Abstract

This paper describes the development and empirical evaluation of a web training for pupils (CIS-WEB, Competent Information Search in the World Wide WEB) which aims to convey prerequisite knowledge and skills that are necessary for a competent search for information on the web. The web training focuses on competent information handling and is based on two theoretical analyses. First, a conceptual analysis of information search from the perspective of media literacy research and information retrieval research was conducted and yielded a set of five pivotal content aspects that need to be covered by a web training. Each of these content aspects is characterized by declarative and procedural knowledge components which are necessary for the pursuit of a competent search for information on the web. Second, we conducted a task analysis which conceptualizes the search for information on the web as a problem-solving process and which allows to systematically distinguish between different types of information problems. In the empirical part of the paper two classroom studies are reported. In Study 1, the widespread training concept of a technically oriented Internet training for pupils was evaluated and it was shown that no substantial improvement of web searching skills can be expected from this type of treatment. In Study 2, it was shown that the web training CIS-WEB improves pupils’ declarative knowledge of the web as well as their search performance, thereby outperforming the conventional Internet training used in Study 1.

© 2007 Elsevier Ltd. All rights reserved.
1. Introduction

The Internet provides numerous services for the user (e.g., email, instant messaging, chat or groupware), however, undoubtedly one of its primary applications is the search for information in the World Wide Web (WWW). Although the birth of the WWW at the European laboratory for particle physics CERN in 1993 took place less than fifteen years ago, the technical development since then (particularly with regard to powerful web browsers, advanced multimedia plug-ins, user friendly authoring tools for web pages, and the physical network structure) enables individuals worldwide to comfortably publish and retrieve information on almost any topic and in almost any representation format imaginable (e.g., text, pictures, audio, video, animation). At the same time today’s information society creates an omnipresent demand for immediate access to information about diverse topics. Thus, the WWW seems to be the information environment of choice to meet these needs. In line with these developments, school children and young adults take the availability of Internet-based information resources for granted. According to the ARD/ZDF-Online-Study 2004 (Van Eimeren, Gerhard, & Frees, 2004) as many as 95% of the individuals between the age of 14 and 19 years regularly use the Internet in Germany. With regard to the Internet availability in school, it has been shown that, for instance, 99% of the US public schools in 2001 had access to the Internet (National Center for Educational Statistics, 2002).

Contrary to these promising introductory remarks with regard to the potential benefits and the ubiquitous availability of the WWW, however, the web is also characterized by several problematic features that impose additional cognitive processing demands and knowledge requirements onto information seekers. Due to these features, the web can be seen as “a unique searching environment that necessitates further and independent study” (Jansen & Pooch, 2001, p. 244). Characteristic features of the WWW are, for instance, its worldwide open access with regard to information retrieval and authorship, its heterogeneity of the user population, the distributed authorship of documents, a lack of a central agency responsible for structure and quality of document collections, a lack of temporal stability, an enormous amount of information available, the heterogeneity of contents, an unordered network structure, and the integration of interactive multimedia components (Blair, 2002; Hölscher, 2000; Jansen & Pooch, 2001; Jenkins, Corritore, & Wiedenbeck, 2003; Kleinberg & Lawrence, 2001). As a result, web users need specific skills ranging from orientation and search planning to information selection and evaluation in order to benefit from the enormous information resources provided by the WWW.

Considering the abovementioned fact that school children quite early encounter the web as an information environment, it seems quite natural that web-related skills should be conveyed and trained in schools. Whereas it seems rather improbable that pupils may acquire these skills and knowledge prerequisites on the fly by engaging in web searches on their own it is a common assumption that attending one of the currently popular short term Internet trainings might provide a sufficient training for competent information search on the web. Nevertheless, it has to be critically noted that existing Internet trainings predominantly focus on technical aspects and neglect most of the abovementioned cognitive aspects of skilled information searches. In our own research, we therefore developed the user-oriented training program CIS-WEB (Competent Information Search in the World Wide WEB) that aims at improving cognitive and metacognitive aspects of pupils’ ability to competently search for information in the WWW.
The development of CIS-WEB is based on two theoretical analyses, which will be described in the subsequent sections in greater detail. First, a conceptual analysis of information search within the fields of media literacy research and information retrieval research was conducted to identify pivotal content aspects that need to be covered by a web training. Second, we conducted a task analysis which conceptualizes the search for information on the web as a problem-solving process in order to systematically distinguish between different types of information problems that need to be addressed by a web training.

2. Conceptual analysis of information search from the perspective of media literacy research and information retrieval research

In a first step we tried to identify important content aspects that need to be covered by a training program for competent information search in the WWW. We conducted a conceptual analysis within the fields of media literacy research and information retrieval research because in both disciplines users competencies to search for information are of focal interest. The dominant approach in media literacy research is to decompose the overarching concept of media literacy into a list of different sub-competencies. In a similar fashion, information retrieval research focuses on processes of information search and tries to identify specific sub-processes that are of pivotal importance for competent information retrieval.

Taking prominent models of media literacy and of information retrieval into account allows to derive a set of uncontroversial content aspects that need to be included in a training for competent information search in the WWW.

2.1. Media literacy research

Based on theoretical analyses and – at least partly – on observational studies, media literacy research decomposes the concept of a global “media literacy” or “information literacy” into a set of sub-competencies (cf. Gapski, 2001, for an overview). These sub-competencies are usually conceived as being independent of the specific medium used and of users’ goals in the context of media utilization. Spitzer, Eisenberg, and Lowe (1998, p. 22) defined that “people trained in the application of information resources to their work can be called information literates. They have learned techniques and skills for utilizing the wide range of information tools”. According to this definition, the concept of information literacy can be easily applied to the specific situation of web-based information search. From the perspective of information literacy, Doyle (1992, p. 8) proposes the following list of important sub-competencies of an information literate user: “Recognizes that accurate and complete information is the basis for intelligent decision making, recognizes the need for information, formulates questions based on information needs, identifies potential sources of information, develops successful search strategies, accesses sources of information including computer-based and other technologies, evaluates information, organizes information for practical application, integrates new information into an existing body of knowledge, uses information in critical thinking and problem solving”. Gapski (2001) conducted a content analysis of 104 definitions of media literacy, revealing that the following six categories are most important to classify the different sub-competencies of users: (1) technical knowledge on how media work, (2) practical knowledge on how to use media, (3) self-reflective competencies with regard to the relation between media usage
and media environment, (4) creative skills when using media, (5) social-reflective competencies with regard to how media influence social responsibility and political action, and (6) affective competencies in coping with media effects. With regard to information search in the WWW, technical and practical knowledge as well as self-reflective competencies can probably be seen as the most relevant categories. When bringing together comprehensive models of media literacy and information literacy, the following four sub-competencies can be derived as the most important content aspects when the focus is on information search in the WWW (Schorr, 2005).

- **Media background knowledge**: Background knowledge with regard to the development and structure of the Internet and with regard to specific features of the WWW as information environment.
- **Media operation skills**: Skills for using computers, the Internet, and the WWW (e.g., how to connect to the Internet, how to use a browser software, how to use search engines and other web tools).
- **Orientation skills**: Ability to keep oriented with regard to the information sources provided by the WWW.
- **Selection and evaluation skills**: Ability to evaluate information provided in the WWW with regard to its relevance in the context of a current information problem as well as with regard to its quality and credibility. Ability to select information according to these evaluation criteria.

### 2.2. Information retrieval research

Compared to the research on media literacy and information literacy the research on information retrieval is more empirically oriented and has a user-oriented focus on the process of information search via information technology (e.g., Belkin, 1978; Fidel, 1991; Ingwersen, 1992; Jansen & Pooch, 2001; Saracevic, 1999). Information retrieval research is rooted in information science and uses methods like surveys, questionnaires, field studies, and controlled experiments. Within this area of research the process of information search is usually segmented into sub-processes that need to be pursued in specific situations like information search in the WWW. Models of information retrieval usually specify a set of those sub-processes and propose an appropriate sequence (e.g., Blair, 2002; Clark, 2002; Eisenberg & Berkowitz, 1996; Kuhlthau, 1993; Marchionini, 1995). Some models of information retrieval are already specified with regard to the WWW as information environment (e.g., Hargittai, 2002; Hölscher, 2000; Laus, 2001; Navarro-Prieto, Scaife, & Rogers, 1999; Shneiderman, Byrd, & Croft, 1998). A critical synopsis of these models leads to the conclusion that five sub-processes of information retrieval are most important. These sub-processes can be described as specification of information requirements, application of search strategies, handling of search systems, selection and evaluation of information and information sources, and monitoring of processes and results of information searches.

### 2.3. Pivotal content aspects that need to be covered by a web training

In an integrative approach, we combined the content aspects derived from both lines of research (media literacy and information retrieval) to come up with a list of five integrated content aspects which were used as the basis to develop the web training CIS-WEB.
• **Representation of the WWW as an information environment**: Includes media background knowledge and media operation skills with regard to the WWW.

• **Specification of information requirements**: Includes segmenting of information problems into sub-problems, sequencing of information sub-problems, formulation of hypotheses with regard to the type and localization of information required.

• **Application of search strategies**: Includes typing in URLs, following links, using search engines, distinguishing appropriate situations for different search strategies.

• **Selection and evaluation of information and information sources**: Includes the evaluation of relevance, quality, credibility, and actuality of information in the context of a current information problem as well as the selection of information according to these criteria.

• **Monitoring of information searches**: Includes knowledge on the monitoring of processes and results of information searches.

3. Task analysis: Information search as problem solving

   Whereas the preceding conceptual analysis of information search from the perspective of media literacy research and information retrieval research allowed deriving a set of content aspects that need to be covered by the CIS-WEB training, an additional task analysis of information search tasks was conducted to come up with a systematic taxonomy of different information problems. One reason for why such a taxonomy is most helpful in designing a web training consists in the fact that a competent search for information comprises – among other things – the adaptation of search strategies to the different types of information problems which are to be solved. In order to describe these processes of strategic adaptation and in order to teach pupils within CIS-WEB how to distinguish appropriate situations for different search strategies, information problems need to be classified systematically.

   As there is no established taxonomy of information problems available yet, we conducted a task analysis based on the problem-space paradigm by Newell and Simon (1972) which allows to describe different types of information problems as different types of path in a problem space. A problem space is characterized by its initial situation, its target situation, and a set of operators that can be applied to transform situations. A sequence of operators that transforms the initial situation into the target situation is a problem solution. Important intermediate situations between the initial and the target situation that have to be solved in order to finally achieve the target situation can be described as sub-goals. Problems that share the sub-goals also share the same solution structure and can thus be classified to belong to the same problem category. Once the sub-goal structure of a problem and the set of applicable operators are clearly characterized, the declarative and procedural knowledge necessary for problem solving can be defined on a formal basis.

   We used a problem-space analysis to define a list of generic sub-goals that cover different types of information problems occurring on the web. These sub-goals were arranged in a sub-goal hierarchy that provided the basis for the identification of four different sub-goal structures. A sub-goal structure can be used to define a category of information problems. All information problems within a category will share the same structure of sub-goals that have to be achieved in order to solve the overall information problem. In this task analysis we identified four pivotal sub-goals of information problems in the WWW, namely, identification of sub-tasks ("What exactly is the search task and which sub-tasks need..."
to be pursued?”), selecting an information provider (“Who can provide the necessary information?”), localization of a website (“Which URL allows to access a known information source?”), and localization of information on a website (“Where is the information on a known site?”), cf. Fig. 1).

These sub-goals can be arranged hierarchically in order to reflect the prerequisite relationships between them (e.g., localizing a website is a prerequisite sub-goal for localizing an information on a website). The sub-goal hierarchy provided the basis for the definition of four sub-goal structures of information problems. While sub-goal structure 1 covered only information problems that require to achieve the sub-goal at the hierarchy’s lowest level – the localization of information on a website – each consecutive sub-goal structure additionally adds sub-goals of the next hierarchical level. Thus, sub-goal structure 4 contains all four sub-goals.

For an example, consider the following search task: Since when is the burger Whopper offered by Burger King (www.burgerking.de)? In this example, the URL of a website is given as information source for the information asked. This task has the lowest possible sub-goal structure 1 and can, therefore, be solved by simply retrieving the website and localizing the information. However, in the case of more complex search tasks additional processing steps are needed in order to find a task’s solution.

In sum, the task analysis illustrated in Fig. 1 involves two basic ideas: First, search tasks can be distinguished according to their complexity (or sub-goal structure) that is determined by the amount of information given for task accomplishment. Second, information search is adaptive in nature, i.e., the complexity of a search task should influence how the task should optimally be processed. Thus, one of the goals of the web training CIS-WEB was to teach pupils (a) to distinguish information problems of different sub-goal structure and (b) to adapt their search strategies to the sub-goal structures of information problems to be solved.

In line with this reasoning, CIS-WEB conveys knowledge about the four different types of sub-goal structures of information problems that have been identified by the task analysis. For the definition of appropriate processing strategies, we extended our task analysis according to the GOMS model (Goals, Operators, Methods, and Results; Card, Moran, 698


Fig. 1. Sub-goals, sub-goal hierarchy, and sub-goal structures of information problems.
& Newell, 1983) which is an architecture for cognitive modeling that is commonly used for cognitive task analyses. GOMS-models contain the following components:

- **Goals.** The goals and sub-goals that are necessary for the solution of a task are organized in a sub-goal hierarchy.
- **Operators.** The processing steps for the solution of a task comprise external and internal actions on a rather low level.
- **Methods.** Sequences of operators are combined into comprehensive methods which are hierarchically organized. Methods are described as operator sequences.
- **Selection rules.** These decision rules are composed of a condition part and an action part, and they guide the choice of goals, operators, and methods.

This extended GOMS task analysis was used to define schematic knowledge on when (in terms of sub-goal structures) certain solution procedures (in terms of operators and methods) are to be applied (see Schorr, 2005). It has often been argued that the availability of this kind of abstract *problem schemata*, that is, representations of *problem categories* together with category-specific *solution procedures* is the most important cognitive prerequisite for successful problem solving in knowledge-rich domains (VanLehn, 1989). Schema-based approaches to complex problems have been seen as a cornerstone of experts performance in many content domains (e.g., Cummins, 1992; Gick & Holyoak, 1983; Marshall, 1995; VanLehn, 1989). In line with these approaches, CIS-WEB conveys knowledge about the sub-goal structures of information problems, the appropriate operators and methods depending on the sub-goal structure as well as the selection rules for deciding when certain operators and methods should be applied. Thus the aim of CIS-WEB is to enable pupils to engage in schema-based information search on the web based on the knowledge conveyed by the training.

### 4. Design of the web training CIS-WEB

In order to design the web training CIS-WEB, we focused on the five pivotal content aspects identified in the conceptual analysis as well as on the four different sub-goal structures of information problems described by the task analysis conducted. Based on these two analyses, six training modules were developed for the target group of pupils and for implementation in the classroom. CIS-WEB used a problem-based training approach according to Reinmann-Rothmeier and Mandl (2000). All six training modules are covered in a total of twelve lessons a 45 min. We will first introduce the contents covered by the six modules before the different instructional methods and the training plan of CIS-WEB are described in greater detail. For a full description of CIS-WEB see Schorr (2005). Most of the training materials can be downloaded online at www.tinaschorr.de/KIS-WEB/index.htm.

#### 4.1. Contents of CIS-WEB

- **Module 1. Representation of the WWW as an information environment:** Basic knowledge on the Internet, the WWW, and search systems on the web
  - Internet: Definition, services, origin, user statistics, access options.
- WWW: Different representational formats of information, contents, authors, lack of quality control, security advice, continuous change of available information, definitions of technical terms (e.g., URL, link, website, homepage), operating web-browser software.
- Search Systems: Definitions (search systems, search catalogs, search engines), functionality of different search systems, selection criteria for the use of different types of search system.

**Module 2. Information problems:** Basic knowledge on information problems, their potential sub-goals and complex sub-goal structures
- Categorization of information problems according to their sub-goal structure and specification of information requirements for the different categories of information problems.
- Basic knowledge on how to adapt search strategies to the sub-goal structures of information problems.
- Introduction to the four pivotal sub-goals of information problems in the WWW: Localization of information on a website, localization of a website, selecting an information provider, and identification of sub-tasks.

Each of the remaining four modules conveys operators, methods, and selection rules with regard to one of the four pivotal sub-goals of information problems.

**Module 3. Localization of an information on a website:** Structure of websites and auxiliary tools (e.g., sitemaps)
- Typical structure of a website.
- Utilization of search functions and sitemaps as specific web tools.
- Strategies for quickly accessing the contents of websites and webpages.
- Evaluating the relevance of information in the context of a current information problem.

**Module 4. Localization of a website:** Methods for retrieving websites (e.g., by inferring their URL or by using search systems)
- Direct retrieval of websites: Inferring URLs on the basis of the typical URL syntax.
- Following links: Inferring auxiliary websites which are expected to contain links to a target website.
- Using search systems: Selection and use of search portals and search engines, selection of search keywords, defining queries, selecting links in query results.

**Module 5. Selecting an information provider:** Evaluation and selection of an information provider with regard to credibility and actuality
- Evaluation of the credibility of an information source.
- Evaluation of the actuality of an information source.

**Module 6. Identifying sub-tasks of information problems:** Breaking down complex information problems into sub-tasks and selecting strategies to solve these sub-tasks
- Decomposition of a complex information problem in sub-tasks by specifying the different information requirements.
- Identifying operators and methods to achieve identified sub-tasks.
- Monitoring processes and results of information searches.
4.2. Instructional methods used in CIS-WEB

CIS-WEB is a computer-based in-class training that uses three different social settings to implement a problem-oriented training approach:

- **Regular classroom teaching.** Basic concepts were introduced and discussed in the classroom based on PowerPoint presentations.
- **Working in dyads.** The details of each training module were worked out in student dyads based on a computer-based hypermedia environment.
- **Individual work.** Exercises and problem-solving assignments based on paper–pencil materials had to be solved by individual learners and were corrected either by student dyads based on worked-out solutions or by the whole class with the help of the teacher.

As instructional methods, CIS-WEB comprises worked-out examples, symbolic and iconic visualizations, interactive multiple-choice questions with feedback, working sheets and exercises.

- **Elaborated worked-out examples.** Example information problems from different content domains were presented together with a step-by-step solution in order to convey the details of finding a solution to information problems.
- **Symbolic and iconic visualization.** Symbolic visualizations were used to illustrate complex and abstract relationships (e.g., structure of the Internet), and iconic visualizations were used to demonstrate important procedures for information search.
- **Interactive multiple-choice questions with feedback.** Interactive questions were introduced to allow pupils to directly check their knowledge and to receive feedback on their answers.
- **Worksheets.** While studying the training modules pupils received paper–pencil based questions related to the contents they were currently studying.
- **Exercise worksheets.** Subsequent to each training module paper–pencil based problem-solving assignments had to be solved that forced pupils to apply their newly acquired knowledge.
- **Recapitulations.** At the beginning of each new training module, the contents of the previous modules where recapitulated orally in the classroom. An overall recapitulation took place at the end of the complete web training.

4.3. Training plan

The training plan of CIS-WEB displayed in Table 1 summarizes the distribution of the training modules over the 12 lessons a 45 min as well as the social settings and the instructional materials used. CIS-WEB can be implemented in schools according to this training plan.

In the remainder of this paper, two empirical studies will be reported that aim at substantiating the instructional value of CIS-WEB. As a first step, Study 1 tests the common assumption that attending one of the currently popular short term Internet trainings in school might be a sufficient intervention to teach web search skills. To do so, Study 1 evaluates the effectiveness of a conventional Internet training (“Surfcheck-Online”) that can be studied within approximately two lessons. Based on the results
of this evaluation, Study 2 investigates the effectiveness of CIS-WEB to convey prerequisite knowledge and skills that are necessary for a competent search for information on the web.

5. Study 1: Evaluation of a popular conventional web training (“Surfcheck-Online”)

Existing Internet trainings predominantly focus on technical aspects of web search and neglect most of the cognitive aspects of skilled information searches. In particular, they do neither rely on the five pivotal content aspects identified in our conceptual analysis nor on the four different sub-goal structures of information problems described by our task analysis. In order to test whether such a conventional and technology-focused approach is sufficient to improve pupils search skills on the WWW, the first empirical study was conducted as an evaluation of a popular technically oriented Internet training. We used the prototypical Internet training “Surfcheck-Online” which is implemented as an online hypermedia environment (www.surfcheck-online.de) and which was explicitly designed for pupils by a well-known German non-profit organization (“Schulen ans Netz”). The training comprises five chapters on Internet access, basic knowledge on the Internet, navigation, communication, and security and needs approximately two lessons of study. It is probably assumed by the designers of such an Internet training that it will improve pupils declarative knowledge with regard to the Internet and that it will facilitate the search for information on the web compared to an unguided exploration of the web. Therefore, we tested the following two hypotheses in this study:

Hypothesis 1. Declarative knowledge acquisition. Participating in a conventional Internet training (compared to an unguided exploration of the web) will increase pupils’ search-relevant and search-irrelevant declarative knowledge.
This hypothesis is based on inspecting different Internet trainings that are available online (e.g., webfuehrerschein.web.de, www.kidstation.de, www.surfcheck-online.de). This inspection reveals that these trainings primarily convey factual knowledge on the Internet, on communication tools, and on aspects of security. Thus, it was expected that participating in an Internet training would result in an increased declarative knowledge. Furthermore, we distinguished between two different types of declarative knowledge: Search-relevant knowledge was defined as knowledge that directly helps in conducting web searches (e.g., declarative knowledge on how to operate a search engine) while search-irrelevant knowledge was defined as knowledge that does not directly help in conducting web searches (e.g., knowledge on the history of the Internet).

Hypothesis 2. Search performance improvement. Participating in a conventional Internet training (compared to an unguided exploration of the web) will improve pupils’ search performance when trying to solve information problems with the help of the WWW, at least if these information problems are not too complex with regard to their sub-goal structure.

This hypothesis expresses the assumption that a conventional Internet training will not only improve factual knowledge but will also have a positive effect on the search performance of pupils working on information problems. However, when considering our task-analytical taxonomy of information problems according to their sub-goal structure, it can be claimed that performance improvements are only to be expected with regard to the two simpler sub-goal structures 1 and 2 but not with regard to the two more complex sub-goal structures 3 and 4. This expectation is based on the fact that conventional Internet trainings usually do not address the issues of evaluating information sources with regard to credibility and actuality and of breaking down complex information problems into sub-tasks and selecting strategies to solve these sub-tasks.

5.1. Method

5.1.1. Participants

The study’s participants were 28 pupils (sixth grade) from a public German high school (21 female, 7 male). Average age was 11.92 years (SD = 0.48).

5.1.2. Materials and procedure

The study took place on three subsequent days (with two lessons a 45 min per day). Day 1 was used for pretesting. All participants had 15 min to fill in a declarative knowledge test with 20 multiple-choice items and 75 min to solve one out of two sets of 16 information problems each. Additionally, a questionnaire was administered to obtain pupils’ computer experience and Internet experience. Day 2 was used to implement the instructional intervention (either the Internet training “Surfcheck-Online” or an unguided exploration of the WWW, depending on the experimental condition). All pupils had a computer with a web browser software at their disposal and had 90 min of time to either study the Internet training online or to explore the WWW on their own. Day 3 was used for the posttest. All participants had to fill in the initial declarative knowledge test again (with a different sequencing of the 20 multiple choice items than in the pretest). Subsequently they had to solve the second set of 16 information problems.
5.1.3. Design and dependent variables

As a first independent variable the instructional condition on Day 2 was manipulated as between-subject variable (Internet training “Surfcheck Online” versus unguided exploration of the WWW).

As a second independent variable the sub-goal structure of the information problems to be solved in the pretest and the posttest was manipulated as within-subject variable. The 16 information problems within each set of information problems comprised four problems of each problem structure: Problem structure 1: Localization of information on a website only (“Where is the information on a known site?”), problem structure 2: Additionally localization of a website (“Which URL allows to access a known information source?”), problem structure 3: Additionally selecting an information provider (“Who can provide the necessary information?”), and problem structure 4: Additionally identification of sub-tasks “What exactly is the search task and which sub-tasks need to be pursued?”, (cf. Fig. 1). Both sets of information problems were used for pretesting as well as for posttesting and where counterbalanced within the instructional conditions. With regard to the content domains of the information problems, six important school subjects were covered (Math, German, Biology, Geography, Art/Music, and Sports).

As a first dependent variable we obtained declarative knowledge gains with regard to search-relevant and search-irrelevant facts. The declarative knowledge test consisted of 20 multiple-choice items (10 items were related to search-relevant facts and 10 items were related to search-irrelevant facts). We used the difference between the number of correct answers in the pretest and the posttest for both types of items as dependent variables.

As a second dependent variable we obtained search performance gains with regard to the four different problem structures. The solution for each of the 32 information problems presented to pupils could be unequivocally rated as correct or false. We used the difference between the number of correctly solved information problems in the pretest and the posttest for the four different problem structures as dependent variables.

Pupils’ computer experience and Internet experience were obtained as control variables by means of a questionnaire.

5.2. Results

All pupils reported to use computers either in school or at home for at least one year. 92% of the pupils reported to use the Internet either in school or at home (at least for one year: 73%; at least once a week: 50%). 92% of the pupils reported to be experienced with regard to information retrieval in the WWW.

Pupils’ prior declarative knowledge with regard to search-irrelevant facts (73.1% correct, SD = 13.2) was quite good and better than their prior knowledge with regard to search-relevant facts (66.9% correct, SD = 11.6; t(25) = 2.22; p < .05, two-tailed).

However, this prior knowledge was obviously not sufficient for solving the information problems as the poor overall pretest performance strongly demonstrated (17.2% correct solutions, SD = 11.5; sub-goal structure 1: 33.0% correct, SD = 25.5; sub-goal structure 2: 12.5% correct, SD = 14.4; sub-goal structure 3: 11.6% correct, SD = 17.3; sub-goal structure 4: 11.6% correct, SD = 17.3).
With regard to the first hypothesis on declarative knowledge gains due to the Internet training, we conducted a MANOVA for the between-subject variable instructional condition and knowledge gains for search-irrelevant vs. search-relevant items. This analysis yielded a significant effect of the instructional condition for search-irrelevant items \( F(1,23) = 6.34, \text{MS}_e = 167.14; p < .05 \) but not for search-relevant items \( F < 1 \).

For search-relevant items, knowledge gains in both instructional conditions were not significantly different from 0 (mean gain = 0.80%; SD = 16.3; \( t(24) = 0.25; p > .80, \text{two-tailed} \)). For search-irrelevant items, pupils in the Internet training condition did not significantly improve their knowledge (mean gain = 4.55%; SD = 12.9; \( t(10) = 1.17; p > .20, \text{two-tailed} \)). Pupils in the unguided exploration condition even showed negative knowledge gains from pretest to posttests (mean gain = −8.57%; SD = 12.9; \( t(13) = −2.48; p < .05, \text{two-tailed} \)). Thus, regardless of instructional condition, there was no increase neither in search-irrelevant, nor in search-relevant declarative knowledge. Consequently, Hypothesis 1 about the increase of declarative knowledge could not be confirmed. Surprisingly, while there was no change in search-irrelevant knowledge in the web training condition, there was even a decrease in the unguided exploration condition which might be attributable to confusion and frustration.

With regard to the second hypothesis on search performance gains due to the Internet training, we conducted a MANOVA for the between-subject variable instructional condition and the within-subject variable sub-goal structure of the information problems. The results showed contrary to our hypothesis that there was no superiority of the Internet training compared to the unguided exploration of the web \( F < 1 \). Additionally, there were no differential knowledge gains for the four different sub-goal structures of the information problems and no interaction between the two factors (both \( F_s < 1 \)).

The mean increase in search performance from pretest to posttest (for all four sub-goal structures and both instructional conditions) was −4.46% (SD = 10.33). Thus, search performance was significantly worse after 90 min of unguided web exploration as well as after 90 min of Internet training \( (t(27) = −2.29; p < .05; \text{two-tailed}) \). Accordingly, Hypothesis 2 about the increase of search performance could not be confirmed. Again, confusion and frustration might be likely explanations for this effect.

**Effect of sub-goal structure on search performance.** In order to explore how search performance was influenced by the variation of task complexity (i.e., sub-goal structure 1–4) a MANOVA was conducted for the within-subject variable sub-goal structure (across instructional conditions and across pretest and posttest results). This analysis yielded a significant effect of the sub-goal structure of the information problems \( F(3,81) = 15.34; \text{MS}_e = 163.62; p < .001 \) which is in accordance with the task analysis on which this manipulation was based.

To sum up, Study 1 demonstrated that pupils face serious problems when searching the web, whereby these difficulties increase substantially when search tasks become more complex. Moreover, neither a conventional Internet training nor an unguided exploration of the web were helpful to improve pupils’ search performance when trying to solve information problems with the help of the WWW. Furthermore, both instructional conditions were also insufficient to improve the declarative knowledge with regard to search-relevant and search-irrelevant facts. Thus, technically oriented Internet trainings – as well as an unguided exploration of the WWW – have to be considered inadequate to facilitate competent information searches on the web. Thus, a more comprehensive training program like CIS-WEB is obviously needed.
6. Study 2: Evaluation of the web training CIS-WEB

Study 2 aimed at exploring whether CIS-WEB – contrary to the conventional web training “Surfcheck-Online” evaluated in Study 1 – is more suitable to improve pupils’ search competencies as well as their declarative web-related knowledge. Probably due to floor effects, the sixth graders participating in Study 1 displayed a very poor search performance. Therefore, participants in Study 2 were recruited from grade seven and grade eight, assuming that pupils search competence on the WWW increases with age. Study 2 uses a repeated-measurement design and mainly addressed four hypotheses:

Hypothesis 1. Age related differences. It is assumed that search competence on the WWW increases with age. Thus, pupils of the eighth grade should outperform seventh grade pupils with regard to their search performance and with regard to their declarative knowledge about the Internet. These differences should disappear after studying CIS-WEB due to training effects.

Hypothesis 2. Declarative knowledge acquisition. CIS-WEB improves pupils search-relevant and search-irrelevant declarative knowledge with regard to the web.

Hypothesis 3. Search performance improvement. Pupils’ search performance improves due to studying CIS-WEB. Search performance is expected to increase for all four sub-goal structures although information problems of more complex sub-goal structures are expected to result in an inferior overall search performance than information problems of simpler sub-goal structures.

Hypothesis 4. Time course of performance improvements and sub-goal structures. The time course of pupils’ search performance improvements depends on the complexity of the information problems. This assumption is based on the modular structure of the web training where the knowledge necessary to solve information problems of more complex sub-goal structures is conveyed later than the knowledge necessary to solve less complex search tasks. Consequently, search performance regarding information problems with more complex sub-goal structures is expected to only increase substantially in later phases of the web training.

6.1. Method

6.1.1. Participants
The study’s participants were 61 pupils (seventh and eighth grade) from a public German high school (30 female, 31 male). Average age was 12.74 years (SD = 0.57). Average age of the 30 pupils in the seventh grade was 12.33 years (SD = 0.48). Average age of the 31 pupils in the eighth grade was 13.30 years (SD = 0.34).

6.1.2. Materials and procedure
The study took place on three subsequent days and comprised the 12 lessons of CIS-WEB as described above plus six additional lessons that were needed for testing (a 45 min). Thus, the study took six lessons on each of the three days.

Day 1 was used for pretesting and for running Module 1 of CIS-WEB. At the beginning a questionnaire was administered to obtain pupils’ computer experience and Internet experience. All participants had 15 min to fill in a declarative knowledge test with 28 multiple-choice items and 75 min to solve one out of five sets of ten information problems each.
(Time of measurement 1, ToM 1). Subsequently pupils studied Module 1 and had another 75 min to solve a second set of ten information problems (ToM 2).

On Day 2, pupils had to fill in the declarative knowledge test with 28 multiple-choice items again (sequence of items manipulated, ToM 2), before they studied Module 2–4 and solved a third set of ten information problems (ToM 3).

Day 3 was used for Module 5 and 6 and the summary of the course. Subsequently the declarative posttest had to be filled in (ToM 3) and pupils had to solve the fourth and the fifth set of ten information problems (ToM 4). The overall procedure of Study 2 is illustrated in Fig. 2.

6.1.3. Design and dependent variables

As a first independent variable pupils’ school grade was manipulated between subjects (seventh grade versus eighth grade). As a second independent variable, time of measurement was manipulated within subjects (ToM 1 to ToM 4 for search performance, ToM 1 to ToM 3 for declarative knowledge, cf. Fig. 2).

As a third independent variable, the sub-goal structure of the information problems to be solved was manipulated as within-subject variable. Five sets of ten information problems were constructed whereby the overall set of 50 information problems comprised 12 problems of sub-goal structure 1: Localization of information on a website only, 14 problems of sub-goal structure 2: Additionally localization of a website, 12 problems of sub-goal structure 3: Additionally selecting an information provider, and 12 problems of sub-goal structure 4: Additionally identification of sub-tasks, (cf. Fig. 1). Each of the five sets with ten information problems contained two or three problems of each sub-goal structure and was used for all four times of measurement. The five sets were counterbalanced to avoid confounding of problem set with the time of measurement. With regard to the content domains of the information problems, the same school subjects were covered as Study 1 (Math, German, Biology, Geography, Art/Music, and Sports).

As a first dependent variable we obtained declarative knowledge with regard to search-relevant and search-irrelevant facts. The declarative knowledge test consisted of 28 multiple-choice items (14 items were related to search-relevant facts and 14 items were related to search-irrelevant facts). We used the percentage of correct answers for both types of items at ToM 1 to ToM 3 as dependent variables.

As a second dependent variable we obtained search performance with regard to the four different problem structures. The solution for each of the 50 information problems presented to pupils was rated independently by two raters with regard to their correctness on a scale from 0 (no answer/completely wrong answer) to 5 (completely correct answer). The mean interclass correlation between the two raters was $r = .94$ (SD = 0.03). We used

Fig. 2. Procedure of Study 2.
the ratings of correctness averaged across the two raters and transformed in percentage correct for all four sub-goal structures at ToM 1 to ToM 4 as dependent variables.

Pupils’ computer experience and Internet experience were obtained as control variables by means of a questionnaire. Additionally, we measured pupils’ learning intensity during their study of Module 1 to Module 6 by rating the correctness of their answers to the worksheets and exercises they had to fill in during the learning phase.

6.2. Results

All pupils reported to use computers either in school or at home (at least for one year: 93%). 98% of the pupils reported to use the Internet either in school or at home (at least for one year: 71%; at least once a week: 58%). Eighty-eight percentage of the pupils reported to be experienced with regard to information retrieval in the WWW. Thus, the computer and Internet experience of the participating pupils of the seventh and the eighth grade seemed to be very high.

In accordance with these results, there was a good starting level of declarative knowledge with regard to facts about the Internet and the WWW. As in Study 1, pupils’ prior declarative knowledge at ToM 1 with regard to search-irrelevant facts (70.5% correct, SD = 13.1) was quite good and better than their prior knowledge with regard to search-relevant facts (60.5% correct, SD = 14.6).

The high level of computer and Internet experience and the solid fact knowledge, however, apparently did not help pupils much in searching the web for information as the overall pretest performance on solving information problems strongly demonstrated (18.0% correct, SD = 13.0).

With regard to the first hypothesis on age related differences, we conducted a MANOVA for the between-subject variable seventh vs. eighth grade for the dependent variables declarative knowledge and search performance at the first ToM as well as at the last ToM of CISWEB (see Fig. 3). Contrary to hypothesis, this analysis yielded no significant effects of pupils grade on declarative knowledge (first ToM: $F(1,56) = 3.51; \ M_{se} = 103.58; p > .05$; last ToM: $F < 1$) or on search performance (both ToMs: $F < 1$).

Thus, with regard to both measures of achievement, there were neither differences between pupils of different grades at the beginning of the training nor at the end of the training. Accordingly, age differences cannot be made responsible for pupils’ fact knowledge with

![Fig. 3. Performance [in % correct] for the declarative knowledge test and solving information problems before and after studying CIS-WEB.](image-url)
regard to the WWW as well as for their performance in solving information problems. Contrary to our hypothesis, it cannot be assumed that a higher grade (or age) more or less automatically (i.e., without a specific training) leads to a more competent search for information in the WWW.

In line with the second hypothesis on declarative knowledge acquisition, a MANOVA for the within-subject variable time of measurement (ToM 1 to ToM 3) with regard to the dependent variables declarative knowledge for search-irrelevant and search-relevant facts (cf. Fig. 4) was conducted that revealed a significant influence of the time of measurement on knowledge of search-irrelevant ($F(2,114) = 29.88; M_{se} = 172.27; p < .001$) as well as on knowledge of search-relevant facts ($F(2,114) = 33.13; M_{se} = 158.31; p < .001$).

Most of the declarative knowledge with regard to search-irrelevant and search-relevant facts is conveyed in the first module of CIS-WEB. Accordingly, specific contrasts yielded a significant increase of the declarative knowledge from ToM 1 (directly before Module 1) and ToM 2 (directly after Module 1).

However, between ToM 2 (directly after Module 1) and ToM 3 (at the end of CISWEB) pupils obviously forgot some of the search-irrelevant facts conveyed so that there was even a decrease of declarative knowledge with regard to these facts ($F(1,57) = 8.99; M_{se} = 364.17; p < .01$). There was no difference between ToM 2 and ToM 3 with regard to search-relevant facts ($F(1,57) = 1.29; M_{se} = 302.00; p > .20$).

In sum, it could be shown – as expected – that CIS-WEB has a positive impact on declarative knowledge acquisition, both for search-irrelevant as well as for search-relevant facts.

In accordance with the third hypothesis on search performance improvement the results show that pupils’ search performance for all four sub-goal structures improves while using CIS-WEB. Additionally, information problems of more complex sub-goal structures resulted in an inferior overall search performance than information problems of simpler Sub-goal Structures. The respective MANOVA with the variables time of measurement (ToM 1 to ToM 4) and sub-goal structure of the information problems (SGS 1 to SGS 4) for the dependent variable search performance (cf. Fig. 5) revealed a significant influence of the time of measurement ($F(3,165) = 11.46; M_{se} = 869.85; p < .001$) and of the sub-goal structure of the information problems ($F(3,165) = 84.28; M_{se} = 382.10; p < .001$).

However, there was no significant interaction between time of measurement and sub-goal structure ($F(3,495) = 1.29; M_{se} = 450.02; p < .20$). This finding is contrary to the
assumed dependence of the time course of performance improvements on the complexity of the sub-goal structures of the information problems to be solved (Hypothesis 4). Thus, the assumption that search performance regarding information problems with more complex sub-goal structures only increases considerably in later phases of the web training could not be confirmed.

Specific contrasts with regard to the time of measurement revealed that there was a significant increase of overall search performance (summarized over the four different sub-goal structures) from ToM 1 (18.00% correct; SD = 13.01) to ToM 2 (28.52% correct; SD = 15.70; $F(1,55) = 18.81; Mse = 370.95; p < .001$). However, both increases of overall search performance from ToM 2 to ToM 3 (29.90% correct; SD = 17.50) and from ToM 3 to ToM 4 (32.90% correct; SD = 17.50) proved to be not statistically significant (both $Fs < 1$).

Additionally, specific contrast with regard to the four different sub-goal structures revealed that there was a significant decrease of search performance (summarized over the four different times of measurement, cf. Fig. 6) from SGS 1 to SGS 2 ($F(1,55) = 26.71$; $p < .001$).
$M_{se} = 156.21; p < .001$) and from SGS 2 to SGS 3 ($F(1, 55) = 56.32; M_{se} = 205.46; p < .001$) but not from SGS 3 to SGS 4 ($F(1, 55) = 1.68; M_{se} = 161.18; p > .20$). The latter finding might go back to the fact that the performance for SGS 3 already shows a substantial floor effect.

To sum up, Study 2 demonstrated that CIS-WEB does not only substantially improve pupils’ declarative knowledge with regard to search-irrelevant and search-relevant facts but also increases their performance on solving information problems of different sub-goal structures. The strongest training effects could be observed for Module 1 which introduces basic knowledge on the Internet, the WWW, and search systems on the web. In order to interpret this finding, it has to be noted, however, that Module 2 to Module 6 heavily rely on pupils engagement with worksheets and exercises. Therefore, learning success in these modules might have depended on pupils’ amount of invested mental effort in working on these exercises and worksheets.

In line with this reasoning, a correlation analysis reveals that there are substantial correlations between pupils’ learning intensity (their performance when filling in worksheets and exercises during the learning phase) and their later learning outcomes in terms of declarative knowledge and search performance (cf. Fig. 7). This pattern of positive correlations also reveals that the positive effect of CIS-WEB on pupils’ abilities for competent information search on the web cannot be interpreted as a simple time or training effect but depends on pupils’ activities. Thus, only pupils who were actively involved in working on Module 2 to Module 6 seem to benefit substantially from those later modules of the CIS-WEB training.

In conclusion, the results of Study 2 showed that CIS-WEB enhances pupils’ declarative knowledge about the web and their search performance far beyond conventional Internet trainings. First positive effects of CIS-WEB are clearly visible after Module 1 (i.e., after three lessons of instruction), whereas further improvements seem to heavily depend on pupils’ effort investment with regard to worksheets and exercises embedded in the training. Therefore, the web training CIS-WEB is to be considered as an effective instrument for the facilitation of pupils’ competent web-based search for information.

Nevertheless, based on the results of Study 2, several suggestions for improving the positive effects of CIS-WEB can be made. For example, the instructional methods used

![Fig. 7. Pattern of correlations between learning intensity during training and later performance for the declarative knowledge test and for solving information problems.](image-url)
in CIS-WEB could be optimized to ensure pupils’ active participation. Additionally, Module 6 that addresses the processing of the sub-goal “identification of sub-tasks” could be extended in order to improve the learning outcomes with regard to this sub-goal. However, the implementation of these suggestions would probably require a temporal extension of the web training beyond its current form.

7. Summary and discussion

The empirical studies reported in this paper demonstrate the necessity of supporting pupils in acquiring skills for the competent search for information in the WWW. Both studies reveal that pupils had great difficulties in their information search despite their substantial Internet experience and their solid declarative basic knowledge. Additionally, these difficulties existed regardless of pupils’ age or grade; thus, even in a long-term perspective it cannot be expected that aging automatically increase pupils’ ability for a competent information search. Accordingly, it can be claimed that a specialized training for information search on the web is an essential educational task in order to enable pupils to autonomously use a wide range of information resources and to face the requirements of a modern information society.

The results of Study 1 demonstrated that neither a technically oriented Internet training nor a free exploration of the web substantially increased pupils ability for competent information search or their declarative knowledge about the web. In contrast to this finding, Study 2 revealed that the web training CIS-WEB is effective to improve pupils’ web search skills and their declarative knowledge with regard to the Internet and the WWW. CISWEB is characterized by a user-oriented (instead of a technically oriented) focus and is based on two theoretical analyses: First, a conceptual analysis in the fields of media literacy research and information retrieval research that helped to identify the content aspects that have to be covered by CIS-WEB. Second, a task analysis of information problems on the web that allowed to derive a taxonomy of four different sub-goal structures of information problems and informed the overall structure of CIS-WEB.

In this final section we will discuss CIS-WEB with regard to its potential for further improvement, for integration in classroom settings, and with regard to the issue of combining its user-oriented approach with a more system-oriented approach in order to further optimize pupils’ abilities to use the WWW as an information resource.

The results of Study 2 clearly demonstrated the effectiveness of the web training CISWEB in a classroom context and simultaneously provided hints how the efficiency of the web training might be further improved. As could be demonstrated, Module 1 already works very efficient with a limited amount of time (three lessons). Module 2–6, however, heavily rely on pupils’ engagement with worksheets and exercises. Accordingly, there are substantial correlations between pupils’ learning intensity during these modules and their later learning outcomes in terms of declarative knowledge and search performance. Only when pupils invest a considerable amount of mental effort in the worksheets and exercises of Module 2–6 it can be expected that they improve their skills and knowledge beyond what they already know after studying Module 1. Thus, two possible options to increase the efficiency of CIS-WEB in a classroom context might be proposed.

One possible option might be to confine the training to Module 1 which takes only about three lessons and improves all performance parameters substantially. This option
would lead to a short term intervention that could be implemented instead of a conventional web training within a classroom setting.

A second and more far-reaching option might be to optimize the instructional methods employed in CIS-WEB in order to ensure that most of the pupils really get engaged with the contents conveyed in Module 2–6. This second option would probably imply to extend the training materials and to space the presentation of the materials over a longer period of time. For instance, it might be useful to switch from a setting where the materials are presented on three successive days with six lessons per day to a setting where pupils attend a course for two lessons a week, resulting in a course that might take several weeks. This spacing of the presentation of the materials might help to avoid that pupils get tired or overwhelmed, and might therefore help to increase the effort that pupils are willing to invest in worksheets and exercises. Additionally, other instructional devices like group work and small problem-based projects might be included in the course in order to encourage more student activities and in order to enrich the course with more external information sources.

It would also be promising to interleaf the teaching of web search competencies by means of CIS-WEB with activities related to other school subjects so that pupils get the opportunity to apply their recently acquired skills to their academic work. In order to implement this idea it might be useful to integrate the training modules of CIS-WEB with an interval of several weeks between the individual modules into the regular classes of other subjects.

Another fruitful line of development would probably be to broaden the target group of the web training from pupils to other populations that are equally in need to acquire basic skills and competencies in information search in the WWW. One example of such a user group are seniors who increasingly use the Internet and who might have some needs for systematic training on web search competencies. Adapting CIS-WEB to a different user population probably would not necessitate a change in the generic design of the training. However, the concrete contents of the training in terms of examples, exercises, and type of language used might need to be adjusted to other populations.

A final issue that will be raised in this discussion concerns the prospects of combining the user-oriented approach of CIS-WEB with a more system-oriented approach in order to further optimize pupils’ abilities for using the WWW as an information resource. Stating that CIS-WEB is based on an user-oriented approach means that this training intends to directly improve the competencies of users to interact with information and to handle existing technical systems for accessing this information. An alternative approach that can be described as system-oriented would aim at facilitating a user’s interaction with the information technology indirectly by means of technical tools and innovations. A typical support measure from a system-oriented perspective would be to improve the design of search systems in order to optimize their usability. For instance, search systems may be improved in that they support users in formulating search keywords or in choosing links from a query result list (cf. the suggestions made by Hölscher (2000)). System-oriented and user-oriented approaches for improving web-search competencies have a differential pattern of characteristics that is illustrated in Table 2.

According to these characteristics, system-oriented and user-oriented approaches might have complementary potentials and problems in facilitating a competent search for information in the WWW. The biggest advantages of system-oriented approaches probably lie in their potential for dissemination, the expectation of short- to medium-term effects, and
the low input needed from users’ sides. User-oriented approaches, on the other hand, have
the advantages of a higher durability of acquired competencies and the potential for the
generalization of the contents learned. A quite natural consequence of this differential pattern
of advantages and disadvantages might be to combine the user-oriented approach of
CIS-WEB with a more system-oriented approach in order to optimize pupils’ abilities for
using the WWW as an information resource. In order to implement this idea it is on the
one hand necessary that information technologies are developed that support users in their
information searches (e.g., by providing spelling corrections for search keywords, by giving
suggestions for the formulation of search keywords, by supporting the selection of
links etc.). On the other hand, training efforts to develop competences on the side of the
users are a necessary complement in order to ensure that users do not become completely
dependent on technical support tools. The latter point is of particular importance because
current information technologies are still not able to substitute elaborated cognitive pro-
cesses needed by users for solving complex information problems. For instance, up to now
technological tools are unable to make well-founded decisions about the relevance
of information in the context of specific information problems or about the credibility
of information – beyond a simple matching of terms or an analysis of the frequency of
use. Therefore, a combination of system-oriented and user-oriented approaches seem to
be a promising path for the future to provide the best possible support for a competent
search for information on the web.

| Table 2 |
| - Typical characteristics of system-oriented and user-oriented approaches |

<table>
<thead>
<tr>
<th></th>
<th>System-oriented approach</th>
<th>User-oriented approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>Starting point</td>
<td>Information technology</td>
<td>User competences</td>
</tr>
<tr>
<td>Intervention</td>
<td>Technical innovation</td>
<td>Training</td>
</tr>
<tr>
<td>Dissemination</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>Effect</td>
<td>Short- to medium-term</td>
<td>Medium- to long-term</td>
</tr>
<tr>
<td>User input</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>Durability</td>
<td>Undetermined</td>
<td>High</td>
</tr>
<tr>
<td>Potential for generalization</td>
<td>Low</td>
<td>High</td>
</tr>
</tbody>
</table>

References

searching based on identifiable document partitions. *Information Processing and Management, 38*, 293–304.
Clark, A. (2002). Global abductive inference and authoritative sources, or, how search engines can save cognitive
Clearinghouse on Information Resources.
Eisenberg, M. B., & Berkowitz, R. E. (1996). *Helping with homework: A parent’s guide to information problem-


