had fulfilled/not fulfilled the learning objectives.

In order to evaluate the effect of scaffolding and anchored instruction a quasi-experiment was carried out. Tutored exercises were carried out in two different ways. In the traditional learning environment different operational search systems (an OPAC, an union catalogue, article reference databases and full-text databases) were used to demonstrate basic functions or IR systems. In the experimental setting instruction full-text newspaper articles from local newspaper along with press image database were used. These sources were used through the IR Game, a system which offers feedback to the searcher on the effectiveness of queries based on recall-base. Same two article reference databases were also used in experimental setting as in traditional instruction. Various ideas of scaffolding and anchored instruction were applied in the experimental learning environment. Detailed descriptions of these pedagogical solutions are provided in the next chapter. The differences between traditional and experimental learning environment are summarized in Table 1.

	Traditional instruction	**Experimental instruction**
Instructional design	Searching databases	Searching in context. Anchored instruction.
	Unintentional scaffolding in the classroom.	Intentional scaffolding in the classroom and in the IR Game.
Performance feedback	Provided by tutor, no performance feedback in operational systems used.	Performance feedback in the IR Game and by tutors.
Systems used	Six operational systems (85% of time) and the IR Game (15%).	The IR Game (85%) and two operational systems (15%)

(

Timing	Six one hour sessions.	Three two hour sessions.

Table 1. Summarization of differences in traditional and experimental learning environment.

Shadish, Cook and Campbell [36] describe quasi-experiments to have participants self-selection to treatment groups, manipulable cause and enumeration of alternative causes. In the present study, the participants of the quasi-experiment were self selected into the groups, but their learning styles, prior conceptions of IR and educational/professional background were analyzed. Learning styles were analyzed with Kolb's "Learning Style Inventory" [37, 38] and analysis of prior conceptions was based on student essays and questionnaire. The methods for and results from, the study of these aspects are reported by Halttunen [39]. Based on this analysis the groups are similar in this respect. Manipulable cause in this study is difference between traditional and experimental learning environment as described earlier.

4.2 Scaffolding in the IR Game and in the classroom

Different kinds of scaffolds were built both in the classroom situation and the instructional tool in the IR Game. The IR Game is based on the idea that test collections used in laboratory-based IR experiments could be used in instruction as a rapid query analysis tool. The IR Game offers features like query performance feedback, visualization of search results, comparison of queries, document browsing and help facilities. The IR Game (also called The Query Performance Analyzer) is described in detail by Halttunen and Sormunen; Sormunen, Halttunen and Keskustalo; and Sormunen et al. [40-42]. The following table describes the instructional scaffolds implemented in the IR Game.

(

Scaffold	Implementation in the IR Game
Giving away parts of solution	Query performance feedback, Relevance bar
Providing cues	Give a hint
Providing cues	Relevant documents
Providing examples	Hall of fame
Providing comparison	Others queries, p/r-curves
Giving away parts of solutions, templates	Search topics

Table 2. Software-based scaffolding implemented in the IR Game

The basic idea and functionality of the IR Game is that the query performance feedback scaffolds the learner by providing information on query performance. The learner receives feedback on query construction, which enables her/him to evaluate different search strategies and tactics. Learners can concentrate on the analysis of effective query formulation not spending lots of time in analyzing results. The "Give a hint" -function provides the learner with one non-retrieved relevant document. This document serves as a cue to selection of appropriate search keys or document structures usable while searching. The possibility of identifying the retrieved relevant documents serves also as a cue in same respect. The hall of fame provides examples of best queries on the current topic over time. The possibility to see one's own performance compared to query performance of other users provides area of comparison and feedback. Search topics can serve as a scaffold, providing examples of search keys and informing about applicable restrictions.

Software-based scaffolds in the IR Game are based on the basic idea of "knowing the right documents" on each topic. In a way this approach pays attention to the end product of process, to the effective query formulation with good precision and recall, but not on the interactive learning process support as such. Face-to-face scaffolding is needed in this situation. Examples of scaffolding in the classroom are presented in Table 3.

Scaffold	Implementation in Classroom
Providing examples	Modeling a search process
Providing hints	Query formulation
Giving away parts of solution	Search terms
Cueing/hinting	Operators, syntax
Coaching comments	Why this happened?
Asking questions	How that affects? What problems that may cause?
Providing a timeline	Search process timeline

Table 3. Scaffolding implemented in classroom.

(

The teacher can model the search process by providing examples not only focusing on end products (efficient query formulations), but also on a process. The teacher provides hints, in initial or ongoing basis, on query formulation. Provision of examples of possible search keys serves as an scaffold in certain situations. Coaching comments are intended for motivation, providing feedback and advice on performance, and provoking reflection. Different kinds of questions can be set to enhance reflection, for example pointing out weaknesses and asking for motivation. Provision of a timeline with fixed timing and goals provides support for goal direction and reflection. The gradual removal of scaffolding, i.e. fading, was based on student performance in exercises. When students were able to construct queries, scaffolds like examples and hints were removed. Coaching comments and questions were used through all exercises to improve reflection and articulation. Relevance feedback provided with the IR Game was removed in one of the search-tasks, when students selected their own viewpoints to the topic.

4.3 Anchored instruction

In the experimental group in the introductory course of information retrieval we built a macro-context for IR activities. We used local newspaper full-text databases, image database of national press agency and national database for journal and newspaper article references. Based on these tools we created a context of journalistic practice. In other words IR activities were situated in simulated work-task situation, were search-tasks were based on the idea of journalist searching local text and image databases in order to find information on certain topics for a forthcoming article. In addition some topics required the reporter to search national databases on certain topics. The following two descriptions of search tasks give an example of different kind of approaches.

First, one of the search tasks covered a news event were two police men were killed in Helsinki. This event consisted of different phases like crime, investigation, imprisonment, national grief, influence of the event to police work and legislation, and so forth. Based on this news event we created a context for learning to narrow down the search by sub-theme, date, section and different kind of linguistic expressions in the different phases of the event. We also used an image database to search for images related to this event. Students were able to explore the topic generally first and after that decided the sub-theme to work with. Efficiency of different search strategies and tactics were presented to the other students.

Second, the learners approached also the national reference database through journalistic practice. First they searched for discussions about the opening hours of shops in local newspaper archive and after that broadened their search to the national reference database. We used this kind of search task in order to demonstrate differences between different kinds of databases and their structures as well as idea of subject representation (descriptors, classification) in the reference database. It was also possible to compare the situation of searching for the same articles either in full-text or based on controlled vocabulary in a reference database.

(

5. Research questions

The research question of the present study is presented as "What is the effect of experimental learning environment to learning experiences and performance in IR instruction?". An experimental learning environment consists of different tools and pedagogical solutions than traditional learning environment as was described in preceding chapter. Both environments share the same elements such as lectures and web exercises, but variations of the tutored exercises form a differentiated environment for learning and instruction.

The main research question can be divided in to following sub questions:
1) What kind of instructional design produces positive or negative learning experiences?
2) Are there differences in experiences in different learning environments?
3) Are there differences in learner performance in using IR Game in different learning environments?

6. Data and methods

Data for the current research was gathered in larger research effort on designing and evaluating modules of an IR learning environment. In the present study three different datasets are analyzed, namely empathy-based stories describing student learning experiences, course feedback questionnaires and finally transaction log files gathered in tutored exercises. The purpose of the data triangulation is to describe both experienced and observed effects of the different learning environments. In the same field experiment data was also gathered and analyzed on conceptions of IR know-how and learning styles and learning outcomes [39, 48].

The first body of data consists of stories written by students attending this course. In these stories they describe good and bad learning experiences. The total number of stories is 65 from which 30 describe positive learning experiences and 35 negative ones. The method of empathy-based stories (MEBS) was applied [43, 44]. The methodology is a variation of non-active role-playing method [45, 46]. This method involves writing of short essays according to instructions given by the researcher. The respondent is given some orientation, which is called the script. This script should be used in conjunction with the respondents' imagination in the writing of the story. The writer of the story either continues the situation detailed in the script or describes what must or may have taken place prior to that situation. Variation is crucial to the use of this method, and there are at least two different versions of the same script which vary with regard to a certain key issue. Variation distinguishes the method of empathy-based stories from many other methods of acquiring data. According to Eskola [44] MEBS is a powerful method to gain rapid feedback from teaching programs or planning future curricula. Since it can readily be carried out in most cases, it is very suitable in exploratory studies. MEBS can also be very useful in supplementing traditional survey methods. The scripts with example stories used in this study are presented in Appendix 1.

(

The second body of data consist of the course feedback questionnaires which were filled out in web-form as an prerequisite for course completion. The questionnaire covers three areas of interest: 1) course design and teaching methods, 2) the learner's self evaluation and role in the course and 3) the teachers role in the course. There were 21 questions out of which 16 were open questions and 5 were multiple choice questions.

The third dataset consists of transaction logs of the IR Game used in the tutored exercises. The total number of queries was 3763 from which the sample was generated. The sample consists of 606 queries from the experimental groups and 620 from the traditional groups. These files cover the same four exercises in both groups, representing all queries of the participants in the experimental group and a random sample of the same number of participants in the traditional group. From each exercise the following aspects were analyzed: a) the number of queries, b) the number of search keys, c) the average number of search keys per query, d) the qualitative evaluation of usage of operators, truncation and field searching, and finally e) the overall efficiency of best queries based on p/r-measures provided in the IR Game. Summary of research questions, data and analytical methods are presented in Table 4.

(

Research question	Data	Analytical method
Good or poor learning experiences.	Method for Empathy Based Stories (MEBS). Course feedback questionnaire.	Theme coding and categorization based on the data. Case-ordered predictor matrix.
Performance in using IR Game.	Transaction logs of search sessions.	Quantitative analysis of queries and search keys. Qualitative assessment of query construction. Efficiency of best queries.

Table 4. Summary of research questions, data and analytical methods.

7. Results

7.1. Analysis of learning experiences

Empathy-based stories were written based on two variations, scripts, describing good and bad learning experiences. Stories were analyzed by theme coding, categorization, and with the aid of matrixes [44, 47]. The type of matrix that was used in this study was case-ordered predictor matrix. Positive and negative learning experiences were contrasted with the traditional and the experimental learning environment where students had their tutored exercises.

Theme coding and categorization produced seven major themes of learning experiences presented in the data. They are study orientation, domain of study, instructional design, teachers, other students, computer skills and factors of everyday life. The theme instructional design has sub-topics lectures, tutored exercises, web-exercises, feedback, examination, and abstraction level. An overview of these themes with the percentage of stories in which they were mentioned is presented in Table 5. An example of theme coding and categorization is presented in Appendix 1.

(

Learning experiences		Positive % (n=30)		Negative % (n=35)	
MEBS themes	Code	trad.	expr.	trad.	expr.
Study orientation	SO	20	23	17	14
Domain of study	DS	20	20	20	20
Instructional design	ID	50	50	20	14
a) Lectures	ID a	50	50	40	46
b) Tutored exercises	ID b	37	50	31	27
c) Web exercises	ID c	27	36	29	17
d) Feedback	ID d	10	27	6	0
e) Examination	ID e	6	3	5	5
f) Abstraction level	ID f	9	4	6	0
Teachers	T	25	25	3	0
Other students	OS	5	8	21	10
Computer skills	CS	4	4	2	1
Everyday life	EL	0	0	34	0

Table 5. Themes from empathy-based stories.

Study orientations were clearly different in their nature. In positive stories learning was presented as the student's own choice and they interpreted topic and methods of the course as important and transferable. Studying in this respect had meaning, while in the negative stories studying was presented as accomplishing a task which had no meaning for actors themselves. There did not seem to be any possibility to transfer learned know-how to any real context. The same kind of basic approach can also be found in the descriptions of the domain or the topic at hand. In positive stories the domain (introduction to information retrieval) was constructed in relation with prior knowledge and usefulness in different study and work settings in the future. In negative stories the domain of study had no connection to the prior knowledge of students, or if it had domain information was presented in such manner that the learner was unable to make those connections. In these stories learners saw the domain quite decontextualized and wanted instruction to solve their current information problems. There was no difference in traditional and experimental learning environments in these themes.

Instructional design as a whole and the different pedagogical solutions were described nearly in every positive story. Instructional design with a process approach, clear structure, and flexible interplay with theory and practice were described as elements of good learning experiences. In negative experiences the large number of details did not construct any meaningful structure. Interplay with theory and practice, with the ability to test and apply knowledge in different situations did not meet the needs of their current information problems in which students seek for an answer. Students in the experimental group clearly pay more attention to the process approach and interplay with theory and practice.

If we consider different elements and pedagogical solutions in the instructional design, the following themes can be found. The same number of good and poor elements of lectures were presented. Main elements, presentation style, lecture

(

materials and atmosphere in the lectures were described. In tutored exercises interaction and feedback between learners and tutors formed a good learning experience. Poor timing and insufficient use of time with insufficient computer facilities resulted in contrary views of learning experiences. The weekly web-exercises were also mentioned both in good and poor learning experiences, but the themes mentioned were contradictory in interesting way. While active working with exercises to applying themes in the lectures in to practice formed a good learning experience, the same activity with the performance approach resulted a poor experience. In the same manner uncertainty of one's own performance carried signs of poor learning experiences, while adequate feedback with task partition, for example possibility to test ideas of lectures in weekly web exercises, resulted good ones. To some extent learning experiences were more positive in experimental group.

Different kinds of feedback both in lectures and exercises form a base for good learning experiences. Instructional design in the current learning environment offers feedback in different situations: feedback on weekly web-exercises, tutored exercises and a feedback mechanism in IR Game, which was exploited especially in the scaffolded experimental group. None of the participants in the scaffolded group mentioned any forms of feedback as negative factor.

Evaluation of learning outcomes was based on exercises and active participation, not on formal examination. This non-examination was described in good learning experiences by non-stressful opportunity to cover all aspects in the course by practicing. Students saw that examination concentrates their attention only on certain topics. On the other hand some students thought that there should be an examination in order to give you motivation. These were the same students, which presented studying as accomplishing a task, which had no meaning for actors themselves.

Social interaction with teachers, tutors and other students formed an interesting contrast in different learning experiences. The impact of teachers and tutors in positive learning experiences was notable. The performance of teachers falls in to three categories. Knowledge and approach to the domain, the understanding and commitment to learning process and social interaction. This positive experience was contrasted with stereotype of stubborn gray haired university teacher. Other students were treated both as obstacles or promoters of learning. In negative learning experiences, fellow students were disturbing the lectures and in the exercises some students were free riders in collaborative learning tasks. In the positive setting collaborative tasks enhance learning. There was also a feeling of support and absence of competition.

Finally, two background themes that affected the formation of learning experiences. Computer skills were mentioned both as an obstacle and a promoter of learning depending on learners' skills. Different happenings of everyday life were present in negative learning experiences. Tiring student life with a tight weekly timetable affected learning experiences as well as troubles in other sectors of life. Interestingly none of these activities seem to have any influence on positive learning experiences. This can be due to tendency to keep away from the reflection of learning experiences by externalizing the causes.

(

7.2. Analysis of course feedback

Web-based course feedback questionnaire with mainly open questions was filled out in the end of semester. Answers were categorized with the aid of matrixes in the same manner as MEBS-stories earlier. There were no remarkable differences in feedback between traditional and experimental groups in overall, but there are clear differences in some aspects. Analysis of these differences and their connection to MEBS data and instructional design is presented next.

The overall feedback was positive in both groups. Versatile instructional design was appreciated as well as domain of the study. Negative remarks of the course differentiated clearly between groups. Participants in experimental groups did not find feedback of tutored and web exercises presented in lectures meaningful as much as students in traditional groups. Participants in experimental groups seem to have enough feedback and scaffolding in tutored exercises. They approached the lectures rather as information source than interactive element of learning environment. The students in traditional group approached lectures in the opposite manner and appreciated given feedback.

Participants estimated the time they spent on weekly web exercises in course feedback. Almost 30 % of students in the traditional group spent more time in these activities than students in experimental group. The average time spent on these web activities was 40 minutes in experimental group and 52 minutes in traditional group in a week.

The major difference in course feedback was in evaluation of tutored exercises, as could be expected. The student in experimental group appreciated and find useful the feedback they received in experimental learning environment while in traditional environment students missed deeper approach to tasks completed. They complained that "the clue" of exercises was sometimes missed and they concentrated too much on different functions of multiple IR systems used, although at the same time they liked the possibility to get to know several systems.

Participants also graded their own performance in the course. The assessment was made in the scale from 1 to 3 (from poor to excellent). There was only a slight difference in participants' self-assessment in this phase. Students in the experimental group average grade in self-assessment was 2.8, while in the traditional group it was 2.4.

7.3. Analysis of transaction logs

Transaction logs in the IR Game were gathered both from the traditional and the experimental learning environment. Logs were analyzed manually with the aid of sorting functions of a standard spreadsheet. Log files consist of identifiers for 1) exercise, 2) topic, 3) date and time, 4) user id, 5) recall, 6) precision and 7) query.

(

1	2	3	4	5	6	7
2	am004	19.10.20 00 15:34	k21	0.75	0.4285	(suomen tietotoimisto# or stt#) and (räsä# or hiiht#) and (dopin# or anabol#)
2	am004	19.10.20 00 15:35	k21	0.75	0.5625	(suomen tietotoimisto# or stt# or hiiht#) and räsä# and (dopin# or anabol#)
2	am004	19.10.20 00 15:36	k21	0.66	0.6666	(suomen tietotoimisto# or stt#) and räsä# and (dopin# or anabol#) and hiiht#

Figure 1. Example of log file in IR Game, # are truncation marks.

Four common exercises from the traditional and the experimental group were selected for analysis. A database containing newspaper articles was used through a Boolean search engine available in the IR Game. Each exercise was assessed with four criteria. First, logical construction of facets with correct use of operators. Second, correct use of truncation and masking. Third, making use of field restrictions. Fourth, the overall efficiency of best query of each student within the exercise measured with the precision/recall -levels provided by the IR Game. In the construction of facets and use of operators, truncation and field searching the following measures were used: 1) major difficulties, 2) minor difficulties and 3) good performance. Example of analysis and application of evaluation measurement is presented in Table 6.

(

Recall	Prec-sion	Query	Construction of facets, selection of search keys and operators	Correct use of truncation and masking	Comments
0,09	1,00	Jönköping AND 1916 AND laiva	1	1	Poor selection of search keys, straight form task, no truncation used.
0,00	0,00	Jönköping AND sampan#	2	2	Appropriate search keys, but typo in sampan#, truncation used only in other key.
0,45	1,00	Jönköping# AND viini#	2	3	Appropriate search keys, change in keys, truncation used in both keys.
0,27	1,00	Jönköping# AND alkohol#	2	3	Change in search keys.
0,82	1,00	Jönköping# AND (alkohol# OR viin# OR samppan#)	3	3	Logical combination of search keys, truncation used properly
0,32	0,80		2,0	2,4	summary (average)
Comment: Student proceeded in exercise well, she/he learned to select appropriate search keys and combined them in good manner in the end. Use of truncation developed into good level after two queries.					
Measures 1=major difficulties, 2=minor difficulties, 3=good performance were used in evaluation of a) construction of facets, selection of search keys and operators, b) correct use of truncation and masking and c) use of field restrictions (not applied in this example).					

Table 6. Example of analysis and evaluation of search exercise log-file.

In the case of overall effectiveness, the average precision and recall was calculated and placed in an matrix. These matrixes are presented in Figures 2 and 3. The summary of log-file analysis is presented in Table 7.

Exercise	Avg queries per exercise		Avg # of search keys per query		Evaluation of facets		Evaluation of truncation		Evaluation of field searching		Avg recall		Avg precision	
	exp.	trad.	exp.	trad.	exp.	trad.	exp.	trad.	exp.	trad.	exp.	trad.	exp.	trad.
002	11,3	9,3	2,5	6,2	2,4	2,3	2,3	2,5	-	-	0,55	0,58	0,75	0,67
003	12,7	15,4	4,9	7,3	2,5	2,9	2,7	2,9	-	-	0,55	0,46	0,39	0,33
004	8,5	7,7	7,2	7,7	2,4	2,4	2,7	2,3	-	-	0,57	0,54	0,63	0,63
011	5,1	6,3	6,8	5,2	2,9	2,4	2,4	2,3	2,9	2,6	0,71	0,59	0,75	0,77

Table 7. Average number of queries, query terms, evaluation of query construction and recall and precision in experimental and traditional groups in four exercises.

Average number of queries per exercise diminished while instruction proceeded. This is due to construction of lengthier queries with more search keys. The exercise 3 on the Y2K-problem was an exception in this trend, probably because of the structure of the topic which called for the use of proximity to make use of textual structure of expressions. This need caused a wide range of different modifications of queries, because it is very difficult to apply several possibilities in a single query. There is no difference between groups in the number of queries. The average number of search keys per query increased in the course of time in the experimental group from 2.5 to 6.8 while in the traditional group there is no clear trend in time variation being from 5.2 to 7.7. Students in the experimental group were starting their learning sessions in

(

this phase while traditional groups had already come across several systems and bibliographic databases. Experience in constructing queries affected the query construction and number of search keys. Students in the traditional group applied greater number of different gaming strategies to reach good precision/recall -levels. Gaming strategies are artificial strategies to improve precision, such as the use of the not-operator to exclude irrelevant search terms and picking up very rare, meaningless terms from documents. These strategies are discussed more detail by Halttunen and Sormunen [40]. Although the traditional group had worked with bibliographic databases earlier their adaptation to full-text environment was good.

Use of operators and construction of facets developed steadily in the experimental group. In the traditional group the situation remained quite similar throughout, with the exception of managing exercise 3 on Y2K much better.

Use of truncation developed in experimental group into good level after first exercises and remained steady at that level. The traditional group used truncation with varied success, being little below the level of the experimental group in exercises 2 and 3. This can be due to adoption to different kind of search settings. The traditional group having been worked with bibliographical databases in earlier exercises and the experimental group practicing their skills in full-text newspaper database environment. The experimental group learned to take into account the variation in natural language along time, which is indicated by better performance in exercises 4 and 11.

Field searching was especially needed in some search tasks. Ideas of restricting searches to certain field in database records were familiar to the traditional group because of their prior experience, although the idea of field structure in newspaper full-text database is different. The experimental group applied more coherent and strict ways to limit their searches to certain fields like section, date and heading and they succeeded better in applying this search facility. They not only adopted the technical and structural idea of field restriction, but also applied different semantics, understanding the structure of newspaper articles. Anchoring instruction to simulated work-tasks as well as scaffolding the understanding of structure of documents enhanced learning.

The overall effectiveness of queries in the exercises was slightly better in the experimental group. They received better or similar effectiveness in three of four cases. The experimental group used fewer unique search keys to reach better results. They exploited more ideas of newspaper context, discourse and structure of documents in their query formulation whereas the traditional group used more gaming strategies, such as the use of the not-operator to exclude irrelevant search terms and picked up very rare, meaningless terms from documents to improve precision.

(

p/r	00-20	21-40	41-60	61-80	81-100	total
81-100		3	2	9	3	23
61-80		5	6	11		good
41-60	1	2	5	8		38
21-40		3	2	2		average
00-20			1		1	3 poor
p/r	00-20	21-40	41-60	61-80	81-100	total

Figure 2. Distribution of average precision and recall in experimental group in four exercises (n=64).

p/r	00-20	21-40	41-60	61-80	81-100	total
81-100	1	2	4	6	1	12
61-80	1	4	7	4	1	good
41-60		3	6	4	3	43
21-40		2	4	2	2	average
00-20	3	2	1	1		9 poor
p/r	00-20	21-40	41-60	61-80	81-100	total

Figure 3. Distribution of average precision and recall in traditional group in four exercises (n=64).

If we compare the traditional and the experimental group through the distribution of average precision and recall of individual exercises we can see that students in experimental group reach good performance in 23, average performance in 38, and poor performance in 3 cases of total 64 cases. In the traditional group the distribution is 12, 43 and 9 respectively. The difference between the groups was statistically significant (χ^2, p=0.0339).

8. Discussion and conclusions

The contribution of this article is the demonstration how scaffolding and anchored instruction can be implemented in an IR learning environment and the analysis of its effect on learning experiences and performance within an instructional episode. Scaffolding was provided by educational software tool, the IR Game, and by tutors in the classroom setting. Anchored instruction was based on the idea of IR activities taking place in a journalistic work-task situation. The analysis was based on the stories describing positive and negative learning experiences and the course feedback questionnaire. Performance in the IR Game was observed through log-files of search sessions. The present study is part of a research project on design and evaluation of IR learning environment. The project consists of four studies, namely: design and pilot evaluation of IR Game [40-42]; analysis of students' conceptions of IR know-

(

how and their implications for the design of learning environments [39]; evaluation of learning experiences and performance (present article); and evaluation of learning outcomes [48].

The analysis of stories of positive and negative learning experiences revealed seven major categories of description which affect learning experiences: study orientation, conceptions of domain, instructional design, teachers, students, computer skills and factors of everyday life. Meaningfulness of learning tasks and their relation to prior know-how created ownership of one's own learning, which was an element of positive learning experiences. Negative learning experiences contained descriptions of de-contextualized motives and learning tasks. This was due to the fact, that students with negative learning experiences could not relate the domain and tasks to their prior knowledge.

Instructional design with process approach, clear structure and interplay with theory and practice were elements of good learning experiences, while in negative experiences the large number of details did not construct any meaningful structure. Some students tried to construct this lacking structure of their current information problems, which ended in an disappointing experiences. Feedback of performance in learning environment was a central element of learning experiences. Study orientation and experimental instructional design were elements that influenced ways to experience different forms of feedback.

There were some differences in learning experiences and performance between participants in traditional and experimental learning environments. Participants in the experimental learning environment paid more attention to the process approach and interplay with theory and practice as elements of positive learning experience. Participants in the experimental group received clearly enough feedback in tutored exercises, because they did not find other forms of feedback necessary. This was indicated both in course feedback and MEBS stories.

Student self-analysis on time investment and learning outcomes indicated slight positive effect of scaffolding and anchored instruction. Students in the experimental group spent little less time in weekly exercises and evaluated their learning outcomes better than students in the traditional group.

Performance in the IR Game indicated that there was no difference between groups in the number of queries, but number of search keys increased in experimental group along the time. Evaluation of query construction, truncation and field searching revealed that students in the traditional group used more gaming strategies, i.e. artificial strategies to improve precision, than experimental groups. The overall effectiveness of queries in the exercises was slightly better in the experimental group. The analysis of distribution of effectiveness measures revealed that students in the experimental group reached good performance in 23 cases, while same performance was reached in 12 cases in the traditional group. The difference between the groups were statistically significant (χ^2, p=0.0339).

Anchoring and scaffolding are promising strategies to make learning experiences meaningful and create learners' ownership to one's learning in contrast with performance orientation without reflection and analysis one's own activities in search

(

exercises. Anchoring search exercises to real-word situations like journalistic practice creates possibilities to create a context were discussion of, for example, linguistic expressions, data structures and concepts of relevance, precision and recall may be studied in connection to prior knowledge.

There are some limitations of the current study, which ought to be discussed. The field experiment in tutored exercises is only a small part of whole learning environment and its effect to learning experiences is not as obvious as it could have be in more extreme situations where the elements of learning environments are totally different. In the present case also the traditional learning environment had of course elements of scaffolding, but they where not used as intentionally as in the experimental environment where it was a central element both in the tool used and in the classroom setting. Anchoring was also implemented in a way, which was possible without substantial technical and economic investment on, for example, video-based interactive tutorials of journalistic practice. We used classroom discussions, modeling and background stories to anchor the IR activities in to a context. These designs can also be seen as strength in current project. Salomon [49] among others has criticized "a horse race approach" of educational research where comparative situations are created between extreme situations. We operated in an operational, authentic environment in an exploratory way to identify effects of scaffolding and anchored instruction to learning experiences and performance in two differentiated learning environments. On the other hand, it is very difficult or even impossible to control all possible factors in an authentic field experiment. In the present study this can be clearly seen, because students' learning experiences were naturally effected by several other factors than those which were implemented in the field experiment. Results of the study, in terms of the benefits of anchoring and scaffolding, are not categorical because the range of intervening variables and the difficulty of setting up a field experiment which tried to be naturalistic but at the same time tried to focus on a specific aspect.

The present study has covered students' learning experiences and performance in naturalistic learning environments, where anchoring and scaffolding were used as instructional strategies. Further research concentrates on learning outcomes as well as their connection to prior conceptions of domain and effect of learning styles in these environments [48]. Learning outcomes are evaluated both in conceptual level and trough performance assessment of IR skills.

Appendix 1

Scripts of MEBS used in the study and example stories with theme coding.

Positive script: Student takes part in the introductory course of information retrieval in information studies. Course consists of lectures, web exercises and tutored exercises. After the course she thinks that she has learned a lot of IR. Use your imagination and write a short - realistic - story or description of what enhanced learning in this course.

(

Example story with theme coding: The course was planned to be interesting[ID] from the very beginning. There was not only lectures, [IDa] but you could practice and test [ID] your learning in web exercises [IDc]. Tutored exercises [IDb] were relaxed and fun occasion where you learned almost unnoticed to find information from different sources with different kind of searches [DS]. The lecturer [T] could present her ideas interestingly. There was also a couple of visiting lecturers [T] who added variety to lectures [IDa]. It was good that there wasn't any "right answers" [ID] to for example in web exercises [IDc], but you could test and practice [ID] as much you wanted to form your answer. This testing and practice [ID] taught a lot of IR. You could also do the tutored exercises [IDb] quite independently [ID], but you received advice [ID] in problem situations. These different parts of course gave me an overall view of IR and I feel that I have learned a lot of IR and things related to it.

Negative script: Student takes part in the introductory course of information retrieval in information studies. Course consists of lectures, web exercises and tutored exercises. After the course she thinks that she has not learned a lot of IR. Use your imagination and write a short - realistic - story or description of what inhibited learning in this course.

Example story with theme coding: Student wasn't very interested in learning [SO] new things. She wanted to earn credits as easily as possible [SO]. She didn't attend lectures [IDa] at all, so she missed a lot. Although she had to read lecture material [IDa] so much that she could do eight compulsory web exercises [IDc]. It forced her to learn a bit of IR, but it was very limited. She was forced to attend tutored exercises [IDb]. In there you had to work in pairs [OS] and she let the pair to do the job [OS]. Well, she gave some ideas for exercises, but she really didn't learn a lot. As a whole she felt that she was forced to learn [SO] a bit of IR in tutored exercises, but due to her own passive attitude [SO] concepts and deeper knowledge remained unlearned.
Table A1

Key to codes:

Code	MEBS themes
SO	Study orientation
DS	Domain of study
ID	Instructional design
ID a	a) Lectures
ID b	b) Tutored exercises
ID c	c) Web exercises
ID d	d) Feedback
ID e	e) Examination
ID f	f) Abstraction level
T	Teachers
OS	Other students
CS	Computer skills
EL	Everyday life

(

(

9. References

[1]   S. Webber and B. Johnston, Conceptions of information literacy : New perspectives and implications, *Journal of Information Science* 26(6) (2000) 381-397.

[2]   S. Harter, *Online Information Retrieval* (Academic Press, Orlando, 1986).

[3]   W.R. Hersh, *Information Retrieval : A Health Care Perspective* (Springer, New York, 1996).

[4]   F.W. Lancaster and A.J. Warner, *Information retrieval today* (Information Resource Press, Arlington, 1993).

[5]   A. Large, L.A. Tedd, and R.J. Hartley, *Information Seeking In The Online Age : Priciples And Practice* (Bowker-Saur, London, 1999).

[6]   I. Hsieh-Yee, Teaching online and CD-ROM resources : LIS educators' views and practices, *Journal of Education for Library and Information Science* 38(1) (1997) 14-34.

[7]   J.M. Still, How online is taught in British and American library schools. In: M. E. Williams (ed.), *National Online Meeting 1993: Proceedings of the 14th National Online Meeting, New York, 4-6 May 1993* (Learned Information, New York, 1993).

[8]   L. Holman, A Comparasion of computer-assisted instruction and classroom bibliographic instruction, *Reference and User Services Quarterly* 40(1) (2000) 53-60.

[9]   B. Bren, B. Hillemann, and V. Topp, Effectiveness of hands-on instruction of electronic resources, *Research Strategies* 16(1) (1998) 41-51.

[10]   N.G. Tomaiuolo, Effective simultaneous hands-on drill for basic electronic database instruction, *Research Strategies* 16(2) (1998) 135-145.

[11]   R.N. Sinn, Library instruction for biology courses: A literature review and survey, *Research Strategies* 16(2) (1998) 103-115.

[12]   C. Wien, Teaching online information retrieval to students of journalism, *Aslib Proceedings* 52(1) (2000) 39-47.

[13]   J.R. Hill and M.J. Hannafin, Cognitive strategies and learning from World Wide Web, *Educational Technology Research and Development* 45(4) (1997) 37-64.

[14]   S.M. Land, and B.A. Greene, Project-based learning with the World Wide Web : A qualitative study of resource integration, *Educational Technology Research and Development* 48(1) (2000) 45-67.

[15]   A.W. Lazonder, Exploring novice users' training needs in searching information on the WWW, *Journal of Computer Assisted Learning* 16(4) (2000) 326-335.

[16]   R. Oliver, The influence of instruction and activity on the development of skills in the usage of interactive information systems, *Education for Information* 14(1) (1996) 7-17.

[17]   R. Oliver and H. Oliver, Using context to promote learning from information-seeking tasks, *Journal of the American Society for Information Science* 48(6) (1997) 519-526.

[18]   C.C. Kuhlthau, *Seeking Meaning: A Process Approach To Library And Information Services* (Ablex, Greenwich, 1993).

(

[19] M.B. Eisenberg and R.E. Berkowitz, *Curriculum Iniative: An Agenda And Strategy For Library Media Programs* (Linworth, Worthington, 1988).

[20] M.B. Eisenberg, *Information Problem-Solving: The Big Six Skills Approach To Library And Information Skills Instruction* (Ablex, Norwood (N. J.), 1990).

[21] A. Irving, *Study And Information Skills Across The Curriculum* (Heinemann Educational Books, London, 1985).

[22] B. Stripling and J. Pitts, *Brainstorms And Blueprints: Teaching Library Research As A Thinking Process* (Libraries Unlimited, Littleton (CO), 1988).

[23] *Information Literacy Competency Standards For Higher Education* (Association of College and Research Libraries, Chigago IL, 2000). Available at: www.ala.org/acrl/ilcomstan.html

[24] C. Bruce, *The Seven Faces Of Information Literacy* (Auslib Press, Adelaide, 1997).

[25] D. Bawden, Information and digital literacies: A review of concepts, *Journal of Documentation* 57(2) (2001) 218-259.

[26] D. Wood, J.S. Bruner, and G. Ross, The role of tutoring in problem solving, *Journal of child psychology and psychiatry* 17(2) (1976) 89-100.

[27] A. Collins, J.S. Brown, and S.E. Newman, Cognitive apprenticeship : teaching the crafts of reading, writing and mathematics, In: L.B. Resnick (ed.), *Knowing, Learning And Instruction : Essays In Honor Of Robert Glaser* (Lawrence Erlbaum Associates, Hillsdale, (N.J.), 1989).

[28] L.Vygotsky, *Mind In Society : The Development Of Higher Psychological Processes* (Harward University Press, Cambridge, Mass., 1978).

[29] K. Winnips and C. McLoughlin, *Six WWW based learner supports you can build*. In: C. Montgomerie and J. Viteli (ed.) *ED-Media 2001 World Conference on Educational Multimedia, Hypermedia & Telecommunications, Tampere, 25-30 June 2001* (AACE, Norfolk, 2001).

[30] M. Gutzdial, Software-realized scaffolding to facilitate programming for science learning, *Interactive Learning Environments* 4(1) (1994) 1-44.

[31] Cognition and Technology Group at Vanderbilt, Anchored instruction and its relationship to situated cognition, *Educational Researcher* 19(3) (1990) 2-10.

[32] Cognition and Technology Group at Vanderbilt, Technology and the design of generative learning environments, *Educational Technology* 31(5) (1991) 34-40.

[33] Cognition and Technology Group at Vanderbilt, The Jasper experiment : An exploration of issues in learning and instructional design, *Educational Technology Research and Development* 40(1) (1992) 65-80.

[34] Cognition and Technology Group at Vanderbilt, Designing learning environments that support learning : the Jasper series as a case study. In: T.M. Duffy, J. Lowyck, and D.H. Jonassen (eds), *Designing Environments For Constructive Learning* (Springer, Berlin, 1993).

[35] P. Honebein, T. Duffy, and B. Fishman, Constructivism and the design of learning environments: context and authentic activities for learning. In: T.M. Duffy, J. Lowyck, and D.H. Jonassen (eds), *Designing Environments For Constructive Learning* (Springer, Berlin, 1993).

[36] W.R. Shadish, T.D. Cook, and D.T. Campbell, *Experimental And Quasi-Experimental Designs For Generalized Causal Inference* (Houghton Mifflin, Boston, 2002)

[37] D. Kolb, *The Learning Style Inventory Technical Manual* (McBer and Company, Boston, 1976)

(

[38]   D. Kolb, *Experiental Learning: Experience as the Source of Learning and Development* (Prentice-Hall, Englewood Cliffs (N.J.), 1984)

[39]   K. Halttunen, Students' conceptions of information retrieval: Implications for design of learning environments, *Library and Information Science Research* 25 (2003) In Press.

[40]   K. Halttunen and E. Sormunen, Learning information retrieval through an educational game: Is gaming sufficient for learning?, *Education for Information* 18(4) (2000) 289-311.

[41]   E. Sormunen, K. Halttunen, and H. Keskustalo, *Query Performance Analyser : A Web-Based Tool for Information Retrieval Research and Instruction* (Department of Information Studies, University of Tampere, Tampere, 2002).

[42]   E. Sormunen, et al. IR Game : a toolfor rapid query analysis in cross-language IR experiments. In: *Joint workshop on cross language issues in artifical intelligence & issues of cross cultural communication* (Kent Ridge Digital Labs, Singapore, 1998).

[43]   A. Eskola, Non-active role-playing: some experiences. In: A. Eskola in collaboration with A. Khilström, D. Kivinen, K. Weckroth and O. Ylijoki (eds), *Blind Alleys in Social Psychology: A Search for Ways Out* (North-Holland, Amsterdam, 1988)

[44]   J. Eskola, *Eläytymismenetelmä sosiaalitutkimuksen tiedonhankintamenetelmänä [The Method of Empathy-based Stories as a Method of Acuiring Data in Social Research]* (TAJU, Tampere, 1998).

[45]   G.P. Ginsburg, Role-playing and role performance in social psychological research. In: M. Brenner, P. Marsh, and M. Brenner (eds), *The Social Context of Method* (Groomhelm, London, 1978).

[46]   G.P. Ginsburg, The effective use of role-playing in social psychological research, In: G.P. Ginsburg (ed.), *Emerging Strategies in Social Psychological Research* (John Wiley & Sons, New York, 1979).

[47]   Miles, M.B. and A.M. Huberman, *Qualitative Data Analysis: An Expanded Sourcebook*. 2nd. ed. (Sage, Thousand Oaks (CA), 1994)

[48]   K. Halttunen and K. Järvelin, Assessing learning outcomes in two information retrieval learning environments : A design experiment, *[article draft]* (2003).

[49]   G. Salomon It's not just the tool, but the educational rationale that counts : invited keynote. In: J. Bourdeau and R. Heller (eds) *Ed-Media 2000 : World Conference on Educational Multimedia, Hypermedia & Telecommunications, Montreal, 26 June - 1 July 2000* (AACE, Charlottesville, 2000).