

## Fostering students' evaluation behaviour while searching the internet

Amber Walraven · Saskia Brand-Gruwel · Henny P. A. Boshuizen

Received: 12 November 2010 / Accepted: 19 March 2012 / Published online: 11 April 2012  
© The Author(s) 2012. This article is published with open access at Springerlink.com

**Abstract** A program for teaching 9th graders course content (history) and how they should evaluate information found on the WWW was designed and tested. Goal of the program was to teach content, evaluations skills, and to achieve transfer of these skills to a different domain. In the design of the program the principles of two transfer theories were combined using a design-based research approach. Results revealed that the program improved students' evaluation behaviour. Compared to the students of the control condition, the evaluation skills of students in the experimental condition improved to a higher level, but did not lead to transfer. Students in the experimental classes scored higher on the final content exam than students in the control class.

**Keywords** Information · Evaluation · Internet · Secondary education · Information skills · Teaching

In our current technological society it is often claimed that our primary and secondary students are living their lives immersed in technology, 'surrounded by and using computers, videogames, digital music players, video cams, cell phones, and all the other toys and tools of the digital age' (Prensky 2001, p.1). However, more and more is known about these so called 'digital natives' and research shows that the two key claims about them (that this distinct generation actually exists and that education should fundamentally change to service them) 'are based on fundamental assumptions with weak empirical and theoretical foundations' (Bennet et al. 2008, p. 777). One of these assumptions is that since they grew up with the Internet, students in secondary education are very capable of working with it. The World Wide Web (WWW) is often the only source of information secondary education students use for accomplishing school assignments (Beljaarts 2006;

---

A. Walraven (✉)  
Faculty of Behavioural Sciences, Curriculum Design & Educational Innovation, University of Twente,  
P.O. Box 217, 7500 AE Enschede, The Netherlands  
e-mail: a.walraven@utwente.nl

A. Walraven · S. Brand-Gruwel · H. P. A. Boshuizen  
Open University of the Netherlands, Heerlen, The Netherlands

Jones 2002). However, ‘working with it’ often refers to the technicalities of the Web, i.e., controlling the mouse, operating a browser. Most students lack the skills to successfully solve information problems such as being able to identify one’s information needs and translate them in questions, to understand a text, to assess the trustworthiness and relevance of a piece of information, to adapt one’s strategy to the results of the search process. So, despite their frequent use of the WWW, students’ search methods are far from ideal; most students do not evaluate their search results, the information they have found and the source of this information (Fidel et al. 1999; Hirsch 1999; Kafai and Bates 1997; Koot and Hoveijn 2005; Lorenzen 2002; Lyons et al. 1997; MaKinster et al. 2002; Wallace et al. 2000; Walraven et al. 2009). Evaluating what one has found on the WWW is crucial, since the WWW lacks centralized control and regulation and its contents can easily be altered (Metzger et al. 2003).

Teachers recognize that students show problematic ‘cut-and-paste’ behaviour when searching the Internet for information to write an essay. Students search, find some information (reliable or not), cut and paste it in a document and hand it over as an essay to the teacher. Teachers agree that instruction in evaluating information found on the Internet is needed, that students need support to use different criteria to evaluate web information, and that instruction must be implemented in domain-specific programs. Unfortunately, educational programs do not give so much attention to stimulating students to evaluate information and to the way a teacher could guide his students on the WWW.

In this article the results of a design experiment conducted in cooperation with four 9th grade teachers aiming at improving students’ use of evaluation skills when searching the WWW will be discussed. In the introduction, first, the grounds for taking evaluation of information on the World Wide Web as the main focus of instruction will be discussed. Second, the design cycle will be described, in which the focus will be on the final design cycle. In this part also the didactical approach based on two transfer theories will be addressed. Finally, the research questions will be presented.

Evaluation of results (the hit list), information and source (the information on the website and the website itself) is part of the larger process of information problem solving (IPS). The process consists of the constituent skills *defining the information problem* (i.e., reading the task, activating prior knowledge), *searching information* (i.e., choose search strategy, specify search terms, evaluate search results), *scanning information* (i.e., read information global, evaluate information and source, elaborate on content), *processing information* (i.e., read in depth, evaluate information and source, store relevant information, elaborate on content) and *organising and presenting the information* (i.e., structure relevant information, realize the product) (Brand-Gruwel et al. 2005). There are three evaluation moments in this process: evaluating search results during the searching phase, evaluating information and source during the scanning and processing phase. Several criteria can be used to evaluate search results, information and source (Barker 2005; Beck 1997; Boekhorst 2000; Kirk 1996; Ormondroyd 2004). Evaluating *search results* will answer the question: Which site am I going to open? Criteria that can be used for evaluating search results are (1) relevance of the content (given by title and description), (2) kind of source (pdf, word-doc etc.), (3) reputation (given by the URL, .com, .org etc.), (4) rank in hit list, (5) familiarity (knowing the URL or organization behind the site) and (6) familiarity with the language (based on the title and description). After evaluating the results and opening a website or file, the *information* on that site or in the file has to be evaluated. This can be done with several criteria, grouped in three categories: usability, verifiability and reliability. Language, connection to task, author, references, kind of information and objectivity are examples of criteria to evaluate information. The *source*

can be evaluated on technical, usability, verifiability and reliability grounds, like speed, appearance, audience and reputation. Using these criteria and thus evaluating results, information and source can help avoiding the use of incomplete, false and biased information. The criteria used in this study can be found in the Appendix Table 4.

From an educational point of view, lack of information skills and a non-critical attitude towards information on the WWW can result in reports and learning that lack quality (Britt and Aglinskas 2002), since students often cut and paste information without evaluating it (Grimes and Boening 2001; Rothenberg 1998). This problematic behaviour does not disappear with age, information seeking that focuses on quick and easy results, without sophisticated searching skills is also shown by undergraduates and college students (Maughan 2001; O'Hanlon 2002). Instruction in information skills seems necessary. Unfortunately, although the importance of instruction in an effective and critical use of the WWW has been recognized for several years, instruction in information skills is rare, not always effective and transfer is often not measured (Walraven et al. 2008).

If students are to become critical users of the WWW and use that ability throughout their lives, it is important that they can use their evaluation skills in multiple contexts and various courses. Instruction in these skills should therefore be aimed at transfer. Transfer can be fostered in several ways (e.g., Gick and Holyoak 1983; Thorndike and Woodworth 1901; Wertheimer 1961).

A widely adopted view of transfer suggests that the "induction or construction of abstract rules and schemata, or other mental representations has been hypothesized to serve as the primary cognitive support for knowledge transfer" (Wagner 2006, p. 2). An important theory in this view is the high road to transfer of Perkins and Salomon (1989), Salomon and Perkins (1989).

According to this theory students have to be stimulated to pay explicit attention to the various steps that have to be taken in a process and to the way these steps can be used flexibly in different situations. The high road to transfer depends on mindful abstraction from the context of learning. It is 'the deliberate, usually metacognitively guided and effortful, de-contextualization of a principle, main idea, strategy, or procedure, which then becomes a candidate for transfer' (Salomon and Perkins 1989, p. 126). The conscious formulation of abstraction means answering questions like: what is the general pattern? What is needed? Which step can I take now? What rules or principles might apply? Abstracting is closely related to metacognitive skills like planning (what am I going to do), monitoring (is the process going according to plan?) and evaluating (what have I learned that I can use a next time?), thus high-road transfer can be fostered by stimulating a persons' metacognitive skills. The high-road transfer can be forward or backward reaching, with the present problem as point of reference. With forward reaching one abstracts situations from the current context to a potential transfer context. An example of forward reaching transfer is a child learning good study habits by setting aside a definite time for certain activities and sticking to it. The child actually schedules his or her activities in this way. When the child grows up and gets a busy job he or she still schedules priority projects in this way, so progress on that project is assured, no matter what happens. The principle (setting definite times) is so well learned that it simply suggests itself appropriately on later occasions (Salomon and Perkins 1989). With backward reaching one abstracts in the transfer context, looking for features of the previous problem where new skills and knowledge were learned. And example of backward reaching transfer is having learned as a child to count to 10 when you felt you were losing your temper. Now, as an adult you notice that you are an impulse buyer and you want to find a way to inhibit your

impulsiveness. You should try to hold back. When you think of remedies, the count to 10 strategy occurs to you. You try it out, and it helps.

Perkins and Salomon (1989) state that high-road transfer is important for skills that call upon strategic knowledge, like thinking skills and problem solving skills. Evaluating results, information and source when searching for information on Internet requires strategic knowledge, since it is part of the heuristic information-problem solving process. The basic assumptions of this transfer theory (conscious formulation of abstraction and stimulating metacognitive skills) could therefore be used to design instruction that fosters the transfer of evaluation skills. Instructional design based on this transfer theory should pay particular attention to strategy explication, emphasizing abstraction and de-contextualization. This means for the skills of evaluation results, information and sources that students should know the steps to be taken, strategies that can be used in the problem solving process, and how to regulate this process.

A contrasting theory on transfer does not focus on the steps of the problem solving process (by abstraction and metacognition) but emphasizes the importance of a good, extensive and well organized knowledge base and the domain specific interpretation of the skills (Larkin et al. 1980; Perkins and Salomon 1989; Simons et al. 2000). This theory, which we will call the rich representation theory, is based on the way that experts organize their knowledge and tend to go about solving problems. An expert's extensive knowledge base includes three representations of the information: conceptual, episodic, and action representations.

Conceptual representations refer to concepts and principles with their defining characteristics (like a cat is an animal with whiskers and a tail). Episodic representations refer to personal experiences with instances of concepts and principles (like I loved the cat I had when I was a child). Action representations refer to the things one can do with the conceptual and episodic information, i.e., using that knowledge to solve a problem (like cats can be kept as a pet). When the three representations have many and strong relationships with each other (e.g., conceptual representations have a relation with concrete experiences) and with representations in other domains, the knowledge base has a high degree of connectedness. These connected, rich representations will make learning outcomes durable, flexible and generalizable. Knowledge and skills 'are not restricted to one context but reach out to other contexts and situations.' (Simons et al. 2000, p. 2), thereby fostering transfer.

For evaluation skills this would mean that students should have deep knowledge of concepts associated with the key concept evaluation. The instruction based on this theory should stimulate students to construct a well structured representation of the criteria to evaluate search results, information and source that can be used in different situations and while solving different tasks. Moreover, students must become aware of the usefulness of the criteria and they should experience that the use of the criteria helps to become critical web searchers. This experience makes the representation of criteria better anchored. In the design of the present research these two theories are used to design an educational program to foster students' evaluation behaviour when searching for information in the WWW.

The program was designed by four secondary education history teachers and one researcher. Teacher A was a very experienced web user and maintained several websites, teacher B had less experience on the Web. Both teachers (from one school) participated in the study because they acknowledged the importance of teaching students how to evaluate information on the WWW. Teacher C did not have a lot of experience with ICT. Teacher D was an experienced web user and a teacher who liked integrating ICT in his lessons. Main reasons for these two teachers, who were from two other schools for secondary education, to participate were acknowledging the need for students to learn how to evaluate

information and, for the first teacher, wanting to learn more about ICT and the Internet. The researcher (R) was an educational technologist and instructional designer with an interest in ICT, and expertise in evaluation of information on the WWW.

In a previous experiment, the team designed two programs, each based on one transfer theory. For the design of a program based on the high road program (the transfer theory of Perkins and Salomon 1989) process worksheets were developed to inform students about the steps to be taken and to support them to reflect on their search process. The guidance given in these sheets faded from high to low. In the rich representation program (based on the transfer theory of Simons et al. 2000) a mind mapping technique with discussion sessions to build a knowledge structure was used. Furthermore in both programs whole tasks were used with a certain kind of variability. That is, tasks requiring from students not only to evaluate information, but to define the problem, search and select information and come up with a product (e.g., essay, role-play, comic). Theme of the lessons for the 9th graders was World War 2, ranging from the Treaty of Versailles at the end of the First World War, to the end of the 2nd World War.

So, the goal of the first experiment was to design two programs, each based on a different transfer theory (high road and rich representation) to teach 9th grade of pre university students (during history classes) how to evaluate search results, information and source when searching for information on the WWW, using different kind of tasks and stimulating students to use these skills in a variety of settings, that is foster the transfer of the skill.

A quasi experimental study was conducted to test and evaluate the effect of both programs on students' evaluation behaviour, that is use of criteria for evaluating results, information and source when solving information problems on the WWW and if the effects of both programs based on two transfer theories (high road versus rich representation) differed in terms of transfer achieved. The teachers gave the lessons they had developed themselves. Each teacher taught one class. So two classes received the lessons based on the high road transfer theory, and two classes received the lessons based on the rich representation transfer theory. Before and after the intervention students were asked to evaluate a hit list, and websites and information in the domain of history and in the domain of biology (transfer tests). Furthermore, 11 students accomplished two tasks (one history, one biology) while thinking aloud during pre and post test. Each class was observed three times by R. During these observations teacher-student interactions, actions of students and actions of teachers were written down by the observer. Researcher and teachers had regular email contact on how the lessons were carried out by the teachers and how they were received by the students.

To determine the effects of the programs on students' use of criteria for evaluating results (hit list), information and source in the domain of instruction (history) repeated measures ANOVA analyses with program as between factor were performed. With regard to evaluating results, there was no significant main effect on the factor 'time',  $F(1,82) = 0.99$ ,  $MSE = 1.20$ , ns, and a significant main effect for 'program',  $F(1,82) = 3.38$ ,  $MSE = 6.40$ ,  $p = 0.05$ ,  $\eta^2 = 0.05$ . The high road students scored higher overall. A marginal interaction effect between 'time' and 'program' was found,  $F(1,82) = 3.07$ ,  $MSE = 3.70$ ,  $p = 0.08$ ,  $\eta^2 = 0.04$ . Because both programs were implemented in two different classes, it was determined if class effects occurred within conditions. No significant class effects were found.

With regard to evaluating information and source, a marginal main effect was found for 'time',  $F(1,82) = 3.65$ ,  $MSE = 217.284$ ,  $p = 0.06$ ,  $\eta^2 = 0.04$ . That is students in both programs slightly improved their evaluation scores. A significant main effect for 'program'

was found,  $F(1,82) = 11.07$ ,  $MSE = 1325.64$ ,  $p = 0.00$ ,  $\eta^2 = 0.11$ . The rich representation condition scored higher overall. No significant interaction effect between 'time' and 'program' was found,  $F(1,82) = 2.13$ ,  $MSE = 126.78$ , ns.

Again it was determined if there were class effects within the conditions, because each condition existed of two classes. No significant difference between classes was found in the high road program. Within the rich representation condition a significant difference between classes was found,  $F(1,43) = 7.03$ ,  $MSE = 357.33$ ,  $p = 0.01$ ,  $\eta^2 = 0.14$ .

To determine the effects of the programs on students' use of criteria for evaluating results, information and source on the transfer task repeated measures ANOVA analyses with program as between factor were performed. With regard to evaluating results, no significant main effect on 'time' was found,  $F(1,82) = 0.40$ ,  $MSE = 0.37$ , ns, and also no main effect on the factor 'program',  $F(1,82) = 0.02$ ,  $MSE = 0.02$ , ns. However, a significant interaction effect between 'time' and 'program' was found,  $F(1,82) = 4.11$ ,  $MSE = 3.57$ ,  $p = 0.05$ ,  $\eta^2 = 0.05$ . The scores of the rich representation condition increased while the scores of the high road condition decreased. Furthermore, no class effects were found within the two conditions.

With regard to the evaluation of information and source a significant main effect on the factor 'time',  $F(1,82) = 5.79$ ,  $MSE = 468.34$ ,  $p = 0.02$ ,  $\eta^2 = 0.07$ , was found. This means that both programs had a positive effect on students' evaluation behaviour. No main effect was found for the factor 'program',  $F(1,82) = 0.96$ ,  $MSE = 130.67$ , ns, and also no interaction between 'time' and 'program' was found,  $F(1,82) = 0.21$ ,  $MSE = 16.58$ , ns. No significant difference between classes was found in the high road program. Within the rich representation condition a significant difference between classes was found,  $F(1,43) = 3.82$ ,  $MSE = 289.54$ ,  $p = 0.06$ ,  $\eta^2 = 0.08$ . For a more elaborate description of the study and its results we refer to Walraven et al. (2010).

After this first design cycle, a new cycle was started. The design team remained the same, as well as the target group, 9th graders. The goal of the second design cycle was to develop one new instructional program, based on the lessons learnt from the first design cycle. It was decided that the good aspects of both programs should be integrated in the new program. In the high road program, the strong points identified were the variety in tasks and the focus on the entire search process with help of process work sheets. The process worksheets were also a weak point, since they were too extensive and students rebelled against filling them out. Strong points of the rich representation program were the focus on one evaluation criterion per lesson and the discussions with students about criteria. A weak point was that not enough attention had been paid to the use of the criteria in various contexts or the way criteria are connected. Teachers also noted that it was very important to convince the students of the importance and significance of evaluating information, and to make sure that the students did not see evaluation of information as just another task during these lessons, but as a skill they should use in more courses in school. It was decided that the new program would use less extensive process worksheets, better structured discussions and the program would start with a confrontation with the importance of evaluating information. The teachers decided on the content of the lessons and divided five themes between them. They agreed to design three lessons per theme; two lessons with WWW and evaluation assignments, and one classical lesson. This decision was made based on the experience with the previous programs; students had complained that they spent too much time behind the computer and longed for a normal lesson. It was decided that R would develop adjusted process worksheets and a detailed teacher manual for the discussions on criteria in order to develop a rich knowledge structure on criteria. During two group sessions, teachers presented their designed lessons and R presented the

new worksheets and the manual. The materials were discussed and adjusted if the group felt this was necessary. A description of the designed program will be presented in the “[Method](#)” section.

The redesigned program was tested and evaluated using a pre test post test control group design. The evaluation of this program will be described in the next sections. The questions addressed are: (1) What are the effects of the program on students’ evaluation behaviour, that is, how do students use criteria for evaluating results, information and source, when solving information problems on the WWW? (2) Does instruction based on a combination of two transfer theories lead to transfer? (3) What is the effect of the program on learning results?

## Method

### Participants

Five 9th grade classes (101 students, age 14–15) of three different secondary schools participated in this study. Four classes received an educational program and one class served as a control class. The four teachers who designed the program taught the four experimental classes.

### Program

#### *Goal and overview of the lessons*

The goal of the program was to teach 9th grade students how to evaluate search results, information and source when searching for information on the WWW in a historical context, using different kinds of tasks and stimulating students to use these skills in a variety of settings, that is foster the transfer of the skill. The general subject of the program was World War II and it consisted of 15 lessons of 50 min. [Table 1](#) gives an overview of the lessons.

As stated earlier, based on experiences with the previous programs, it was decided that the new program would use less extensive process worksheets, better structured discussions and the program would start with a confrontation with the importance of evaluating information. So, the first lesson was a confrontation lesson. Students were confronted with the importance of evaluating results, information and source. Students had to answer four questions about the treaty of Versailles. They could only use seven sites provided by the teacher. These sites contained contradictory information, information on a different treaty of Versailles, newspaper articles etc. For instance, the first site was a Wikipedia site about the treaty of Versailles of 1783. This was not the treaty the assignment was about. The second site was a newspaper article about the correct treaty of 1919, but with different information than the third site, with an unknown source. Goal of this lesson was confronting students with incorrect, false, and biased information and having them think about the importance of evaluating information. After this confrontation lesson, three lessons on the treaty of Versailles, followed by two lessons on Weimar and art, three lessons on the rise of Hitler, three lessons on chronology/1938, one on daily life in the war, one on persecution of Jews, and finally a lesson about the war in the region of the school.

**Table 1** The 'Evaluation of Internet information'—program

Lesson	Theme	Content
1	Confrontation	Students answer four questions about the treaty of Versailles, using seven websites provided by the teacher. These sites contain contradictory information, information on a different treaty of Versailles, newspaper articles etc. After 30 min a discussion is held on what students noticed about the sites and criteria for evaluating results, information and source
2–4	Versailles	Students act out the negotiations of the Treaty of Versailles. The class is divided in groups and each group searches for information on the Treaty and the viewpoints of the main characters of these negotiations (France, US etc.) When enough information has been collected students act out the negotiations. Students receive a process worksheet that focuses on the first step of the process: define the problem
5–6	Weimar and art	Students read a text on art in Germany at the beginning of the War. And choose a subject to write an article on art and war. The process worksheet focuses on the second step: search information, with special attention paid to the evaluation of results
7–9	The rise of Hitler in 1933	Students draw a comic on the rise of Hitler. They search information about this rise, write a scenario for the comic and draw the comic. Instead of drawing they can also use pictures they find. The process focuses on the third step in the process, scan information, with special attention paid to the evaluation of information and source
10–11	Chronology	Students play a card game (happy families/old maid/kwartet (Dutch)) Each set contains four events during a certain year (e.g., 1939: Occupation of the Czech republic, Molotov Ribbentrop pact, invasion of Poland, Russians occupy east of Poland). Students play the game and organize the events according to chronology
12	1938	Students act out the convention of Munich, in the same way they acted out the negotiations on the treaty of Versailles. Process worksheets focuses on evaluation of information and source
13	Daily life in the war: collaboration and resistance	Students write an interview with a Dutch collaborator or a hero from the resistance. They base their questions and answers on information they find on the WWW
14	Persecution of the Jews	Normal lessons with textbook or movie
15	The war in our own region	Students make a picture of a war monument in their home town and write an article about it

### Reader

Students received a reader on information problem-solving and how to evaluate search results, sources and information. This reader was based on the skills decomposition of the Information-Problem Solving skill by Brand-Gruwel et al. (2005), described the necessary phases for information-problem solving (define the problem, search information, scan information, process information and organize and present information) and the steps per phase (e.g., in the search information phase the steps are: select search strategy, define search terms, and evaluate search results). It also provided information on how and why the phases and steps should be taken and also provided rules of thumb concerning evaluation



criteria. In the scan information phase during the step evaluate information and source, students were given hints like: check if you can see when the site was last updated.

### *Process worksheet*

Students received a process worksheet with the assignment for the coming lessons and some questions they had to answer at the start of each theme (Versailles, Weimar and art, the rise of Hitler and chronology/1938). The questions on the sheets were linked to the phases of information-problem solving and corresponded with the phases and steps in the reader. In the first three lessons the first phase (define the information problem) was focused on, the next three lessons the scanning information phase, and so on. So, instead of filling out questions for each phase like on the work sheets in the previous high road program, students only filled out question from one phase of the information problem solving process. To make the entire problem solving process visible to the students, and to point out which phase of the process was focussed on, every worksheet started with a figure of the process solving process with the central phase highlighted. Figure 1 presents an example of a process worksheet.

### *Discussion*

At the end of every theme teachers and students had a discussion on evaluation criteria. The goal of these discussions was to develop a mind map or knowledge structure. Teachers received a manual for these discussions, based on the theory of Klausmeijer (1990). According to Klausmeijer several steps have to be taken when teaching concepts: (1) Providing an orientating instruction, by focussing students attention, pointing out the importance of the concepts to be learned, helping students developing a scheme of the concepts, providing students with a strategy to learn the concepts and creating the intention to learn the concepts. (2) Providing or eliciting a definition or defining characteristics. (3) Recalling what was learned from long term memory. (4) Using examples and non-examples. (5) Helping students with discovering defining characteristics. (6) Providing strategies for distinguishing examples and non-examples. (7) Giving feedback.

The manual prescribed how teachers should structure the discussions on evaluation criteria. The first discussion would start with telling the students that during the next 15 lessons they would learn more about criteria to evaluate information on the WWW. Next students were asked why they thought it is important to learn about criteria, and the teachers explained why he or she finds it important that students learn about criteria. The teacher then asked students which criteria they knew and used for evaluating information. The teachers drew these criteria on the blackboard and created a knowledge structure together with the students. When a student mentioned a criterion, the teacher asked for a definition of the criterion or characteristics of the criterion (i.e., a student mentions objectivity, and the teacher asks how can I tell if information is objective?). Students copied the knowledge structure in their notebooks. The discussion ended with the teacher stating that in the next 15 lessons they will do several assignments on the Internet, and that he will return to this knowledge structure every three lessons, and that he expects that they can enrich the structure together and that students are able to explain the criteria and know how they can use them more and more. This first discussion followed steps 1 and 2 of the theory of Klausmeijer (1990).

The discussions after this first discussion always followed the same routine: (1) Take a look at the knowledge structure, and let students summarize what they already know about

## Opdracht Vrede van Versailles

Tijdens deze cyclus van drie lessen gaan jullie de besprekingen van Versailles, die de afronding waren van de Eerste Wereldoorlog, naspelen. Deze besprekingen vonden plaats in de eerste helft van 1919.

De klas wordt vooraf door de docent in twee groepen verdeeld, binnen elke groep spelen vier leerlingen de rol van een land. Dit zijn de belangrijkste landen die aanwezig waren bij de vredesonderhandelingen en de ondertekening van het verdrag van Versailles. De rest van de subgroep ondersteunt bij het voorbereiden van het spel.

Allereerst worden per groep de rollen verdeeld. Daarvoor moeten jullie eerst gaan uitzoeken welke hoofdrolspelers er waren in de onderhandelingen. Gebruik hiervoor je lesboek of het internet. Tijdens deze opdracht maak je gebruik van het proces voor het oplossen van informatieproblemen uit je theorieboek. Hieronder vind je het schema van dat proces.



Fig. 1 Process worksheet

Tijdens deze opdracht staat de fase ‘definieer het probleem’ centraal. Op deze lesbrief vul je de stappen in die je neemt in die fase. Straks wordt niet alleen het eindresultaat beoordeeld, maar vooral het proces, de manier waarop je tot je resultaat komt. Vergeet niet om voor je zelf alle stappen uit het proces te nemen, ook degene die je niet op dit formulier hoeft in te vullen!!

### Lees de opdracht

Lees de opdracht goed door, en bepaal wie welk land straks in het rollenspel vertegenwoordigt.

### Formuleer een hoofdvraag

Welke vraag moet je beantwoorden? Kijk nog eens op pagina 10 en 11 van het theorieboek informatieproblemen oplossen. Maak zo nodig een mindmap om je gedachten te ordenen. Je mindmap kun je hieronder tekenen.

Schrijf je hoofdvraag en eventuele deelvragen hier onder op:

.....

.....

.....

.....

### Stel vast wat je al weet van het onderwerp

Weet je al iets over de Vrede van Versailles, en de rollen van de verschillende landen? Schrijf dat hier onder op. Je mag ook een mindmap maken.

### Fig. 1 continued

criteria (e.g., step 3 of Klausmeijer); (2) Ask if students have new additions to the knowledge structure. This could be new criteria or additions to definitions of criteria); (3) Try to provide examples and non-examples of the criteria. (So examples of an objective website, but also an example of a very subjective website) and discuss the differences and similarities between the examples (Klausmeijer step 4, 5 and 6), and (4) Discuss whether the criteria are equally important for every research question, or course.

### Measurements

#### *Evaluation hit list*

To measure how students evaluated a hit list four information problems with a manufactured hit list of 14 results on paper were developed. Two tasks were in the domain of history (domain of instruction) and two in the domain of biology (transfer domain). The topics of the history tasks were ‘Anastasia Romanov’ and ‘the Watergate affair’, and the

topics of the biology tasks were ‘Super Size Me’ and ‘influence of sex before a sports match’. Per hit list, students had to select three sites they would open and three sites they would not open. They could highlight and circle parts of the results they based their decision on. Participants received a point per website if their evaluation was correct. That is, a point for choosing an appropriate site they wanted to open and a point for choosing not to open an inappropriate site. Maximum score was six points per hit list.

### *Evaluation of websites and information*

To measure how students evaluated websites and information four information problems and booklets of eight printed websites each were developed. Two tasks were in the domain of history (domain of instruction) and two in the domain of biology (transfer domain). The first history information problem regarded whether the Bush administration was behind the attacks of 9/11, and the second regarded whether the NASA was responsible for the first landing on the moon. The biology tasks involved whether the Dutch non-smoking policy was effective enough and whether or not teenagers were more often infected with sexually transmitted diseases. Students were asked which sites and what information they would or would not use, given the problem provided and were informed that it was not impossible to find 5–10 features they based their decision on. They could highlight those features. If students had circled a certain area on the site or written down a comment like: “Site is old” they received a point for recognizing the criterion, and if their evaluation was correct (e.g., the site was indeed old) they received another point. So, student could receive points for recognizing a criterion and for using a criterion in the correct way. Maximum score on all tasks was 200.

### *Learning results*

To determine students’ history knowledge a final exam about the topic 2nd World War was developed by one teacher. This test consisted of 10 content related open questions. Two of the questions also paid attention to information evaluation. A correction model was provided to the teachers in order to score the exams of their class in a correct way. Scores could range for 0–10

### *Field notes*

In each experimental class three lessons were observed. Field notes of these observations served as secondary material for possible explanations of the results. During these observations special attention was given to the interaction between the students and between the students and their teacher concerning evaluation behaviour and the use of evaluation criteria.

### *Design and procedure*

A pre test post test control group-design was used to determine the effect of the program on students’ evaluation behaviour (e.g., evaluation of hit list, websites and information). Table 2 presents the design of the experiment.

Before the first lesson, all students (experimental and control condition) made the hit list and website evaluation tasks (one history and one biology). These tasks were

**Table 2** Design of the Study

O1	X1	O2	N = 80
O1	X2	O2	N = 21

O1 = Two tasks evaluation hit list (history and biology), two task evaluation information and source (history and biology)

X1 = Intervention program (three observations per class)

X2 = Regular lessons on the history content

O2 = Two tasks evaluation hit list (history and biology), two task evaluation information and source (history and biology), final exam, reflective reports

**Table 3** Means and Standard Deviations of Hit List and Website Evaluation Task Score

	Experimental condition (N = 80)		Control condition (N = 21)	
	Pre test	Post test	Pre test	Post test
Hit list history	3.2 (2.2)	3.2 (2.3)	3.8 (1.7)	1.3 (2.0)
Websites history	14.9 (8.8)	18.8 (9.5)	14.8 (7.7)	12.8 (7.6)
Hit list biology	3.1 (2.1)	3.4 (2.3)	4.2 (1.4)	1.6 (2.2)
Websites biology	16.5 (9.0)	18.4 (8.4)	14.2 (9.4)	15.2 (9.2)

counterbalanced and rotated for the pre and post test. Half of the students received the first history tasks (hit list and website) and the first biology tasks (hit list and websites) during the pre test, and the remaining half received the second history tasks and the second biology tasks. Furthermore, half of the students started with the history task, and the other half started with the biology task. For completing the tasks students got a maximum time of 50 min. After the pre-test the experimental classes received the designed program and the control class received regular lessons on the 2nd World War. In each experimental class three lessons were observed by the first author. The observer wrote down the actions of teacher and students (e.g., 'Teacher starts with explaining the assignment. After 10 min students start to work on the assignment.'). Extra attention was given to interactions between teacher and students when talking about evaluation of information and sites. Questions students asked, remarks the teachers made about this topic when helping a student. The field notes became short narrative summaries of the lessons. A week after the last lesson all students completed the parallel forms of evaluation tasks (different information problem). During the post test students received a different task than the pre test task. Students who had made task 1, now made task 2 and vice versa. Again, the order of tasks (starting with history or biology) differed between students.

## Results

### Evaluation tasks hit list and websites

Table 3 provides the means and standard deviations of hit list and website evaluation task score. Scores are provided for the history tasks and biology tasks. The latter are the transfer tasks.

### *Effects of the instruction*

To answer the first research question on the effects of instruction on students' use of criteria for evaluating results (hit list) a repeated measures ANOVA analysis with condition as between factor was performed for the history task (the domain of the instruction). The analysis on the evaluation of the *history hit list* showed a significant main effect of 'time',  $F(1,99) = 14.60$ ,  $MSE = 49.56$ ,  $p = 0.00$ ,  $\eta^2 = 0.13$ , and no main effect for 'condition',  $F(1,99) = 2.21$ ,  $MSE = 13.36$ , ns. There was a significant interaction effect between 'time' and 'condition',  $F(1,99) = 14.30$ ,  $MSE = 48.55$ ,  $p = 0.00$ ,  $\eta^2 = 0.13$ . This interaction effect is not caused by an increase in scores of the experimental classes, but by a decrease in scores in the control class. The scores in the experimental classes remain constant. No class effects were found between experimental classes.

The second part of the research question focuses on effects of instruction on students' use of criteria for evaluating information and source (website). Again, a repeated measures ANOVA was performed with the results on the *history websites evaluation task* and condition as between factor. No main effect of 'time' was found,  $F(1,99) = 0.73$ ,  $MSE = 32.36$ ,  $p = 0.40$ ,  $\eta^2 = 0.01$ , and also no main effect for 'condition' was established,  $F(1,99) = 2.73$ ,  $MSE = 307.30$ , ns. However, a significant interaction between 'time' and 'condition' was found,  $F(1,99) = 6.46$ ,  $MSE = 287.29$ ,  $p = 0.01$ ,  $\eta^2 = 0.06$ . Scores in the experimental classes increase over time, while scores in the control class decrease. No class effects were found between experimental classes.

### *Transfer effect of instruction*

To answer the second research question, whether instruction based on two transfer theories achieves transfer, a repeated measures ANOVA analysis with condition as between factor was performed. This analysis on the evaluation of the *biology hit list* showed a main effect on 'time',  $F(1,99) = 11.62$ ,  $MSE = 45.12$ ,  $p = 0.00$ ,  $\eta^2 = 0.11$ , and no main effect on 'condition',  $F(1,99) = 0.75$ ,  $MSE = 4.11$ , ns. There was a significant interaction between 'time' and 'condition',  $F(1,99) = 19.34$ ,  $MSE = 75.06$ ,  $p = 0.00$ ,  $\eta^2 = 0.16$ . This interaction effect is not caused by an increase in scores of the experimental condition, but by a decrease in scores in the control condition. Furthermore, a class effect was found between the four experimental classes,  $F(3, 76) = 2.68$ ,  $MSE = 10.22$ ,  $p = 0.05$ ,  $\eta^2 = 0.10$ . Post hoc analysis revealed that the scores in one class decreased from pre to post test, while the scores in the other three classes increased from pre to post test.

To analyse the effects of instruction on transfer of evaluation of information and source an ANOVA was performed with the results on the *biology websites evaluation task* and condition as between factor. No main effect on 'time',  $F(1,99) = 2.00$ ,  $MSE = 69.95$ ,  $p = 0.16$ ,  $\eta^2 = 0.02$ , and also no main effect on 'condition',  $F(1,99) = 2.00$ ,  $MSE = 244.23$ , ns, was found. There was no interaction effect between 'time' and 'condition',  $F(1,99) = 0.19$ ,  $MSE = 6.74$ ,  $p = 0.66$ ,  $\eta^2 = 0.00$ . This indicates that there is no significant difference in scores between the students in the experimental and control condition, and thus no transfer effect was found. No class effects were found in the experimental classes.

### Learning results

The average score on the final exam was higher ( $M = 6.3$ ,  $SD = 1.1$ ) in the experimental classes than in the control class ( $M = 5.8$ ,  $SD = 1.1$ ). This difference was significant,  $t(99) = 1.97$ ,  $p = 0.05$ ,  $r = 0.19$ .

## Field notes

### *Experimental class A: the students*

During the first lesson, students were confronted with websites they had to use to solve a task. Some of these websites contradicted each other or provided information on a different subject. Students noticed that there was something wrong with the sites, but did not seem to adjust their actions accordingly. They wanted to finish the assignment. In the discussion that followed the assignment, students mentioned criteria they used to select information. In general, things like using a site they had used before, the site has to look good, there have to be sources or references mentioned on the site, were mentioned by them.

During the next lessons, it became clear that some students in this class caused problems and had a negative influence on the other students. During the final lesson, where students presented good and bad websites to each other, students showed that despite the fact that they did not really show it during the lessons, they actually had developed knowledge on how to evaluate websites (they could mention more and more sophisticated criteria) and were able to discuss websites with fellow students and defend their choice for a good or bad website.

### *Experimental class A: the teacher*

During the group discussions, the teacher did not follow the manual. Since his students had some difficulties with the assignments, the teacher had to spend more time on summarizing the historical content. Therefore, little time was left for discussions concerning the use of criteria when evaluating results, sources and information. The teacher did not check whether students filled out the process worksheet correctly.

### *Experimental class B: the students*

During the first lesson, students worked concentrated. Many students walked into the Wikipedia-trap. A Wikipedia site was presented as the first site, but this site did not give information about the task at hand (The treaty of Versailles, 1919), but provided information about a different treaty of Versailles, in 1782. Most students used this Wikipedia site to complete the assignment. During the group discussion students mentioned that they found contradictory information, sites without references and sites with poor lay-out. One student remarks: "If I have no knowledge on a subject, how do I know if the information on the site is reliable?" This is exactly the critical attitude we wanted to achieve with this first lesson. During a group discussion in one of the following lessons only a small group of students was prepared to discuss, while others remained silent. Not all students were happy with the new lessons; they feared that, since they had to find their own information, everyone would learn something else. These students would rather have more classical lessons. One student remarked: "I have never wondered about information not necessarily being correct, I believed everything and now they make me distrust everything!" During the final lesson, students were able to discuss websites with fellow students and defend their choice for a good or bad website.

### *Experimental class B: the teacher*

The teacher had to adjust the lesson program due to illness and classes being cancelled for meetings. The teacher tried to give the lessons according to the descriptions and followed

the manual during the discussions on the use of criteria when evaluating information. Due to cancellation of lessons, the group discussions were not held as frequently as planned. The teacher did not check whether students filled out the process worksheet correctly.

#### *Experimental class C: the students*

Students in this class were not very enthusiastic about the new lessons. They felt they already knew everything about the WWW. When asked what they had noticed about the websites during the first lesson, they mentioned that there were no references on the sites and layout was bad. Some students felt they shouldn't use Wikipedia because it is an open source, others mentioned you can use Wikipedia because it is better than most sites and the content is checked. During this discussion some students were busy playing games or talking to each other. Only a few students were active. This didn't change in the following lessons; most students were not motivated and complained about the assignments. During the final group discussion in the last lesson, students were not able to mention more criteria than the few they mentioned during the first lesson.

#### *Experimental class C: the teacher*

The teacher in this class tried to follow the manual for group discussions but was not always sure how to structure the discussions. The teacher admitted not feeling competent enough for leading the discussions, since developing a knowledge structure in this way was new to this teacher. The teacher encouraged the children to fill out the process worksheets.

#### *Experimental class D: the students*

Due to a miscommunication the first lesson was not given according to plan. Students had already answered the questions, without receiving the sites they had to use. When they were confronted with the sites for this assignment, reading and evaluating the sites was done very quickly, since they had already answered the questions. After a while students were asked to explain what they noticed about the sites. They mentioned that a newspaper is not always a reliable source, and that information differed between sites. In the next lessons, students seemed to understand what the lessons hoped to accomplish and why evaluating information is important, but they admitted they did not want to change their usual ways. They felt it took them more time to finish an assignment if they filled out the work sheets and evaluated every site. During the group discussion in the final lesson, some students were able to mention more criteria; some students did not engage in the discussion.

#### *Experimental class D: the teacher*

There was not always time to have a discussion after three lessons. The first time, the teacher mentioned more criteria than students. During the final discussion this had improved and the teacher made sure the students did the work. The manual was not always followed. The teacher did not check whether students filled out the process worksheet correctly.

In summary, the instructional program was only partly executed as planned. Most teachers did not check whether students filled out the process worksheet correctly and there was too little time for the group discussions on evaluation criteria every three lessons. For students the used approach in the lessons was new and they had to get used to it. Although students did not



seem to change their way of evaluation results, information and source, during the lessons, the final discussions showed students indeed had more knowledge on evaluation criteria.

## Discussion

Design-based research was used to develop and test an educational program to teach 9th grade students not only historical content, but also to evaluate search results (a hit list), information and source (information on a website and the website itself) when searching for information on the WWW. Furthermore, the program should also stimulate students to use these skills in a variety of settings (e.g., lead to transfer). Effects of the program on knowledge and use of criteria in the domain of instruction, on the transfer of the skills to another domain, and on the learning results concerning history content were determined.

We can conclude that instruction improved students' evaluation of information and source (websites). Students' scores of the experimental condition increased from pre to post test, while the scores of the control group decreased from pre to post test.

Instruction seems to have had no effect on students' evaluation of search results (hit list); students in the experimental condition maintained their scores from pre to post test and the scores of the students in the control condition decreased from pre to post test. This could not be due to the difficulty of the tasks used during the post test, because tasks were counterbalanced and rotated. An explanation could be that since the lessons were not always executed according to the lesson plans and the discussions were mostly focused on the evaluation of websites and information, students' use of criteria of evaluating hit lists was not triggered. This can explain the maintenance in scores of the students in the experimental condition. The scores of the students in the control condition decreased; this could be due to the fact that only one class participated in the control condition. Less time, or other interferences could have biased the results.

Furthermore, it can be concluded that the instruction did not achieve a transfer effect with regard to the use of criteria for the evaluation of websites. There was no difference in the scores on the post test between students in the experimental classes and the control class. It was expected that the experimental classes would score higher on the post test, if the program had led to transfer. Instruction did not have a transfer effect on students' use of criteria for evaluating search results as well. Students in the experimental condition maintained their scores from pre to post test and the scores of the students in the control condition decreased from pre to post test. The same explanation as stated above holds in this case. While it is positive that instruction improved students' evaluations of websites, it was unexpected that transfer was not achieved. The program was based on two transfer theories and the combination of both theories was hypothesized to achieve transfer. The results of the earlier study (Walraven et al. 2010), in which the two programs, each based on one transfer theory, gave rise to positive transfer effects, led the design team to believe that a combination of the strong points of both programs could even give rise to more transfer. The most likely explanation for not achieving a transfer effect is that the new program was not implemented to the full extent in all experimental classes. This is confirmed by the observations of the implementation of the program.

Another question concerned the effect of the instruction on students' learning results. It was hypothesised that the scores on the final test concerning the history content would not differ between the two conditions. Teachers of the experimental classes had some doubt about embedding the evaluation skills, because it would mean less time for history content, and as a consequence maybe lower grades for history. But results reveal that the score on

the final test of the experimental classes is significantly higher than the score of the control class. So, embedding evaluations skills does not cause lower grades.

Regarding the implementation of the program, the design team believed that short process worksheets and structured discussions to develop a rich and well connected knowledge structure would be the best way to achieve transfer. In the first design cycle, the long process worksheets caused problems for the students (Walraven et al. 2010). The team expected that if they provided students with shorter worksheets in the current study, students would not rebel against the worksheets, and use them in the intended way and transfer would be fostered. Unfortunately, the teachers did not encourage the students enough to fill out the sheets. Some students did not fill out sheets at all, others filled out sheets after completing the assignment instead of during, and only a few students filled out the sheets correctly. Focus of the students was still more on product, than on process. This is also found in a study by Julien and Barker (2009). Despite specific questions addressing the process of their search task (comparable with the process worksheets in our study), 11th and 12th grade students found it hard to recall their actions and choices, similar to our students who tried to fill out the sheets after the assignment.

The study described in this article is also in line with research by Brem et al. (2001) who state that improving 'student skills will not only be a matter of practicing evaluative strategies, but also of reflecting on the value of particular moves' (p.211). In our previous study we also concluded that the impact of these reflections (in form of the process worksheets) is moderated by the correct use of the sheets (Walraven et al. 2010). Teachers in the current study again did not always focus on the worksheets, which could have influenced the results in a negative way. Barranoik (2001) also found that teachers should pay more attention to the process, since students are mostly focused on product.

Another important educational measure in the program was the use of discussions on evaluation criteria to develop a rich and well connected knowledge structure about evaluation criteria. Important in these discussions was paying attention to the use of the evaluation criteria in various contexts or the way criteria are connected. A manual for the teachers had to make sure that all discussions were held in the same way and all aspects would receive the right amount of attention. Unfortunately, the discussions were often shortened due to time constraints and teachers did not always follow the manual.

The fact that two crucial factors of the educational program (filling out the sheets and the group discussions) were not completely implemented according to plan could—as stated—provide an explanation for the fact that no transfer was achieved. That the program was not executed according to plan is confirmed by the teachers who mentioned that the execution of this second program was more difficult than the execution of the first program. Julien and Barker (2009) also showed that having an explicit curriculum 'which explicates sound information searching skills, is clearly insufficient to ensure that students are learning these skills' (p. 15). Although Julien and Barker did not explore actual classroom practices, they suggest that teachers may believe that students learn these skills on their own, or teachers lack skills themselves and are unable to teach it. Our study showed that indeed teachers who are uncomfortable with the curriculum (e.g., the parts they did not design themselves) indeed had difficulties implementing the curriculum.

Both groups show high standard deviations on both website tasks. These deviations can be explained by a difference in skills between students or a difference in task execution. The latter is the most probable explanation. Students were asked which sites and what information they would or would not use, given the problem provided and were informed that it was not impossible to find 5–10 features they based their decision on. They could highlight those features. If students had circled a certain area on the site or written down a

comment like: “Site is old” they received a point for recognizing the criterion, and if there evaluation was correct (e.g., the site was indeed old) they received another point. So, student could receive points for recognizing a criterion and for using a criterion in the correct way. Students differed a lot in the number of features that highlighted. Some students only used one, others used more causing a wide spread in points scored.

A limitation of the study is the difference in number of students in experimental condition and control condition. There were four experimental classes and only one control class. Future research concerning the effect of the—further refined program—should be set up according to a more experimental design, where teachers try to stick to the lesson plan as close as possible.

Another focus of research could be testing whether the improvement in evaluations in the domain of instruction is still present a few months after the last lesson. It is not enough that students improve from pre to post test, students should acquire and use these evaluation skills throughout their lives. The role of the teacher to accomplish this should be further investigated. The teacher is an important factor, because the teacher is the one who must make students aware of the importance of evaluating information when searching the WWW for information. Future research should explore the role of the teacher in more detail and test whether being stricter to students with regard to the process worksheets influences the results. Next, the role of all teachers in the school should also be investigated. Integrating this evaluation skill throughout the curriculum is essential to foster transfer and prepare students for lifelong learning. Although the programs of the first design cycle achieved transfer to another domain, it was not tested whether the skills would still be used after time. It is critical that teachers in different domains pay attention to these skills and integrate instruction on these skills in their lessons. An important first step would be teaching teachers how to evaluate and how they can support their students to become critical web searchers.

**Acknowledgments** This research was supported by the Netherlands Organisation for Scientific Research (NWO) under project number 411-03-106.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

## Appendix

See Table 4.

**Table 4** Evaluation criteria

Sub skill	Criteria	
Evaluate search results	1. Relevance	What is the title given by the search engine and what is the contents of the summary?
	2. Kind (site/PDF)	What kind of source is it, a website, a word document, a PDF file?
	3. Reputation	What is the address? Is it a.com or.org address?
	4. Rank in hit list	How many results are there in total and what is the rank of the result I am evaluating?
	5. Familiarity	Have I used this site before, or have I heard good or bad things about it?
	6. Familiarity with the language	Is the site in a language I prefer and/or understand?

**Table 4** continued

Sub skill	Criteria		
Evaluate information	A. Usability		
	1. Language	In what language is the information written? Are there many grammatical or type errors? Is it filled with domain specific language?	
	2. Connection to task/relevance	Does the information answer (part of) the information problem?	
	3. Audience	Is the information aimed at a specific group of readers?	
	4. Currency	Is the information up to date?	
	5. Amount	Is there enough information on the page? Or only a part of the information I'm looking for?	
	B. Verifiability		
	1. Author	Who has written the information? Can I contact him/her?	
	2. References	Are there references on the page to used sources? Or links to more websites on the same subject?	
	3. Information agrees with more sites	Can I find the same information on more pages or is this information only available on this site?	
	4. Information agrees with prior knowledge	Does the information confirm what I already know?	
	5. Organization	Which organization is behind the information. A governmental organization, or a health organization? Can I find their logo on this site?	
	C. Reliability		
	1. Kind of information	What kind of information is it? A newspaper article or a forum? Is it an opinion or results from research?	
	2. Objectivity	Is the information objective or coloured by a certain point of view? Are there a lot of advertisements on the page?	
	3. Primary/Secondary	Is the information first hand or is it someone telling about someone who did something?	
	4. Goal	What does the (author of) information want to achieve. Sell something? Convince me of something or just inform me?	
	Evaluate source	A. Technical	
		1. Appearance	Does the site appeal to me? Does it have pictures or only text?
		2. Speed	Does it take a lot of time to load the page?
B. Usability			
1. Language		In what language is the site written? Are there many grammatical or type errors?	
2. Connection to task/relevance		Does the site have a connection to (part of) my information problem?	
3. Audience		For whom is the site meant? Who are it's visitors?	
4. Currency		Is the site updated regularly?	
C. Verifiability			
1. Reputation		Is this site famous or infamous? Does it have a good/bad reputation?	
D. Reliability			
1. Kind (site/PDF)		What kind of source is it, a website, a word document, a PDF file?	
2. Primary/secondary		Is the site an original source or a site telling about what is written somewhere else?	

## References

- Barker, J. (2005). *Critical evaluation of resources*. Retrieved May 16, 2006 from UC Berkeley Library Web site <http://www.lib.berkeley.edu/TeachingLib/Guides/Evaluation.html>.
- Barranoik, L. (2001). Research success with senior high school students. *School Libraries Worldwide*, 7, 28–45.
- Beck, S. (1997). "Evaluation criteria": The good, the bad & the ugly: Or, why it's a good idea to evaluate web sources. Retrieved May 10, 2006 from <http://lib.nmsu.edu/instruction/evalcrit.html>.
- Beljaarts, M. (2006). *Internet, een populair medium voor het zoeken van informatie bij schoolopdrachten [Internet, a popular medium for searching information for school assignments]*. 's-Hertogenbosch: Malmberg.
- Bennet, S., Maton, K., & Kervin, L. (2008). The digital natives' debate: A critical review of the evidence. *British Journal of Educational Technology*, 39, 775–786.
- Boekhorst, A. K. (2000). *Informatievaardig worden in het onderwijs, een informatiewetenschappelijk perspectief: Een vergelijkende gevallenstudie in Nederland en Zuid-Afrika [Becoming information literate in education, an information science perspective: a comparative case study in The Netherlands and South Africa]*. Unpublished dissertation. Retrieved May 10, 2006 from <http://cf.hum.uva.nl/akb/publicaties/prom-akb-tot.pdf>.
- Brand-Gruwel, S., Wopereis, I., & Vermetten, Y. (2005). Information problem solving: Analysis of a complex cognitive skill. *Computers in Human Behavior*, 21, 487–508.
- Brem, S. K., Russel, J., & Weems, L. (2001). Science on the Web: Students' evaluations of scientific arguments. *Discourse Processes*, 32, 191–213.
- Britt, M. A., & Aglinskias, C. (2002). Improving students' ability to identify and use source information. *Cognition and Instruction*, 20, 485–522.
- Fidel, R., Davies, R. K., Douglass, M. H., Holder, J. K., Hopkins, C. J., Kushner, E. J., et al. (1999). A visit to the information mall: Web searching behavior of high school students. *Journal of the American Society of Information Science*, 50(1), 24–37.
- Gick, M., & Holyoak, K. (1983). Schema-induction and analogical transfer. *Cognitive Psychology*, 12, 1–38.
- Grimes, D. J., & Boening, C. H. (2001). Worries with the Web: A look at student use of Web resources. *College and Research Libraries*, 62(1), 11–23.
- Hirsch, S. G. (1999). Children's relevance criteria and information seeking on electronic resources. *Journal of the American Society for Information Science*, 50, 1265–1283.
- Jones, S. (2002). *The Internet goes to college. How students are living in the future with today's technology*. Washington, DC: Pew Internet and American Life Project.
- Julien, H., & Barker, S. (2009). How high-school students find and evaluate scientific information: A basis for information literacy skills development. *Library & Information Science Research*, 31, 12–17.
- Kafai, Y., & Bates, M. J. (1997). Internet web-searching instruction in the elementary classroom: Building a foundation for information literacy. *School Library Media Quarterly*, 25, 103–111.
- Kirk, E. E. (1996). *Evaluating information found in the Internet*. Retrieved May 10, 2006 from Johns Hopkins University, the Sheridan Libraries Web site <http://www.library.jhu.edu/researchhelp/general/evaluating/index.html>.
- Klausmeijer, H. J. (1990). Conceptualizing. In B. F. Jones & L. Idol (Eds.), *Dimensions of thinking and cognitive instruction* (pp. 93–138). Hillsdale: Erlbaum.
- Koot, U., & Hoveijn, M. (2005). Google, de magische antwoordmachine. [Google, the magical answer-machine]. 28, 11, 18–22.
- Larkin, J., McDermott, J., Simoa, D. P., & Simon, H. A. (1980). Expert and novice performance in solving physics problems. *Science*, 208, 1335–1342.
- Lorenzen, M. (2002). The land of confusion?—High school students and their use of the World Wide Web for research. *Research strategies*, 18, 151–163. Retrieved September 7, 2006 from <http://www.libraryinstruction.com/confusion.html>.
- Lyons, D., Hoffman, J., Krajcik, J., & Soloway, E. (1997). *An investigation of the use of the World Wide Web for on-line inquiry in a science classroom*. Paper presented at the meeting of the National Association for Research in Science Teaching, Chicago, IL.
- MaKinster, J. G., Beghetto, R. A., & Plucker, J. A. (2002). Why can't I find Newton's third law? Case studies of students' use of the web as a science resource. *Journal of Science Education and Technology*, 11, 155–172.
- Maughan, P. D. (2001). Assessing information literacy among undergraduates: A discussion of the literature and the University of California-Berkeley assessment experience. *College and Libraries*, 61, 71–85.

- Metzger, M. J., Flanagin, A. J., & Zwarun, L. (2003). College student Web use, perceptions of information credibility, and verification behavior. *Computers & Education, 41*, 271–290.
- O'Hanlon, N. O. (2002). Net knowledge: Performance of new college students on an Internet skills proficiency test. *The Internet and Higher Education, 5*, 55–66.
- Ormondroyd, J. (2004). *Critically analysing information sources*. Retrieved May 10, 2006 from Cornell University Library, Reference Department, Instruction, Research, and Information Services (IRIS), Ithaca, NY, USA. Web site <http://www.library.cornell.edu/olinuris/ref/research/skill26.htm>.
- Perkins, D. N., & Salomon, G. (1989). Are cognitive skills context-bound? *Educational Researcher, 18*, 16–25.
- Prenkys, M. (2001). Digital natives, digital immigrants. *On the Horizon, 9*(5), 1–6.
- Rothenberg, D. (1998). How the web destroys student research papers. *Educ Digest, 63*, 59–61.
- Salomon, G., & Perkins, D. N. (1989). Rocky roads to transfer: Rethinking mechanisms of a neglected phenomenon. *Educational Psychologist, 24*, 113–142.
- Simons, R. J., Van der Linden, J., & Duffy, T. (2000). New learning: Three ways to learn in a new balance. In R. J. Simons, J. Van der Linden, & T. Duffy (Eds.), *New learning* (pp. 1–20). Dordrecht: Kluwer.
- Thorndike, E. L., & Woodworth, R. S. (1901). The influence of improvement in one mental function upon the efficiency of other functions. *Psychological Review, 8*, 247–261.
- Wagner, J. F. (2006). Transfer in pieces. *Cognition and Instruction, 24*(1), 1–71.
- Wallace, R. M., Kupperman, J., Krajcik, J., & Soloway, E. (2000). Science on the web: Students online in a sixth-grade classroom. *The Journal of the Learning Sciences, 9*(1), 75–104.
- Walraven, A., Brand-Gruwel, S., & Boshuizen, H. P. A. (2008). Information-problem solving: A review of problems students encounter and instructional solutions. *Computers in Human Behavior, 24*, 623–648.
- Walraven, A., Brand-Gruwel, S., & Boshuizen, H. P. A. (2009). How students evaluate information and sources when searching the World Wide Web for information. *Computers & Education, 52*, 234–246.
- Walraven, A., Brand-Gruwel, S., & Boshuizen, H. P. A. (2010). Fostering transfer of web searchers' evaluation skills: A field test of two transfer theories. *Computers in Human Behavior, 26*, 716–728.
- Wertheimer, M. (1961). *Productive thinking*. London: Tavistock.