Computers in Human Behavior 62 (2016) 90–104

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

Full length article

# Completion strategy or emphasis manipulation? Task support for teaching information problem solving

Jimmy Frerejean*, Johan L.H. van Strien, Paul A. Kirschner, Saskia Brand-Gruwel

*Welten Institute, Open University of the Netherlands, Valkenburgerweg 177, 6419 AT Heerlen, Netherlands*

ARTICLE INFO

*Article history:*
Received 16 September 2015
Received in revised form
12 March 2016
Accepted 15 March 2016

*Keywords:*
Completion strategy
Emphasis manipulation
Prompting
Instructional design
Information literacy
Information problem solving

ABSTRACT

While most students seem to solve information problems effortlessly, research shows that the cognitive skills for effective information problem solving are often underdeveloped. Students manage to find information and formulate solutions, but the quality of their process and product is questionable. It is therefore important to develop instruction for fostering these skills. In this research, a 2-h online intervention was presented to first-year university students with the goal to improve their information problem solving skills while investigating effects of different types of built-in task support. A training design containing *completion tasks* was compared to a design using *emphasis manipulation*. A third variant of the training combined both approaches. In two experiments, these conditions were compared to a control condition receiving conventional tasks without built-in task support. Results of both experiments show that students' information problem solving skills are underdeveloped, which underlines the necessity for formal training. While the intervention improved students' skills, no differences were found between conditions. The authors hypothesize that the effective presentation of supportive information in the form of a modeling example at the start of the training caused a strong learning effect, which masked effects of task support. Limitations and directions for future research are presented.



## 1. Introduction

Searching the web for information seems effortless for students; they simply navigate to a popular search engine, type in a couple keywords, and select some of the sources that appear to be relevant (MaKinster, Beghetto, & Plucker, 2002). Most students easily find their way without any explicit instruction. They paraphrase, cite, or – in the worst case – copy and paste some of the text into their own document and the job is done (De Vries, van der Meij, & Lazonder, 2008). The abundance of information on the internet is a bliss. While this may be viewed as a successful process in the eyes of the student, from an educational perspective it can be a waste of time. If the student is not equipped with the necessary skills, such as advanced search strategies and the ability to critically scrutinize information sources to determine relevance and reliability, chances are that the search process and the product fall short of what the teacher intended. It may be true that younger generations of students appear to quickly master the skills needed to navigate online information sources, but it is premature to claim that they automatically develop the skills to find correct and reliable online sources and learn from them (Kennedy, Judd, Churchward, Gray, & Krause, 2008; Kirschner & van Merriënboer, 2013; Rosman, Mayer, & Krampen, 2016).

While most educational institutions acknowledge information problem solving (IPS) as an essential academic skill, they often struggle with implementation (Badke, 2010). To promote transfer of IPS to daily practice, it is advisable to practice these skills in different contexts and across different domains throughout the whole curriculum. This is problematic, and most schools experience great difficulty in finding a suitable place and time in the curriculum. Many, in turn, resort to providing nothing more than a short library training. To support teachers and faculty in embedding IPS skills in educational curricula, it is desirable to investigate which instructional approaches work well for IPS skills. This paper takes a first step in that direction, describing the development and empirical testing of instruction for IPS skills, based on a solid instructional design model for teaching complex skills. Implications are discussed for both the domain of instructional design and information problem solving.

* Corresponding author.
*E-mail addresses:* jimmy.frerejean@ou.nl (J. Frerejean), johan.vanstrien@ou.nl (J.L.H. van Strien), paul.kirschner@ou.nl (P.A. Kirschner), saskia.brand-gruwel@ou.nl (S. Brand-Gruwel).

http://dx.doi.org/10.1016/j.chb.2016.03.048
0747-5632/

## 2. Theoretical framework

### 2.1. Information problem solving

In educational settings, teachers often use information problems, where the necessary information to solve the problem is lacking, as an educational approach. The student is required to gather the missing information from external sources and combine the findings to construct a solution. Simple information problems, such as looking up the average monthly temperature in a country, pose little challenge for most students. Complex information problems, such as writing an essay on the effects of global warming on biodiversity, are a far more difficult challenge, because students will need to find, evaluate, and process sources of information that can vary greatly in terms of their trustworthiness, bias, reliability, or can contain contradictory information. Teachers often expect that having students search for information will automatically lead to their learning (Kirschner, Sweller, & Clark, 2006). But correctly and efficiently solving an information problem is a complex higher-order cognitive competence requiring a broad range of different cognitive skills that these students might not possess. The range of skills has been summarized as a 5-step model (see Fig. 1) in which students iterate between the stages 'define the problem', 'search information', 'select information', 'process information', and 'present information', each step consisting of several constituent skills (Brand-Gruwel, Wopereis, & Vermetten, 2005; Brand-Gruwel, Wopereis, & Walraven, 2009).

To solve an information problem, the learner first needs to reach an understanding of the task and identify the needed information to define and delimit the task domain. In this step, formulating a clear and concise question is essential to stay focused and avoid unnecessary deviations while searching. Second, search terms need to be generated and tried out in a search engine. By identifying key concepts from the question and then systematically changing, adding, or removing terms while correctly using the available Boolean operators, the learner maximizes the chance to find relevant information sources. Third, it is important to maintain a critical attitude while evaluating the search results page, the subsequently visited information sources, and the information itself. Critical scrutiny avoids spending time on irrelevant websites or becoming occupied with information that is outdated, false, or which originates from unreliable or biased sources. Fourth, when relevant and reliable sources are found and stored, the learner needs to process their contents, deal with overlapping and conflicting information, and synthesize the different elements chosen from the separate sources. Finally, the solution can be presented in a product such as an essay or a presentation, depending on the task. It is important that the product clearly answers the question that was defined earlier in the task. Moreover, during all of these steps, the learner should regulate the search process, decide whether sufficient useful information has been found, and steer the process to avoid deviations or distractions.

Previous research indicates students may quickly develop the instrumental skills needed to operate digital devices and use software and internet browsers, but IPS skills are generally underdeveloped or absent. In a comparison of experts and novices, Brand-Gruwel et al. (2005) found that novices took less time for orientation, chose less effective keywords, judged and evaluated sources less often, and hardly regulated their process. In a literature review, Walraven, Brand-Gruwel, and Boshuizen (2008) discuss several studies that show execution of IPS skills leave much room for improvement for all age groups. Similarly, studies by Van Deursen and van Dijk (2009) and Van Deursen and van Diepen (2013) show users of all ages experience problems with query formulation, evaluation of search results and processing of information.

Two things become clear from these findings. First, IPS is a complex higher-order cognitive skill. Successful problem solving depends on the existence of knowledge, the mastery and coordination of a set of skills and the adoption of a critical attitude. Second, research shows clear deficiencies in students of almost all ages. In general, students' IPS skills are often overestimated or expected to develop naturally over time. These IPS skills may not be of the level that is often expected of the student problem solver, or from the so-called 'digital natives' (see also: Kirschner & van

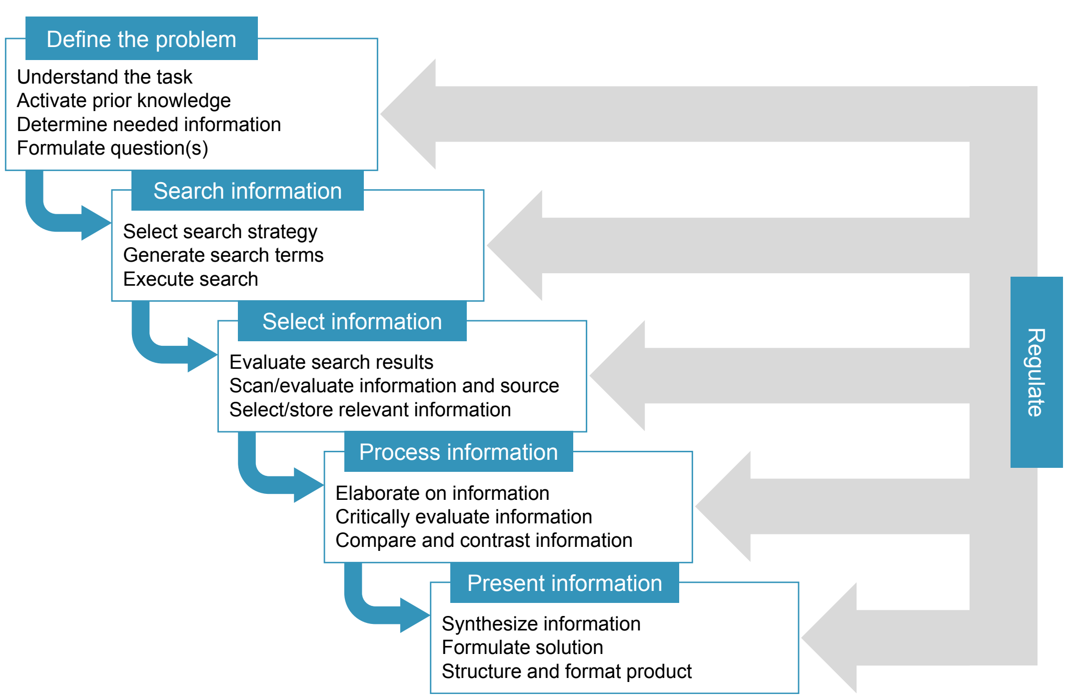


**Fig. 1.** Decomposition of the skill 'information problem solving' (based on Brand-Gruwel et al., 2005).

Merriënboer, 2013). Providing students with a complex task for which they do not possess the required skills risks overloading their memory systems and lowering task performance and learning. Therefore, the development of evidence-based instruction for fostering IPS skills is warranted.

### 2.2. Instructional design for complex learning

Complex learning is defined as “the integration of knowledge, skills and attitudes; coordinating qualitatively different constituent skills; and often transferring what was learned in school or training to daily life and work” (Kirschner & van Merriënboer, 2009, p.244). The *Four Component Instructional Design* (4C/ID) model provides an extensive blueprint and approach for developing instruction to achieve complex learning, based on solid psychological and educational research (Van Merriënboer & Kirschner, 2013). First, the model advocates the use of authentic, whole tasks that require integration of knowledge, skills and attitudes, and coordination of constituent skills. Second, it provides guidelines to correctly provide the information needed to solve the problems: domain knowledge and a structured approach to solve the problem. Third, it advises providing just-in-time procedural information during the tasks to aid problem-solvers with routine tasks. The fourth component, part-task practice, is necessary when performance of these routine tasks needs to be automated.

This task-centered approach confronts learners with a series of whole-tasks in the learner's zone of proximal development. Task complexity increases to keep up with learner progress. However, especially in the early phases of learning, tasks can be too complex for the learner, because they introduce too many interacting elements or the learner's knowledge schema are insufficiently developed. In these cases, the learner's memory system may become overloaded, which can negatively impact learning (Paas & van Merriënboer, 1994). In situations of complex learning and authentic tasks, there are many elements that potentially increase the amount of cognitive load experienced by the student. It is therefore essential that instructional designers take great care to reduce unnecessary load, while maintaining activities that induce germane load and lead to learning.

For IPS specifically, task complexity is not the only factor that influences the demands on working memory during problem solving, and in consequence, learning and instruction. Rouet (2009) summarizes additional factors in a conceptual framework comprising three dimensions: individual variables, information resources, and problem context. Instructional designers should be aware that personal factors, such as an individual's domain-specific knowledge (Monchaux, Amadieu, Chevalier, & Mariné, 2015), age (Chevalier, Dommes, & Marquié, 2015), attitudes and biases (Ford, Miller, & Moss, 2005; Van Strien, Brand-Gruwel, & Boshuizen, 2014), epistemic beliefs (Kammerer, Bråten, Gerjets, & Strømsø, 2012), and reading skills (Rouet, Ros, Goumi, Macedo-Rouet, & Dinet, 2011) can affect the learning process and outcomes. Similarly, source factors (DeStefano & LeFevre, 2007) and task type (Wirth, Sommer, von Pape, & Karnowski, 2015) may influence variables in the learning process. While most of these factors lie outside the designer's influence, they all affect the demand imposed on working memory during the IPS process.

For situations where tasks may be too demanding for a learner to complete successfully, the problem-solving process must be supported (Van Merriënboer, 2013). The 4C/ID model stresses the importance of built-in task support. While learners can be supported in many ways (i.e. with case studies, modeling and/or worked examples, inducing reflection, etc.), the current experiments focus on two approaches that appear most applicable to IPS instruction, namely *the completion strategy* and *emphasis manipulation*.

#### 2.2.1. Completion tasks

A completion task is a problem where the learner is provided with a given state and a partial solution. After studying the partial solution and the given information, the learner then has to complete the remaining solution steps in order to solve the problem (Van Merriënboer, 1990; Van Merriënboer & De Croock, 1995). This approach is effective for several reasons. First, completion tasks inherently stimulate active processing of the given solution steps because they contain essential information the learner needs to process before being able to continue. In addition, the provided solution steps are examples of a correct systematic approach to solving the problem. This enables learners to study the examples and by induction generate schemas of correct solution strategies themselves (Van Merriënboer, 2013). Studying correct examples (albeit partial solutions) can often be more effective than solving whole problems, especially early in the learning process (Renkl & Atkinson, 2003). When learners lack the necessary schemas and strategies, they will fall back to naïve and ineffective strategies such as means-end analyses or trial-and-error to solve the problem. Providing sufficient worked-out steps in this phase can avoid this (Van Gog, Paas, & van Merriënboer, 2004).

The second benefit of using completion tasks is that a designer can change the number of worked-out steps to adapt the task to the learner's level. For learners in an early learning phase, it would be beneficial to increase the number of worked-out steps (e.g., one or even no steps missing), providing ample examples of correct complete or partial solutions and allowing the learner to induce the necessary schemas and strategies (Atkinson, Renkl, & Merrill, 2003; Renkl, Atkinson, & Große, 2004). In later learning phase, learners benefit more from more conventional tasks that contain just a few worked-out steps. Offering too many worked-out steps to these learners would create the risk of inducing the expertise-reversal effect (Kalyuga, Ayres, Chandler, & Sweller, 2003; Kalyuga & Sweller, 2004). By gradually reducing the number of worked-out steps as a learner progresses, the amount of support that is offered corresponds more closely to the amount of support that is actually needed. In the context of IPS, this fading of solution steps can only be applied backward, meaning that worked-out solution steps late in the process will always fade before solution steps early in the process. To illustrate, consider the opposite: A worked-out example where the solution and information sources are given but the student needs to define the problem and generate search terms. Such a backward information problem is unrealistic, and practicing it has little purpose. In conclusion, a gradual transition from completely worked-out problems to conventional problems would be a good strategy for instruction: an approach dubbed the *completion strategy*.

Wopereis, Frerejean, and Brand-Gruwel (2015) implemented the completion strategy in a university-level IPS training program. In their training, an example completion task provides students with a problem orientation, a well-formulated problem statement and research question, and a partial list of search terms. In this case, the step ‘problem definition’ is completely worked out, and the step ‘searching’ is partially worked out. Students are required to process the problem orientation to become familiar with the task domain and to activate any prior knowledge. The given problem statement and research question provide a clear direction for the search and inform them which information is needed, and consequently, which information is not. Based on this orientation, students then extend the list of search terms and proceed with the search for information and the remaining solution steps (‘select information’, ‘process information’, and ‘present information’). Compared to a conventional task where students perform the whole task, this approach requires less decision making - and therefore less room for error - and provides an additional example to learn from. The expectation here is that such tasks will impose fewer cognitive demands than conventional problems.

#### 2.2.2. Emphasis manipulation

Students can also be supported by guiding them in the allocation of their attention to a certain skill (i.e., generating search terms) or a step in the process (i.e., select information) within a learning task. Students then perform the whole task from beginning to end, but just one aspect of the solution procedure is *emphasized*, often by instructions and feedback. In subsequent tasks, the emphasis and thus the allocation of the learner's attention shifts to a different aspect of the task. Note that the task is not broken up into part-tasks, but only the relative emphasis of the selected aspect varies. All skills are still performed in the context of the whole task. This approach, called *emphasis manipulation* or emphasis change (Gopher, 2006; Gopher, Weil, & Siegel, 1989), reduces strain on working memory because not all instruction needs to be kept available in working memory, and attention is focused on a single aspect, not divided over all aspects.

The *emphasis change approach* was effective in a training regime for a high-workload computer game called Space Fortress and in several dual-task settings (Gopher, 2006). In other research, students who received whole-task training with emphasis change were less easily disrupted by a concurrent task than students receiving part-task training (Fabiani et al., 1989). In addition, Yechiam, Erev, and Gopher (2001) demonstrated that an emphasis change approach is more effective than guided instruction in settings where searchers quickly converge to suboptimal strategies. The idea here is that problem solvers make only small changes to their current, suboptimal, strategy and insufficiently explore more diverse solution strategies, a process called *melioration* (Yechiam, Erev, Yehene, & Gopher, 2003). Emphasis change protocols facilitate the exploration of other, potentially more effective strategies.

The errors that can be observed when novices search the web may be a sign of melioration. Lacking sufficient skill, they employ naive strategies that will find *some* results (partly due to increasing quality of search engines), even though it may not be the information they are looking for. This will then lead them to obtaining suboptimal information, which in turn leads to a suboptimal solution to the task. Students experience the success of solving the problem, which reinforces their current behavior and leads to a similar approach to the next problem. Students see no reason to expend extra effort to significantly change their strategy. Emphasis change can encourage students to explore other strategies, such as more extensive planning, or using thesauri to generate keywords, which increases the chance of a more effective or efficient problem solving process.

Placing emphasis on specific aspects of a task can be done by incorporating instruction and feedback during those specific aspects of the learning task. A simple and effective method to provide instruction and feedback in an online environment is by using prompts (see: Stadtler & Bromme, 2008). In the case of IPS, three types of prompts are effective: anticipative prompts delivered before execution of the targeted skill, instructional prompts delivered just in time before execution of the targeted skill, and reflection prompts delivered after performing the skill.

Consider a student working on a learning task where the skill *evaluating sources* is emphasized and therefore accompanied by prompts. Before she starts evaluating sources (i.e., the targeted skill), she is prompted: “Describe your approach to the next step. Where will you focus your attention?” By articulating her upcoming actions before performing the skill, anticipative reasoning, a skill found in effective problem solvers, is stimulated (Renkl, 1997). The student answers: “I'll look at the result list and click on some of the titles that seem interesting. I'll then read that text. If it seems relevant, I'll probably use it.” The answer reveals that her solution schema is still incomplete, and that she has not yet learned to evaluate a search engine results page or an information source. Merely activating knowledge is therefore not sufficient. Her current schemas or strategies need to be corrected or completed.

She is prompted again, this time simply with instructions. The instructional prompt explains how to evaluate search results (i.e., pay attention to domain names, publication dates, snippets) before clicking a link and how to judge information sources (i.e., take into account author reputation, target audience, information goal, publication date). It essentially gives general feedback on her previous answer. The student will acknowledge that her previously articulated approach was incomplete and that she should not merely click ‘interesting’ links and use ‘relevant’ information. She learns that there are many more criteria to use to discriminate between interesting and relevant. She then processes this information and immediately carries out the solution step, with this new knowledge in memory. The subsequent application of the new knowledge stimulates assimilation into knowledge schemas.

To enforce this process, a reflection prompt can be delivered after the step is performed: “How did it go? Did you encounter any problems?” This prompt induces reflection and forces her to look back at how she applied the new knowledge, which should reinforce the use of a correct or more effective solution strategy (Saito & Miwa, 2007; Stark & Krause, 2009). Taken together, this combination of three prompts, the *prompt triad*, fulfills the purpose of emphasis manipulation by first lowering cognitive demand by focusing student attention to a particular aspect of the task while leaving the whole task intact and then promoting improvements in strategies by activating and correcting current knowledge schema.

## 3. The present study

Seemingly little research has focused on the development of holistic instruction for IPS. Most studies either focus on elements of instruction, such as feedback (e.g., Timmers, Walraven, & Veldkamp, 2015), restrict the search space to prefabricated portals (De Vries et al., 2008), or focus instruction on elements of the skill, such as source evaluation (Walraven, Brand-Gruwel, & Boshuizen, 2010). Some are focused on classroom interventions (e.g., Argelagós & Pifarré, 2012; Kuiper, Volman, & Terwel, 2008). In the current study, we adopt a holistic approach for teaching the complete skill in individual (online) instruction and take a first step towards developing instruction based on whole-tasks with built-in task support. Two experiments were conducted to investigate the effects of two forms of task support (*completion strategy* vs. *emphasis manipulation*) on the acquisition of IPS skills in a short online training. This training was embedded as a standalone practical assignment in university students' first-year curriculum. As an intervention in a naturalistic setting, this training aimed to develop students' IPS skills while detecting differences in the extent of learned skills due to the different methods of support. It was expected that students who receive at least one form of task support (i.e., completion tasks and/or emphasis manipulation) will perform better than students who do not receive task support (Hypothesis 1) and students who receive a combination of both forms of support will perform better than students who receive only a single form of support (Hypothesis 2). To help explain differences in learning outcomes, students were asked to report the required mental effort at several points during the learning phase.

## 4. Experiment 1

### 4.1. Method

#### 4.1.1. Participants

A total of 96 students between 18 and 24 years old ($M_{age} = 18.7$ years) participated in this experiment, 89 of whom were female (92.7%) and 7 were male (7.3%). All participants were first-year Pedagogical Science students at a Belgian university.

#### 4.1.2. Experimental design

The experiment was a regular pretest-posttest design with four conditions. All conditions received a 2-h online training consisting of an instructional video, a modeling example, and four learning tasks. Each condition received a different form of task support during three of the four learning tasks. The first condition received task support in the form of the completion strategy combined with emphasis manipulation (CS + EM). The second condition received completion tasks, but no emphasis prompts (CS). The third condition received emphasis prompts, but no completion tasks (EM). The fourth condition was a control condition and received conventional learning tasks without support. The different forms of task support are further detailed in the section 'Task support'.

#### 4.1.3. Materials

*4.1.3.1. Online training.* In a 2-h classroom session, students received an online training that started with a 14-min instructional video introducing the five steps of the IPS process (i.e., 'define', 'search', 'select', 'study', 'present') including their constituent skills. The instructional video was followed by a modeling example: a 10-min screencast in which a fictitious expert showed a systematic approach to solving an information problem. This modeling example was split into four short fragments that ended with the questions "What do you think of the actions of the expert?" and "How does this differ from your current approach?" intended to stimulate students to formulate explanations and stimulate active processing of the example (Atkinson et al., 2003; Renkl & Atkinson, 2002). These elements formed the supportive information component in the 4C/ID model.

The training further comprised four learning tasks in the form of a web search exercise. Students received a problem description and had approximately 15 min to search the web for information and formulate a solution to the problem. The topics were: effects of stretching before sports, effects of electromagnetic radiation from cell phones, effects of violence in videogames, and effects of using media devices before sleeping. The learning tasks guided students through the problem solving steps with on-screen instructions. Students were asked to explicitly formulate research questions and search terms, and list the URL of four sources that contributed to their solution, along with an explanation of why they chose these sources. At the end of the learning task they formulated a solution in a few sentences. Each of the experimental conditions received a different form of support during learning tasks 1 to 3. A fourth and final task was presented that did not include any support or guidance, but simply gave a problem description and a textbox for an answer. This task was identical for all students and contained no explicit instruction.

*4.1.3.2. Task support.* For the three experimental conditions, learning tasks 1 to 3 contained built-in task support in the form of completion tasks, emphasis prompts, or both. These tasks were designed in a way to support the problem-solving process without overloading the student. While the IPS model (Brand-Gruwel et al., 2009) describes a five-step approach, to comply with time constraints in this experiment the steps 'select information' and 'process information' were merged to a single step and no task support was supplied on the final step: 'present information'. Presenting information can be done in countless ways, and providing support on this skill would be very time-consuming. Additionally, students likely benefit more from support on the first four steps than from support on presenting information. In conclusion, task support is offered on the steps 'define the problem', 'search information', and 'select & process information'.

The control condition received no task support at all, meaning that they work through each learning task by following the guidance on the screen that take them through the steps 'define the problem', 'search information', 'select & process information' and 'present information'.

The EM condition received emphasis manipulation, meaning that each learning task contained one solution step that was emphasized with a prompt triad: an anticipative prompt and an instructional prompt before execution of the step, and a reflective prompt afterwards. In learning task 1, the step 'select & process information' was emphasized, in learning task 2 the step 'search information', and in learning task 3 the 'step define the problem'.

The CS condition received completion tasks. In these tasks, some solution steps are already worked out and replaced with a very short video (approximately 1–2 min) of a fictitious expert reasoning through the solution step. No further action was required. The worked-out steps were faded backwards, meaning each subsequent learning task contained one less worked-out step and students were therefore required to perform one step more in each learning task. In learning task 1, the steps 'define the problem' and 'search information' were worked out and students had to select and process sources from the given search results to formulate a solution. In task 2, only 'define the problem' was worked out and all other steps had to be performed. In task 3, no steps were worked out.

Finally, in the CS + EM condition completion tasks and emphasis prompts were combined, meaning that the prompt triad was added to the first step that followed the worked-out steps. This entails that in task 1, the first two steps were worked out: 'define the problem' and 'search information'. The next step, 'process & select information' was emphasized with a prompt triad. The final step had to be performed without support. In task 2, only the first step was worked out and emphasis shifted to 'search information'. In the third task, no worked-out steps were given and 'define the problem' was emphasized. See Fig. 2 for a graphical overview of the experimental design.

*4.1.3.3. Measurement of information problem solving skill.* To measure IPS skills in such a short timeframe, an online skills test was developed that aimed to reveal the student's level of performance without requiring the performance of another whole-task. The tests confronted students with seven fabricated situations that occur during an information problem and asked them to formulate their next action. This closely mimics a realistic task situation. To ensure validity, the items were based on important subskills in the first four steps of the IPS model by Brand-Gruwel et al. (2009). The step 'present information' was not measured because presentation of a problem solution is a multifaceted skill too difficult to measure quickly, and the training did not include support on this step. See Table 1 for an overview of the pretest and posttest items. For example, the item corresponding to 'select information' showed a fabricated SERP (search engine results page) and asked students to indicate which sources they would select and why. Answers were scored blindly, based on a task-specific rubric that resulted in a maximum subscore of 4 points per step, for a maximum total of 16 points. The scoring sheet and procedure are included in Appendix 1. The items in the pretest concerned the topic (i.e., problem domain)

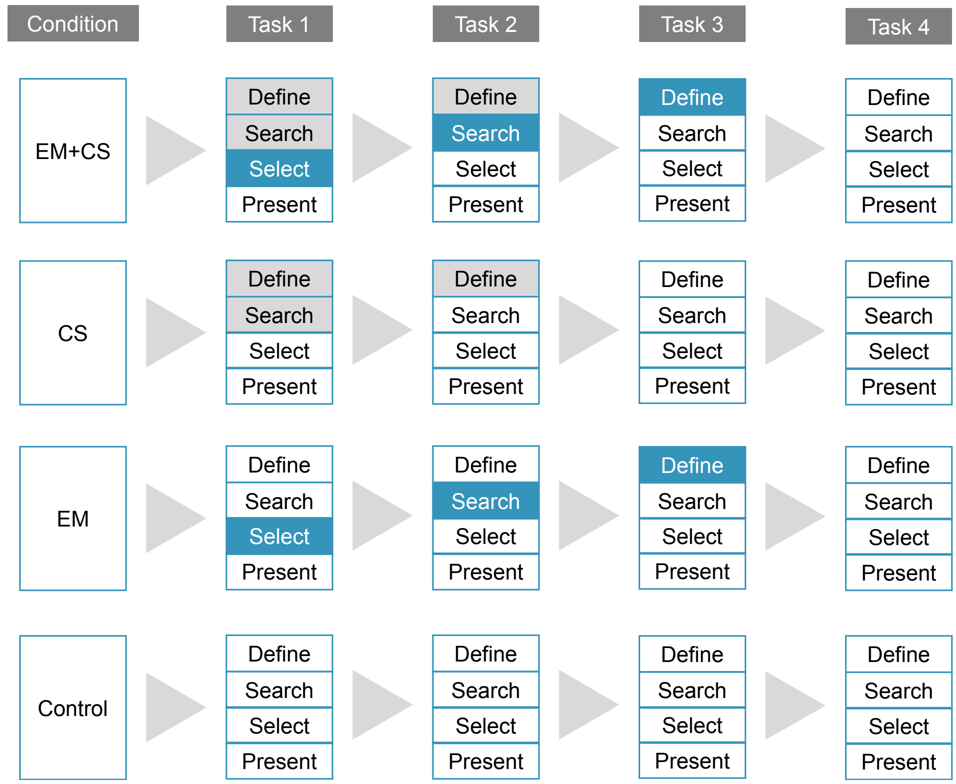


**Fig. 2.** Overview of the experimental design: four conditions (rows) received four learning tasks (columns) that consist of four steps. Worked-out steps in these tasks are marked with gray, emphasized steps are colored. Steps that are not colored contained no built-in support. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

**Table 1**
Overview of pretest and posttest.

| Item | Step | Subskill | Given | Question |
|---|---|---|---|---|
| 1 | Define the problem | Problem orientation | A problem description | How would you start this task? What is your first step and why? |
| 2 | Define the problem | Formulating a question | A problem description | Which problem statements would you formulate? Why do you choose these? |
| 3 | Search information | Generating search terms | A problem description | Which search query would you type into Google? Formulate two alternative search queries. |
| 4 | Select information | Evaluating search results | A fabricated SERP | Which three websites would you select? Why did you select these websites? |
| 5 | Process information | Scanning a source | A screenshot of a text-rich website, zoomed out so the text is unreadable | What do you do when you visit a text-rich website and want to find out if it contains relevant information? How do you proceed? |
| 6 | Process information | Evaluating information | A short text fragment containing an argument given by an expert | Which criteria do you use to determine whether information is useful for your task? What are your conditions for use? |
| 7 | Process information | Contrasting information | Two short, contradicting arguments | How do you deal with contradicting information? How does this affect your solution? Explain. |

of gender-specific education. The posttest items were identical to the pretest items, but on the topic of the malleability of intelligence. A second experimenter rescored 20 randomly chosen participants in order to obtain a measure of inter-rater agreement.

*4.1.3.4. Mental effort rating.* Integrating and coordinating the skills, knowledge and attitudes that are required to effectively and efficiently solve an information problem is a complex activity that places high demands on the learner's memory system. To bring down this complexity, built-in task support is incorporated in the learning tasks. It can be expected that different types of support impose different amounts of cognitive load on the students. Lacking an objective, direct way to measure cognitive load, experienced mental effort was measured as a proxy. During the learning phase, each learning task ended with a short measurement of experienced mental effort: a 9-point mental effort rating scale (Paas, 1992): *How much effort did it take to perform this task?* While all students were instructed to spend approximately 15 min on each learning task, working through the extra prompts and worked-out steps may have increased time on task for those students and perhaps put the students under time pressure. Performing the task under high time pressure might cause an increase in experienced mental effort. Therefore, time pressure was explicitly measured with the *temporal demand* item from the NASA-TLX (Hart & Staveland, 1998): *How hurried or rushed was the pace of the task?*

*4.1.4. Data analysis*

The scores on the pretest and posttest were analyzed with a repeated measures analysis of variance with *type of support* (CS + EM vs CS vs EM vs Control) as a between-subjects variable and *time of test* (pretest vs posttest) as a within-subjects variable. The same analysis was conducted on the subjective *mental effort* rating and the *time pressure* rating but with *learning task* as a within-subjects variable. In addition, an analysis of variance was conducted on the ratings per learning task to investigate differences in required mental effort between conditions.

*4.1.5. Procedure*

The training was embedded in the students' current curriculum as a practical assignment and offered in four different timeslots. Students were free to choose a timeslot that fit their schedule. During the 2-h training session, students took place at a computer in the university computer room and logged in to the online learning environment. After logging in, students first filled out a short preliminary questionnaire and were automatically randomly assigned to one of four conditions. They were instructed to work individually through the tasks they received on screen and informed that their screen content could differ from that of the other students. The experimenter asked students to spend approximately 15 min on each learning task, comparable to similar tasks used in other research (Lazonder, Biemans, & Wopereis, 2000; Lazonder, 2000). They then received the following: pretest, instructional video, modeling example, four learning tasks, and posttest. Each learning task concluded with the mental effort and time-pressure ratings. The instructional video and modeling example remained available via a link during the learning tasks. Before the posttest, students filled out a short evaluation and a final mental effort rating for the training as a whole. After the posttest, students signed for informed consent, received course credit and were subsequently dismissed. A debriefing with preliminary results followed 8 weeks later.

*4.2. Results*

The four randomly generated conditions did not differ significantly on any of the items on the preliminary questionnaire, such as age or prior education. They reported equal amounts of time spent behind a computer per day, and no differences in the use of the computer for information retrieval (either for personal or educational goals), news, social media, chatting, and entertainment. The sample can therefore be considered homogeneous. Some data were scored as missing due to the fact that students answered questions with a dash or a space, and some data were lost due to incidental technical problems. On the posttest, missing values were substituted for their corresponding scores on the pretest as a best-guess – and indicating no progress – under the condition that only one value in that step was missing. If more values were missing, the corresponding subscore was also classified as missing data. Total scores on the posttest were treated the same: if more than one of the four subscores was missing, they were classified as missing value, otherwise the total was calculated over the remaining subscores.

*4.2.1. Pretest and posttest scores*

Inter-rater agreement on the scoring rubric for pre- and posttest was measured with a two-way mixed, absolute, single measure intra-class correlation and amounted to 0.878, indicating a reliable measure. Students scored rather low on the pretest, achieving a mean score of 41.86% ($SD = 9.86$). The scores varied between 18.75% and 62.5%. On the posttest, the mean score improved to 60.55% ($SD = 11.16$) with a range from 31.25% to 81.25%. Table 2 shows the mean scores per condition for the pretest and posttest. The repeated measures analysis showed that the between-subjects factor was not statistically significant: $F(3, 92) = 0.97$, $p = .410$, meaning that there was no effect of support and the scores did not depend on the type of support received. Indeed, the mean scores in Table 2 reveal that the four groups show a similar progression. The within-subjects factor did reveal a significant effect: $F(1, 92) = 187.46$, $p = 0.000$, $\eta^2_{\text{partial}} = 0.671$, indicating there was a substantial effect of training on the test scores.

**Table 2**
Overview of scores (in percentage) per condition.

| Condition | Pretest (*SD*) | Posttest (*SD*) |
|---|---|---|
| EM | 43.75 (11.89) | 63.07 (10.19) |
| CS | 41.25 (9.02) | 62.25 (11.05) |
| CS + EM | 41.75 (8.79) | 58.50 (12.48) |
| Control | 40.89 (10.09) | 58.59 (10.56) |
| Total | 41.86 (9.86) | 60.55 (11.16) |

*4.2.2. Mental effort ratings*

The mental effort ratings showed a similar pattern: significant changes over learning tasks, but not between the conditions. The repeated measures analyses revealed no significant between-subjects effect $F(3, 90) = 0.64$, $p = 0.593$, but a significant within-subjects effect $F(3, 90) = 9.60$, $p = 0.000$, $\eta^2_{\text{partial}} = 0.100$. Contrast analysis further revealed that reported mental effort drops significantly from 5.21 ($SD = 2.03$) in learning task 3 to 4.36 ($SD = 1.89$) in learning task 4: $F(1, 90) = 18.14$, $p = .000$, $\eta^2_{\text{partial}} = 0.174$. Univariate ANOVAs per learning task revealed no differences between conditions. Fig. 3 shows mean mental effort ratings for each condition and each learning task.

*4.2.3. Time pressure ratings*

Analysis of time pressure showed that although scores were relatively high (all means above 5 on the 7-point scale), there were no within-subjects differences: $F(3, 89) = 1.01$, $p = 0.391$ or between-subjects differences: $F(3, 89) = 0.16$, $p = 0.923$. Therefore, students experienced similar time pressure in all conditions and in all learning tasks. Univariate ANOVAs per learning task confirmed this finding: on all four learning tasks, differences between conditions were not statistically significant. Fig. 4 shows time pressure ratings for each condition and each learning task.

*4.3. Discussion*

This experiment was designed to explore whether the acquisition of IPS skills was affected by different forms of task support. However, the results show that all groups show similar increases in skill. These findings do not provide support for the hypotheses that 1) supported students show higher learning outcomes than unsupported students, and 2) two forms of support lead to higher learning outcomes than just one form of support. As a matter of fact, the control group, which merely received conventional tasks without any built-in support, performed just as well as the three groups who received task support. There was a significant increase in scores from pretest to posttest for all conditions, showing that the intervention clearly caused a learning effect. From this finding, it can be concluded that even a short online training, much like the training sessions often offered by schools, can be effective for fostering IPS skills. While the results clearly show a short-term learning effect, it is unclear whether there is potential to achieve a long-term effect. Additionally, the different types of support might have different effects on retention, which only manifest when measured after sufficient delay, or are induced by testing (i.e., a testing effect: Dirkx, Kester, & Kirschner, 2014). No such delayed

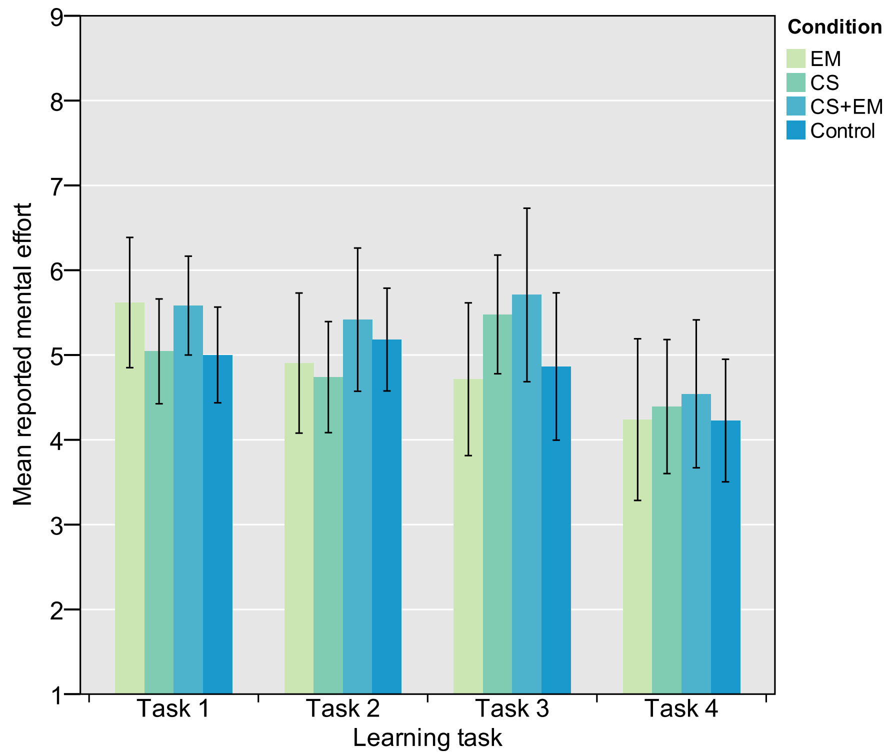


**Fig. 3.** Reported mental effort per learning task for all conditions.

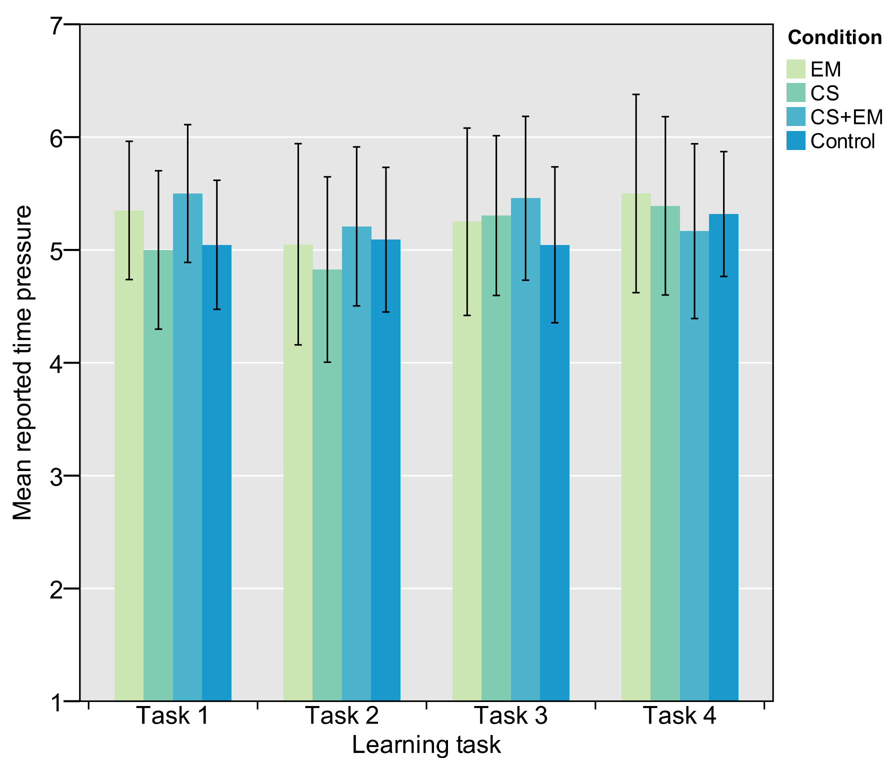


**Fig. 4.** Reported time pressure per learning task for all conditions.

measurement was undertaken in this experiment. It would therefore be interesting to investigate delayed learning effect with a retention test.

In line with these findings, students reported a similar amount of required mental effort in all conditions. For this category of students and in this particular setting, online learning tasks with or without built-in task support, whether that is completed steps or emphasized aspects, are equally demanding in terms of mental effort. From this self-report of mental effort, it is only possible to gauge the *total* amount of experienced cognitive demand, but not changes in the *underlying* types of cognitive load. If worked-out steps reduced intrinsic cognitive load but required students to invest additional mental effort to process and self-explain the worked-out steps, it replaced intrinsic with germane cognitive load and there might be no change in the total amount of experienced cognitive load (Paas, Tuovinen, Tabbers, & van Gerven, 2003). Similarly, if prompting leads to more extraneous cognitive load and less invested energy in learning, germane load is reduced but the total amount of cognitive load remains the same. However, when intrinsic or extraneous load is replaced with germane cognitive load, this hypothetically leads to increased learning (Van Merriënboer & Ayres, 2005). It is unlikely that this has happened, because increased learning would manifest as higher scores on the skills tests, which were not found. From the current data, the only valid conclusion is that the different types of support have no effect on the total amount of experienced cognitive load as measured by reported mental effort.

The high scores on the time pressure item revealed that many students experienced time pressure to finish the experiment. In the short evaluation at the end of the training 43 out of 96 participants made a remark about experienced time pressure. From their comments it became clear that the lack of time affected their concentration and performance during the learning tasks, or answer quality on the posttest. These students reported they took less time to think about and formulate their answers, thereby perhaps leaving out parts of the reasoning and missing points. This makes it likely that the learning outcomes are affected and possibly lowered because of time pressure. Given more time per task, students would perhaps have scored differently.

Inspection of students' solutions on the learning tasks revealed a great variation in answers. However, there was little instruction on presenting a solution incorporated in the training, so it cannot be expected that these outcomes correspond strongly to the level of their searching skills. Performance on the learning tasks was not part of the experimental design, and therefore, students' products were not scored and analyzed. For this reason it is not possible to comment on the students' performance during the learning phase.

## 5. Experiment 2

A second experiment was conducted with the same goal as the first experiment: to investigate differences in learning outcomes due to different types of task support. The same design and conditions were used as in the first experiment, but an additional questionnaire was used and a retention test was added. Some procedures were adapted to reduce time pressure.

### 5.1. Method

#### 5.1.1. Participants

A total of 115 students between 18 and 46 years old participated in the replication ($M_{\text{age}} = 20.7$ years), 82 of which were female (71.3%) and 33 male (28.7%). These were all first-year Psychology students at a Dutch university. Of these 115 students, three had a Belgian nationality (2.6%) and 48 were German (41.7%). The remainder was Dutch.

#### 5.1.2. Materials

*5.1.2.1. Measurement of information problem solving skill.* The same pretest and posttest were used as in Experiment 1, but a retention test was added. This retention test was identical to the existing pre- and posttests, but handled the topic of health benefits of red wine. Furthermore, a self-report questionnaire was added to the pretest, posttest and retention test.

*5.1.2.2. Self-report questionnaire.* The self-report questionnaire was based on an existing questionnaire (Van Meeuwen, 2008) and contained 30 items to measure students' systematic approach and evaluation behavior; for example: "I check whether a page is up-to-date before I use its information". Students responded to these items by selecting 'Never', 'Sometimes', 'Often', or 'Always'. The questionnaire included an 'I don't know' option to reduce guessing.

#### 5.1.3. Data analysis

The pretest, posttest, and retention test were scored as in Experiment 1 and subjected to a repeated measures analysis of variance with *type of support* (CS + EM vs CS vs EM vs control) as a between-subjects variable and *time of test* (pretest vs posttest vs retention test) as a within-subjects variable. Mental effort and time pressure ratings were analyzed with a repeated measures analysis of variance with *learning task* as a within-subjects variable. In addition, a univariate analysis of variance was conducted on the *mental effort* and *time pressure* items per learning task to investigate differences in required mental effort between conditions.

For the self-report scale, a principle component analysis with oblimin rotation was conducted on the 30-item scale in a larger sample size ($n = 250$) to extract underlying clusters and form scales. A mean value was calculated for each cluster by averaging the scores on the corresponding items. The 'I don't know' answer was treated as a missing value, and averages were only calculated if there was no more than one missing value. Scores were analyzed with a repeated measures analysis of variance.

#### 5.1.4. Procedure

As in Experiment 1, the training was embedded in the students' current curriculum as a practical assignment. Participation was voluntary, but strongly stimulated by granting research participation credit and informing students that the content of the training corresponded strongly to one of the course tasks about problem solving. The session was offered in eight different timeslots. Again, students were free to choose a timeslot that fit their schedule. Unlike in Experiment 1, the pretest was now administered in advance and was filled out at home, 1 week before the training. The retention test was also filled out at home, 1 week after the training. The length of the training session remained 2 h, which allowed students to spend approximately 20 min on each learning task; compared to 15 min in Experiment 1. Further procedures were identical to those in Experiment 1. After finishing the final evaluation, students signed a form to obtain research participation credit and were reminded to fill out the retention test after 1 week. They were then dismissed. A debriefing followed in a lecture 2 weeks after the retention test.

### 5.2. Results

As in Experiment 1, analysis of the answers on the preliminary questionnaire revealed a homogeneous group in terms of age and prior education. No notable differences arose in computer usage patterns or time spent behind the computer per day. Again, some data was missing, which was handled in the same way as in Experiment 1.

#### 5.2.1. Pretest, posttest, and retention test

The scores on the pretest ranged between 12.5% and 62.5% with a mean of 35.14% ($SD = 11.18$). For the posttest, scores ranged between 37.5% and 83.33% with a mean score of 61.58% ($SD = 11.15$). On the retention test the mean score was 60.6% ($SD = 13.73$) with a minimum score of 25% and a maximum score of 87.5%. Table 3

**Table 3**
Means and standard deviations of scores on the skills test (in percentages), systematic approach ratings (0–3), and evaluation behavior ratings (0–3) per condition on the pretest, posttest, and retention test.

| Condition | | Pretest | Posttest | Retention test |
|---|---|---|---|---|
| EM | Score | 34.72 (11.93) | 58.33 (11.88) | 58.80 (12.41) |
| | Systematic | 1.22 (0.36) | 1.28 (0.38) | 1.39 (0.43) |
| | Evaluation | 1.57 (0.42) | 1.76 (0.33) | 1.89 (0.44) |
| CS | Score | 34.25 (8.86) | 63.58 (9.14) | 60.50 (14.96) |
| | Systematic | 1.08 (0.41) | 1.22 (0.43) | 1.32 (0.45) |
| | Evaluation | 1.53 (0.47) | 1.78 (0.57) | 1.94 (0.42) |
| CS + EM | Score | 36.22 (12.63) | 63.06 (12.20) | 64.90 (13.93) |
| | Systematic | 1.20 (0.44) | 1.41 (0.41) | 1.32 (0.40) |
| | Evaluation | 1.74 (0.50) | 1.87 (0.47) | 1.93 (0.46) |
| Control | Score | 35.27 (12.41) | 61.09 (10.79) | 57.14 (14.00) |
| | Systematic | 1.24 (0.43) | 1.34 (0.39) | 1.33 (0.49) |
| | Evaluation | 1.52 (0.47) | 1.95 (0.35) | 1.92 (0.42) |
| Total | Score | 35.14 (11.18) | 61.58 (11.15) | 60.60 (13.73) |
| | Systematic | 1.19 (0.39) | 1.32 (0.39) | 1.34 (0.42) |
| | Evaluation | 1.57 (0.48) | 1.83 (0.46) | 1.93 (0.42) |

shows the mean scores per condition for the three tests. The results resemble those of the first experiment and show an increase in scores after training, but little difference between the conditions. The repeated measures analysis confirms that there was no significant difference between the groups: $F(3, 102) = 1.09$, $p = 0.358$ but a significant difference on the within-subjects factor: $F(2, 102) = 236.40$, $p < 0.001$, $\eta^2_{partial} = 0.699$. This confirms that there was a substantial effect of training on the test scores. A planned contrast revealed that the increase in scores from pretest to posttest was statistically significant: $F(1, 102) = 383.03$, $p < 0.001$, $\eta^2_{partial} = 0.790$, but the scores did not change significantly on the retention test: $F(1, 102) = 0.72$, $p = 0.400$. There were no significant interaction effects.

### 5.2.2. *Self-report questionnaire*

The Kaiser-Meyer-Olkin measure and sphericity measure indicated adequate sampling and sufficient correlations between items: KMO = 0.789, $\chi^2$ (435) = 1544.54, $p = 0.000$. An initial analysis of eigenvalues and interpretation of the scree plot justified retaining two components for the final analysis. Table 4 shows the factor loadings and correlations after rotation. These loadings create two clusters that can be labeled as *systematic approach* and *source evaluation behavior*. Six items were discarded: four with both loadings below 0.32 and two with equal factor loadings on both components (Tabachnick & Fidell, 2007). The scales yielded reliability scores of $\alpha = 0.85$ and $\alpha = 0.62$ respectively. See Table 3 for an overview of means and standard deviations for both variables.

For the *systematic approach* data, Mauchly's test revealed that the assumption of sphericity had been violated, $\chi^2(2) = 21.19$, $p = .000$. Therefore, the Huynh-Feldt correction was applied to the degrees of freedom. The test showed a significant increase in scores: $F(1.74, 99) = 13.58$, $p = 0.000$, but a small effect: $\eta^2_{partial} = 0.125$. Subsequent contrast analysis showed that scores increased significantly from pretest to posttest: $F(1, 99) = 16.78$, $p = 0.000$, $\eta^2_{partial} = 0.150$, but did not change significantly on the retention test. There were no significant differences between conditions: $F(3, 99) = 0.40$, $p = 0.756$. For the *evaluation behavior* data, the Huynh-Feldt adjustment was necessary as well: $\chi^2(2) = 8.40$, $p = .015$. Results show a significant within-subjects effect $F(1.93, 94) = 32.98$, $p = 0.000$, $\eta^2_{partial} = 0.268$, but no significant between-subjects effect: $F(3, 94) = 0.44$, $p = 0.726$. Contrast analysis shows a strong increase in scores from pretest to posttest: $F(1, 94) = 38.79$, $p = 0.000$, $\eta^2_{partial} = 0.301$, and another small increase on the posttest. The latter just fails to reach significance: $F(1, 94) = 3.90$, $p = 0.051$, $\eta^2_{partial} = 0.024$.

### 5.2.3. *Mental effort ratings*

The experienced mental effort during learning tasks shows a significant within-subjects effect: $F(3, 88) = 8.31$, $p = 0.000$, $\eta^2_{partial} = 0.090$, indicating that scores change significantly over time. However, a significant interaction effect reveals that the effect depends on the type of support the student received: $F(9, 88) = 2.74$, $p = 0.005$, $\eta^2_{partial} = 0.089$. Separate repeated measures ANOVAs for each condition showed significant effects only in the

**Table 4**
Exploratory factor analysis results for the IPS self-report: factor loadings (correlations).

| | Systematic approach | Evaluation behavior |
|---|---|---|
| I work according to a predetermined plan when searching, selecting, and processing information | **0.75** (0.72) | −0.12 (0.09) |
| I make an overview (a list or table) of the needed information | **0.72** (0.68) | −0.17 (0.03) |
| I plan where I am going to search for which information | **0.67** (0.61) | −0.23 (−0.05) |
| I make a list of steps to follow | **0.67** (0.62) | −0.19 (−0.00) |
| I mostly work intuitively and do not use a predetermined plan[a] | **0.66** (0.65) | −0.03 (0.16) |
| I make an overview of possible keywords | **0.61** (0.58) | −0.12 (0.05) |
| I just search for information without thinking about it too much[a] | **0.58** (0.57) | −0.05 (0.11) |
| I make a time schedule for performing the task | **0.57** (0.56) | −0.06 (0.10) |
| I systematically keep track of the keywords I have used | **0.51** (0.52) | 0.04 (0.18) |
| I regularly check whether I am searching correctly | **0.46** (0.49) | 0.12 (0.25) |
| While searching, I try to keep an overview of the search process | **0.45** (0.47) | 0.08 (0.21) |
| I deliberately check what I do not know yet in relation to the task | **0.43** (0.50) | 0.24 (0.26) |
| I present the information in an organized and ordered fashion | **0.42** (0.47) | 0.16 (0.28) |
| After visiting a site, I check which information is still needed | **0.41** (0.43) | 0.06 (0.17) |
| At the end, I check again whether I have all the information | **0.39** (0.45) | 0.24 (0.35) |
| I mostly work on and see how far I get[a] | **0.36** (0.45) | 0.31 (0.41) |
| I make sure that I organize all relevant information well | **0.35** (0.42) | 0.26 (0.36) |
| I keep the desired end product in mind | **0.33** (0.40) | 0.25 (0.34) |
| By looking at the URL (Uniform Resource Locator) I can see if a site is reliable | −0.23 (−0.06) | **0.63** (0.56) |
| To decide which site to open, I look at the URL (Uniform Resource Locator) | −0.13 (0.04) | **0.62** (0.58) |
| I check whether the site is up-to-date before I use the information | 0.01 (0.17) | **0.55** (0.56) |
| I check whether information I have found overlaps with previously found information | −0.05 (0.10) | **0.52** (0.51) |
| Before I open a site, I check its reliability | 0.11 (0.24) | **0.49** (0.52) |
| I check whether information I have found contradicts previously found information | 0.09 (0.22) | **0.47** (0.50) |

*Note.* Bold values indicate the item is included in the corresponding scale.
[a] Reverse-coded item.

EM condition: $F(3, 18) = 0.5.50$, $p = 0.002$, $\eta^2_{partial} = 0.244$, and in the CS condition: $F(3, 23) = 7.76$, $p = 0.000$, $\eta^2_{partial} = 0.261$. Subsequent contrast analysis indicated that the mental effort ratings in these groups only changed significantly on the fourth learning task. In the EM condition, scores dropped from 4.72 ($SD = 2.16$) on task 3 to 3.17 ($SD = 2.01$) on task 4: $F(1, 18) = 7.84$, $p = 0.012$, $\eta^2_{partial} = 0.316$. In the CS condition, scores dropped from 4.74 ($SD = 2.34$) to 2.83 ($SD = 1.64$): $F(1, 23) = 12.88$, $p = .002$, $\eta^2_{partial} = 0.369$. Fig. 5 shows mental effort ratings for each condition and each learning task.

#### 5.2.4. Time pressure ratings

Analysis of time pressure ratings revealed no significant changes over time and no differences between conditions. Separate univariate ANOVAs for each learning task showed that the average amount of time pressure on each learning task was the same in each group. Fig. 6 shows time pressure ratings for each condition and each learning task.

### 5.3. Discussion

The second experiment replicated the first with some improvements. First, it measured additional variables with a self-report questionnaire to achieve a more complete impression of the students' skill level. Second, it set out to reduce the experienced time pressure by administering the pretest before the training session. And finally, it included a retention test to measure IPS skill one week after training. With these improvements, the findings display a similar pattern as in the first experiment. The significant increase in scores from pretest to posttest leads to the conclusion that the intervention was effective for fostering IPS skills. However, the results do not back the claim that the type of support has an effect on the learning outcomes. None of the groups that received support, whether completion strategy, emphasis manipulation, or both, outperformed the control group.

This was also true for scores the self-report questionnaires. For *systematic approach*, students scored around 1.19 on the pretest, a value closer to 'Sometimes' than to 'Often', indicating that students are aware that they do not work very systematically when solving information problems. This score showed a small increase to an average of 1.32 on the posttest. While statistically significant, the effect of the training is small, and type of support again showed no effect. For *evaluation behavior*, a similar pattern emerges, but with larger effects. Average scores increase from 1.59 before training to 1.84 after the training, showing a large effect size. From these results it can concluded that the training significantly improved students' scores on self-reported systematic approach and source evaluation behavior, but again, there were no significant differences between the conditions. This corroborates previous research that shows evaluation skills can be trained in classroom settings (Britt & Aglinskas, 2002; Walraven et al., 2010).

In general, scores on the retention test results show a similar picture for all measured variables. While they increase from pretest to posttest, they do not change much one week later. All the differences between posttest scores and retention test scores are statistically insignificant and show small effect sizes. However, some conditions show a small increase in scores after a week, while others show a decrease in scores. It would be interesting to see if this difference develops into a significant effect over a longer period of time. From these findings, it can be concluded that the learning effect caused by this intervention is sufficiently robust to last one week.

Compared to Experiment 1, the mean reported mental effort and time pressure is generally lower. This is an expected finding as students in Experiment 2 were given more time to perform the learning tasks. On the fourth learning task - a conventional problem without support or guidance - the CS and the EM conditions reported significantly less mental effort than the CS + EM and control conditions. This might be a hint that these students have become more efficient in their problem solving and require less mental

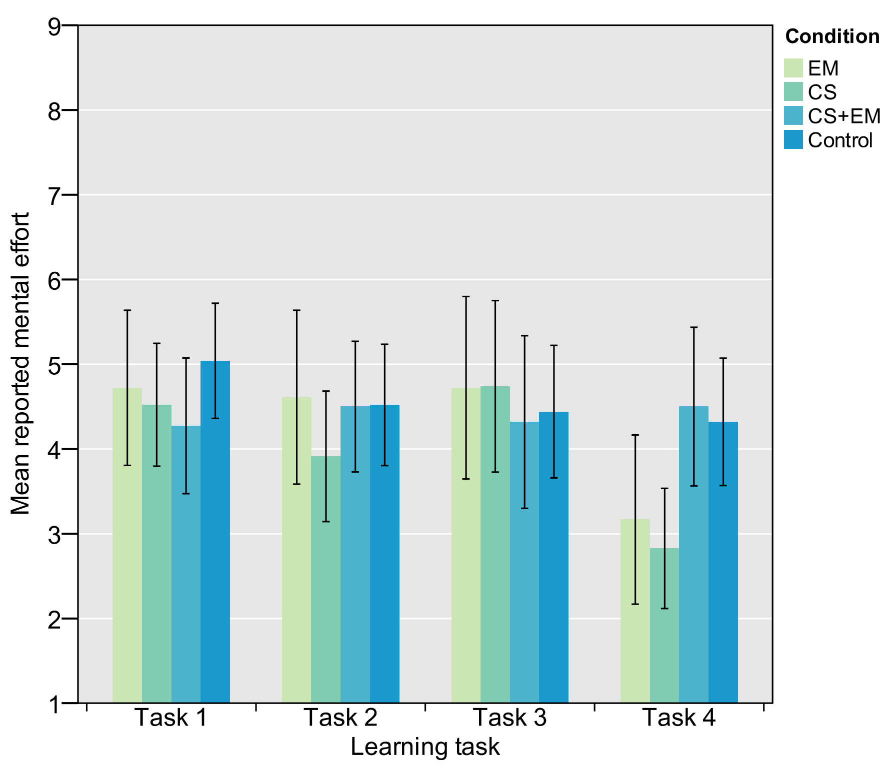


**Fig. 5.** Reported mental effort per learning task for all conditions.

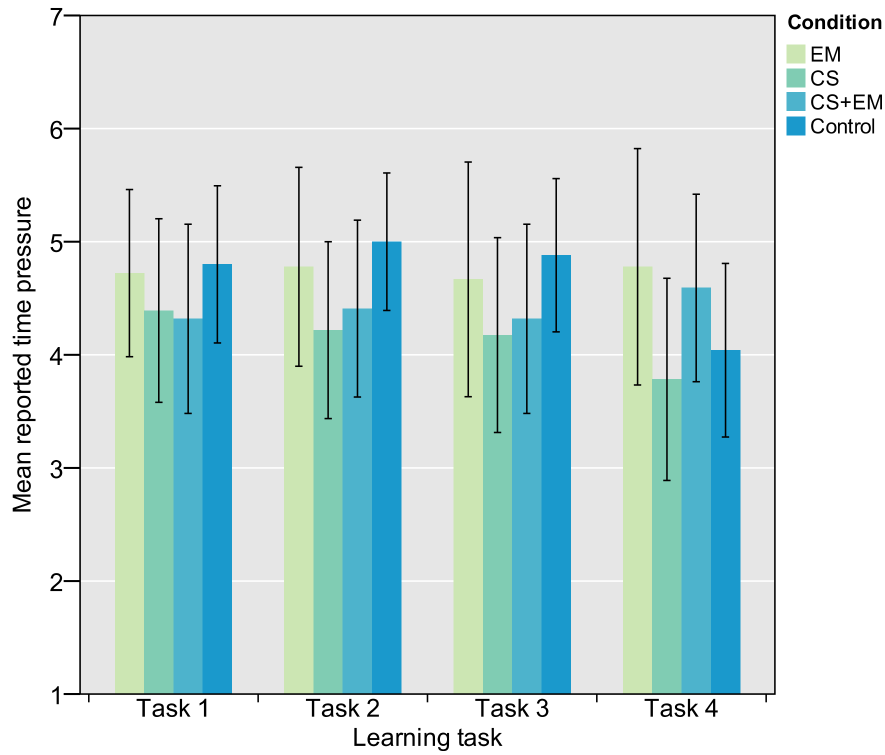


**Fig. 6.** Reported time pressure per learning task for all conditions.

effort to reach the solution. However, without performance data on the learning task, this is impossible to determine (Hoffman & Schraw, 2010). The subsequent posttest did not show any differences in performance between the conditions.

In short, Experiment 2 yielded no support for the hypothesis that supported students show higher learning outcomes than students who receive no support. Mean scores in all conditions did not differ significantly. While the EM and CS conditions reported less mental effort on a conventional learning task at the end of training, it is difficult to draw any solid conclusions from this finding.

## 6. General discussion

The experiments reported on here investigated the hypothesis that students who receive task support while acquiring IPS skills, either in the form of completion tasks or emphasis prompts, show better learning outcomes than students who do not receive task support. The findings do not support this claim. Students who receive no task support performed just as well as those who did. While Experiment 1 suffered from some methodological issues, a revised version of the experiment confirms the pattern of results and provides more confidence in this conclusion.

These findings have some implications for the domain of information problem solving. The current experiment once again confirms that IPS skills are underdeveloped, even in university-level students. Pretest scores are low in both experiments. In fact, the slightly younger group of students in the first experiment scored higher on the pretest than their counterparts in the second experiment. This difference shows a discrepancy in prior knowledge between both samples. While the exact cause of this is unclear, these differences likely originate from prior experience, practice, or instructions concerning IPS skills, such as a library training. However, the most important conclusion to draw from these findings is that this generation of first-year university students do not show very well-developed IPS skills. The scores, which lie well in the lower half of the range, can only refute claims that students are ‘digital natives’, a new generation technologically skilled students in need of adapted education. These findings agree with research challenging the existence of the digital native (Bennett, Maton, & Kervin, 2008; Kirschner & van Merriënboer, 2013; Smith, 2012) and underline our claim that IPS instruction in schools is a necessity.

The good news is that the current experiments show that a short online intervention can increase IPS skills. The online training session was successful, as shown by the significant increase in scores between pretest and posttest. After the training, students from both experiments scored slightly over 60% on average, which leads to two conclusions. First, a 2-h online training including an instructional video, a modeling example, and four short whole-tasks can increase students' IPS skills. As shown by retention test scores, this increase is maintained for at least a week. Second, effect sizes are not very large, and a 60% average score after training indicates that there is still much room to grow. However, the encouraging result of this short training indicates that a scaled-up version with more content, more task classes containing tasks of increasing complexity, offered over a longer period of time and embedded in a multitude of contexts, might prove very effective.

The findings of these experiments also lead to implications for the field of instructional design. Concerning the effect of built-in task support, the hypothesis that task support would lead to better performance was not confirmed: students who received no support showed performance equal to that of supported students. There are two possible explanations for this. First, it might be the case that both forms of support were ineffective for different reasons. Previous research has shown that completion tasks can lead to an expertise reversal effect in situations where learners have high prior knowledge (Kalyuga et al., 2003). However, this effect is less likely to occur in less structured domains (Nievelstein, van Gog, van Dijck, & Boshuizen, 2013), which, in combination with the low pretest scores, makes it unlikely that an expertise reversal effect occurred. The other method of support, prompting, can be ineffective when prompts are not used as intended, in which case they show reduced effects on learning outcomes and reported mental effort (Bannert & Reimann, 2011). Although answers to the prompts were generally short (i.e., approximately one sentence), they indicated that the prompts were used as anticipated. These findings lead to the conclusion that the task support methods were implemented correctly.

The second explanation suggests that a maximum learning effect for this setting was achieved. It could be the case that the learning effect in this experiment can be attributed to the viewing of the instructional video in combination with the modeling example and self-explanation prompts. Modeling examples are very powerful learning tools when employed correctly (Bjerrum, Hilberg, van Gog, Charles, & Eika, 2013; Hoogerheide, Loyens, & van Gog, 2014). Perhaps, after viewing both videos, students were sufficiently equipped to complete the learning tasks, and had no need of support. It follows then that the built-in support in those learning tasks had little value. A video-based modeling example is intuitively a very suitable method of instruction for teaching these skills, as most of the IPS process happens on-screen. An expert can easily record a screencast while working and reasoning through a problem and offer this as an example to students. The effects of using a modeling example for teaching IPS skills presents an interesting venue for future research.

In the context of this short online training, task support did not lead to higher learning outcomes. However, Rosman et al. (2016) show that working memory capacity moderates the acquisition of IPS skills. In a holistic approach to learning IPS, task support, such as completion tasks or prompts for emphasis, is essential to avoid overloading the learner during complex task performance. However, in situations where the learner's skill level is sufficient or the tasks are less complex, it might have no beneficial effects. Therefore, the results of these experiments should not convince instructional designers that task support does not matter. Instead, it should stimulate them to seek closer alignment of the learner's skill level, task complexity, and built-in task support. When designing instruction for IPS on a larger scale and over a longer period of time with increasing levels of complexity, managing the cognitive load imposed on the learner remains a crucial aspect of instructional design.

Several limitations of these experiments should be regarded when interpreting and generalizing these conclusions. The IPS training was offered in a single 2-h session with learning tasks of the same type and complexity. In educational practice, students are confronted with a great variety of tasks. The current intervention did not include different task types (c.f. Gerjets & Hellenthal-Schorr, 2008), which makes it less likely that far transfer occurred. To achieve far transfer, students would benefit from more learning tasks: more practice with varying task demands and task complexity, yet without added time pressure. An embedded approach, where instructional designers combine IPS instruction with domain-specific instruction in an extensive curriculum, appears appropriate for this task (Argelagós & Pifarré, 2012; Wopereis, Brand-Gruwel, & Vermetten, 2008).

The current intervention focused on learning in an online environment without involvement of a teacher and without feedback on the learning tasks. Considering the multitude of factors that can increase task complexity and cognitive demand, personalized feedback on performance would be beneficial for students, as this allows them to learn from their mistakes. Research has shown a positive effect of feedback on development of metacognitive skills in online learning environments (Van den Boom, Paas, van Merriënboer, & van Gog, 2004) and therefore presents another interesting direction for future research. For example, interventions could be improved with the addition of a cognitive feedback element in which teachers provide students with adapted feedback on their performance (Timmers et al., 2015; Wopereis et al., 2015).

To conclude, this experiment makes clear that first-year university students are not as information literate as many assume, and that their IPS skills need to be trained. The 2-h online intervention in this experiment shows a promising learning effect. While it was expected that different types of task support would vary in their effect on the learning outcomes, this proved not to be the case. The authors hypothesize that a powerful modeling example is most likely responsible for a large proportion of the learning effect and increased students' skill level, thereby reducing the value of the task support in the subsequent learning tasks. A follow-up study will investigate whether modeling examples are indeed a powerful learning tool for IPS.

## Appendix 1. Scoring rubric for information problem solving assessment

| **Question 1: What is your first step and why?** | |
|---|---|
| Maximum points: 2 | |
| 0 points | for statements that reflect that the student starts searching right away |
| Add 1 point | for statements reflecting orientation activities: activating prior knowledge, planning, thinking, etc. |
| Add 1 point | for statements concerning task demands: determining information needs, types of sources, formulating a question, etc. |

| **Question 2: Which problem statements would you formulate?** | |
|---|---|
| Maximum points: 2 | |
| 0 points | for statements that are irrelevant for the task |
| 1 point | for statements that are relevant, but incomplete or formulated vaguely |
| 2 points | for statements that contain all three relevant concepts (comparable to "*What is the influence of X on Y?*") |

| **Question 3: Which search query would you type into Google?** | |
|---|---|
| Maximum points: 4 | |
| Award a point for each relevant search term or synonym thereof. If the student shows a systematic search pattern, award an additional point. | |
| Pretest | gender-specific education, influence, school performance |
| Posttest | intelligence, change, age |
| Retention test | red wine, health, influence |

| **Question 4: Which three websites would you select? Why?** | |
|---|---|
| Maximum points: 4 | |
| Pretest | sources #3, #4, and #7 yield 2 points, sources #6 and #8 yield 1 point. |
| Posttest | sources #4, #5, and #6 yield 2 points, sources #3 and #8 yield 1 point. |
| Retention test | sources #3, #6, and #8 yield 2 points, sources #4 and #5 yield 1 point. |
| If the sum of these points is 5 or 6, award 2 points for this question. | |
| If the sum of these points is 2, 3, or 4, award 1 point for this question. | |
| If the sum of these points is lower than 2, award no points for this question. | |
| Award an additional point, but no more than 2 points, for all selection criteria that are mentioned in the comment that do not refer to "relevance". For example: reliability, author, publication date, reputation, etc. | |

| | |
|---|---|
| **Question 5: What do you do when you visit a text-rich website and want to find out if it contains relevant information?** | |
| Maximum points: 1 | |
| 1 | point |
| for | mentioning a scanning strategy, such as reading headlines only or using the search function (Ctrl + F) |

| | |
|---|---|
| **Question 6: Which criteria do you use to determine whether information is useful for your task?** | |
| Maximum points: 2 | |
| 1 point for each of the following criteria | goal of the text, reliability, author reputation, publication date, language/style, compares to other sources |

| | |
|---|---|
| **Question 7: How do you deal with contradicting information?** | |
| Maximum points: 1 | |
| 1 | point |
| For | statements that reflect critical scrutiny, for example searching for more information or investigating reliability, or if the answer reflects that both sides of the story are incorporated in the solution. |

| | |
|---|---|
| **Calculating the score** | |
| Subscore for step 1: Define the problem | The sum of scores for questions 1 & 2 |
| Subscore for step 2: Search information | The score for question 3 |
| Subscore for step 3: Select information | The score for question 4 |
| Subscore for step 4: Process information | The sum of scores for questions 5, 6, & 7 |
| **Total score:** | The average of these four subscores forms the final score for the test and is expressed as a percentage of the maximum score (4 points) |

## References


Argelagós, E., & Pifarré, M. (2012). Improving information problem solving skills in secondary education through embedded instruction. *Computers in Human Behavior, 28*, 515–526. http://dx.doi.org/10.1016/j.chb.2011.10.024.

Atkinson, R., Renkl, A., & Merrill, M. (2003). Transitioning from studying examples to solving problems: effects of self-explanation prompts and fading worked-out steps. *Journal of Educational Psychology, 95*, 774–783. http://dx.doi.org/10.1037/0022-0663.95.4.774.

Badke, W. (2010). Why information literacy is invisible. *Communications in Information Literacy, 4*, 129–141. http://dx.doi.org/10.1109/ICELMACH.2010.5608206.

Bannert, M., & Reimann, P. (2011). Supporting self-regulated hypermedia learning through prompts. *Instructional Science, 40*, 193–211. http://dx.doi.org/10.1007/s11251-011-9167-4.

Bennett, S., Maton, K., & Kervin, L. (2008). The "digital natives" debate: a critical review of the evidence. *British Journal of Educational Technology, 39*, 775–786. http://dx.doi.org/10.1111/j.1467-8535.2007.00793.x.

Bjerrum, A. S., Hilberg, O., van Gog, T., Charles, P., & Eika, B. (2013). Effects of modelling examples in complex procedural skills training: a randomised study. *Medical Education, 47*, 888–898. http://dx.doi.org/10.1111/medu.12199.

Brand-Gruwel, S., Wopereis, I., & Vermetten, Y. (2005). Information problem solving by experts and novices: analysis of a complex cognitive skill. *Computers in Human Behavior, 21*, 487–508. http://dx.doi.org/10.1016/j.chb.2004.10.005.

Brand-Gruwel, S., Wopereis, I., & Walraven, A. (2009). A descriptive model of information problem solving while using internet. *Computers & Education, 53*, 1207–1217. http://dx.doi.org/10.1016/j.compedu.2009.06.004.

Britt, M. A., & Aglinskas, C. (2002). Improving students' ability to identify and use source information. *Cognition and Instruction, 20*, 485–522. http://dx.doi.org/10.1207/S1532690XCI2004_2.

Chevalier, A., Dommes, A., & Marquié, J.-C. (2015). Strategy and accuracy during information search on the web: effects of age and complexity of the search questions. *Computers in Human Behavior, 53*, 305–315. http://dx.doi.org/10.1016/j.chb.2015.07.017.

De Vries, B., van der Meij, H., & Lazonder, A. W. (2008). Supporting reflective web searching in elementary schools. *Computers in Human Behavior, 24*, 649–665. http://dx.doi.org/10.1016/j.chb.2007.01.021.

DeStefano, D., & LeFevre, J. A. (2007). Cognitive load in hypertext reading: a review. *Computers in Human Behavior, 23*, 1616–1641. http://dx.doi.org/10.1016/j.chb.2005.08.012.

Dirkx, K. J. H., Kester, L., & Kirschner, P. A. (2014). The testing effect for learning principles and procedures from texts. *The Journal of Educational Research, 107*, 357–364. http://dx.doi.org/10.1080/00220671.2013.823370.

Fabiani, M., Buckley, J., Gratton, G., Coles, M. G. H. H., Donchin, E., & Logie, R. (1989). The training of complex task performance. *Acta Psychologica, 71*, 259–299. http://dx.doi.org/10.1016/0001-6918(89)90012-7.

Ford, N., Miller, D., & Moss, N. (2005). Web search strategies and human individual differences: cognitive and demographic factors, internet attitudes, and approaches. *Journal of the American Society for Information Science and Technology, 56*, 741–756. http://dx.doi.org/10.1002/asi.20168.

Gerjets, P., & Hellenthal-Schorr, T. (2008). Competent information search in the world wide web: development and evaluation of a web training for pupils. *Computers in Human Behavior, 24*, 693–715. http://dx.doi.org/10.1016/j.chb.2007.01.029.

Gopher, D. (2006). Emphasis change as a training protocol for high-demand tasks. In A. F. Kramer, D. A. Wiegmann, & A. Kirlik (Eds.), *Vol. 101. Attention: From theory to practice* (pp. 209–224). http://dx.doi.org/10.1093/acprof:oso/9780195305722.003.0015.

Gopher, D., Weil, M., & Siegel, D. (1989). Practice under changing priorities: an approach to the training of complex skills. *Acta Psychologica, 71*, 147–177. http://dx.doi.org/10.1016/0001-6918(89)90007-3.

Hart, S. G., & Staveland, L. E. (1998). Development of NASA-TLX (task load index): results of empirical and theoretical research. In P. Hancock, & N. Meshkati (Eds.), *Human mental workload* (pp. 139–183). Amsterdam: North Holland.

Hoffman, B., & Schraw, G. (2010). Conceptions of efficiency: applications in learning and problem solving. *Educational Psychologist, 45*, 1–14. http://dx.doi.org/10.1080/00461520903213618.

Hoogerheide, V., Loyens, S. M. M., & van Gog, T. (2014). Effects of creating video-based modeling examples on learning and transfer. *Learning and Instruction, 33*, 108–119. http://dx.doi.org/10.1016/j.learninstruc.2014.04.005.

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist, 38*, 23–31. http://dx.doi.org/10.1207/S15326985EP3801_4.

Kalyuga, S., & Sweller, J. (2004). Measuring knowledge to optimize cognitive load factors during instruction. *Journal of Educational Psychology, 96*, 558–568. http://dx.doi.org/10.1037/0022-0663.96.3.558.

Kammerer, Y., Bråten, I., Gerjets, P., & Strømsø, H. I. (2012). The role of Internet-specific epistemic beliefs in laypersons' source evaluations and decisions during web search on a medical issue. *Computers in Human Behavior, 29*, 1193–1203. http://dx.doi.org/10.1016/j.chb.2012.10.012.

Kennedy, G. E., Judd, T. S., Churchward, A., Gray, K., & Krause, K. L. (2008). First year students' experiences with technology: are they really digital natives? *Australasian Journal of Educational Technology, 24*, 108–122. doi:10.1.1.85.9526.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: an analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist, 41*, 75–86. http://dx.doi.org/10.1207/s15326985ep4102_1.

Kirschner, P. A., & van Merriënboer, J. J. G. (2009). Ten steps to complex learning. In T. L. Good (Ed.), *21st century education: A reference handbook* (pp. 244–253). Thousand Oaks, CA: Sage. Retrieved from http://hdl.handle.net/1820/2327.

Kirschner, P. A., & van Merriënboer, J. J. G. (2013). Do learners really know best? Urban legends in education. *Educational Psychologist, 48*, 169–183. http://dx.doi.org/10.1080/00461520.2013.804395.

Kuiper, E., Volman, M., & Terwel, J. (2008). Integrating critical Web skills and content knowledge: development and evaluation of a 5th grade educational program. *Computers in Human Behavior, 24*, 666–692. http://dx.doi.org/10.1016/j.chb.2007.01.022.

Lazonder, A. W. (2000). Exploring novice users' training needs in searching information on the WWW. *Journal of Computer Assisted Learning, 16*(4), 326–335. http://dx.doi.org/10.1046/j.1365-2729.2000.00145.x.

Lazonder, A. W., Biemans, H. J. A., & Wopereis, I. G. J. H. (2000). Differences between novice and experienced users in searching information on the World Wide Web. *Journal of the American Society for Information Science and Technology, 51*, 576–581. http://dx.doi.org/10.1002/(sici)1097-4571(2000)51, 6<576::aid-asi9>3.0.co;2–7.

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. http://dx.doi.org/10.1023/a:1014617530297.

Monchaux, S., Amadieu, F., Chevalier, A., & Mariné, C. (2015). Query strategies during information searching: effects of prior domain knowledge and complexity of the information problems to be solved. *Information Processing & Management, 51*, 557–569. http://dx.doi.org/10.1016/j.ipm.2015.05.004.

Nievelstein, F., van Gog, T., van Dijck, G., & Boshuizen, H. P. A. (2013). The worked example and expertise reversal effect in less structured tasks: learning to reason about legal cases. *Contemporary Educational Psychology, 38*, 118–125. http://dx.doi.org/10.1016/j.cedpsych.2012.12.004.

Paas, F. G. W. C. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: a cognitive-load approach. *Journal of Educational Psychology, 84*, 429–434. http://dx.doi.org/10.1037/0022-0663.84.4.429.

Paas, F., Tuovinen, J. E., Tabbers, H., & van Gerven, P. W. M. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist, 38*, 63–71. http://dx.doi.org/10.1207/S15326985EP3801_8.

Paas, F. G. W. C., & van Merriënboer, J. J. G. (1994). Instructional control of cognitive load in the training of complex cognitive tasks. *Educational Psychology Review, 6*, 351–371. http://dx.doi.org/10.1007/BF02213420.

Renkl, A. (1997). Learning from worked-out examples: a study on individual differences. *Cognitive Science, 21*, 1–29. http://dx.doi.org/10.1207/s15516709cog2101_1.

Renkl, A., & Atkinson, R. K. (2002). Learning from examples: fostering self-explanations in computer-based learning environments. *Interactive Learning Environments, 10*, 105–119. http://dx.doi.org/10.1076/ilee.10.2.105.7441.

Renkl, A., & Atkinson, R. (2003). Structuring the transition from example study to problem solving in cognitive skill acquisition: a cognitive load perspective. *Educational Psychologist, 38*, 15–22. http://dx.doi.org/10.1207/s15326985ep3801_3.

Renkl, A., Atkinson, R., & Große, C. (2004). How fading worked solution steps works – a cognitive load perspective. *Instructional Science, 32*, 59–82. http://dx.doi.org/10.1023/B: TRUC.0000021815.74806.f6.

Rosman, T., Mayer, A.-K., & Krampen, G. (2016). A longitudinal study on information-seeking knowledge in psychology undergraduates: exploring the role of information literacy instruction and working memory capacity. *Computers & Education, 96*, 94–108. http://dx.doi.org/10.1016/j.compedu.2016.02.011.

Rouet, J.-F. (2009). Managing cognitive load during document-based learning. *Learning and Instruction, 19*, 445–450. http://dx.doi.org/10.1016/j.learninstruc.2009.02.007.

Rouet, J.-F., Ros, C., Goumi, A., Macedo-Rouet, M., & Dinet, J. (2011). The influence of surface and deep cues on primary and secondary school students' assessment of relevance in web menus. *Learning and Instruction, 21*, 205–219. http://dx.doi.org/10.1016/j.learninstruc.2010.02.007.

Saito, H., & Miwa, K. (2007). Construction of a learning environment supporting learners' reflection: a case of information seeking on the web. *Computers & Education, 49*(2), 214–229. http://dx.doi.org/10.1016/j.compedu.2005.07.001.

Smith, E. E. (2012). The digital native debate in higher education: a comparative analysis of recent literature. *Canadian Journal of Learning and Technology, 38*(3), 1–18.

Stadtler, M., & Bromme, R. (2008). Effects of the metacognitive computer-tool met.a.ware on the web search of laypersons. *Computers in Human Behavior, 24*, 716–737. http://dx.doi.org/10.1016/j.chb.2007.01.023.

Stark, R., & Krause, U. M. (2009). Effects of reflection prompts on learning outcomes and learning behaviour in statistics education. *Learning Environments Research, 12*, 209–223. http://dx.doi.org/10.1007/s10984-009-9063-x.

Tabachnick, B. G., & Fidell, L. S. (2007). Principal components and factor analysis. In *Using multivariate statistics* (5th ed., pp. 607–675). Boston: Pearson Education.

Timmers, C. F., Walraven, A., & Veldkamp, B. P. (2015). The effect of regulation feedback in a computer-based formative assessment on information problem solving. *Computers & Education, 87*, 1–9. http://dx.doi.org/10.1016/j.compedu.2015.03.012.

Van Deursen, A. J. A. M., & van Diepen, S. (2013). Information and strategic internet skills of secondary students: a performance test. *Computers & Education, 63*, 218–226. http://dx.doi.org/10.1016/j.compedu.2012.12.007.

Van Deursen, A. J. A. M., & van Dijk, J. A. G. M. (2009). Using the internet: skill related problems in users' online behavior. *Interacting with Computers, 21*, 393–402. http://dx.doi.org/10.1016/j.intcom.2009.06.005.

Van Gog, T., Paas, F., & van Merriënboer, J. J. G. (2004). Process-oriented worked examples: improving transfer performance through enhanced understanding. *Instructional Science, 32*, 83–98. http://dx.doi.org/10.1023/B:TRUC.0000021810.70784.b0.

Van Meeuwen, L. W. (2008). *Tracking the brain*. Master's thesis. Eindhoven: Technische Universiteit Eindhoven. Retrieved from http://repository.tue.nl/634895.

Van Merriënboer, J. J. G. (1990). Strategies for programming instruction in high school: program completion vs. program generation. *Journal of Educational Computing Research, 6*, 265–285. http://dx.doi.org/10.2190/4NK5-17L7-TWQV-1EHL.

Van Merriënboer, J. J. G. (2013). Perspectives on problem solving and instruction. *Computers & Education, 64*, 153–160. http://dx.doi.org/10.1016/j.compedu.2012.11.025.

Van Merriënboer, J. J. G., & Ayres, P. (2005). Research on cognitive load theory and its design implications for e-learning. *Educational Technology Research and Development, 53*, 5–13. http://dx.doi.org/10.1007/BF02504793.

Van Merriënboer, J. J. G., & De Croock, M. B. M. (1995). Strategies for computer-based programming instruction: program completion vs. program generation. *Journal of Educational Computing Research, 8*, 365–394. http://dx.doi.org/10.2190/MJDX-9PP4-KFMT-09PM.

Van Merriënboer, J. J. G., & Kirschner, P. A. (2013). *Ten steps to complex learning: A systematic approach to four-component instructional design* (2nd ed.). New York: Routledge.

Van Strien, J. L. H., Brand-Gruwel, S., & Boshuizen, H. P. A. (2014). Dealing with conflicting information from multiple nonlinear texts: effects of prior attitudes. *Computers in Human Behavior, 32*, 101–111. http://dx.doi.org/10.1016/j.chb.2013.11.021.

Van den Boom, G., Paas, F., van Merriënboer, J. J. G., & van Gog, T. (2004). Reflection prompts and tutor feedback in a web-based learning environment: effects on students' self-regulated learning competence. *Computers in Human Behavior, 20*, 551–567. http://dx.doi.org/10.1016/j.chb.2003.10.001.

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. http://dx.doi.org/10.1016/j.chb.2007.01.030.

Walraven, A., Brand-Gruwel, S., & Boshuizen, H. P. A. (2010). Fostering transfer of websearchers' evaluation skills: a field test of two transfer theories. *Computers in Human Behavior, 26*, 716–728. http://dx.doi.org/10.1016/j.chb.2010.01.008.

Wirth, W., Sommer, K., von Pape, T., & Karnowski, V. (2015). Success in online searches: difference between evaluation and finding tasks. *Journal of the Association for Information Science and Technology*. http://dx.doi.org/10.1002/asi.23389.

Wopereis, I., Brand-Gruwel, S., & Vermetten, Y. (2008). The effect of embedded instruction on solving information problems. *Computers in Human Behavior, 24*, 738–752. http://dx.doi.org/10.1016/j.chb.2007.01.024.

Wopereis, I., Frerejean, J., & Brand-Gruwel, S. (2015). Information problem solving instruction in higher education: a case study on instructional design. In S. Kurbanoğlu, J. Boustany, S. Špiranec, E. Grassian, D. Mizrachi, & L. Roy (Eds.), *Communications in computer and information science: Vol. 552. Information literacy: Moving toward sustainability* (pp. 293–302). Basel, Switzerland: Springer. http://dx.doi.org/10.1007/978-3-319-28197-1.

Yechiam, E., Erev, I., & Gopher, D. (2001). On the potential value and limitations of emphasis change and other exploration-enhancing training methods. *Journal of Experimental Psychology: Applied, 7*, 277–285. http://dx.doi.org/10.1037/1076-898x.7.4.277.

Yechiam, E., Erev, I., Yehene, V., & Gopher, D. (2003). Melioration and the transition from touch-typing training to everyday use. *Human Factors, 45*, 671–684. http://dx.doi.org/10.1518/hfes.45.4.671.27085.