Ifenthaler, D. (2012). Determining the effectiveness of prompts for self-regulated learning in problem-solving scenarios. *Educational Technology & Society*, *15* (1), 38–52.

Determining the effectiveness of prompts for self-regulated learning in problem-solving scenarios

Dirk Ifenthaler
Fakultät für Sozialwissenschaften, Universität Mannheim, D-68159 Mannheim, Germany // ifenthaler@uni-mannheim.de

ABSTRACT

Cognitive scientists have studied internal cognitive structures, processes, and systems for decades in order to understand how they function in human learning. In order to solve challenging tasks in problem situations, learners not only have to perform cognitive activities, e.g., activating existing cognitive structures or organizing new information, they also have to set specific goals, plan their activities, monitor their performance during the problem-solving process, and evaluate the efficiency of their actions. This paper reports an experimental study with 98 participants where effective instructional interventions for self-regulated learning within problem-solving processes are investigated. Furthermore, an automated assessment and analysis methodology for determining the quality of learning outcomes is introduced. The results indicate that generic prompts are an important aid for developing cognitive structures while solving problems.

Keywords

Reflection, metacognition, prompting, HIMATT

Introduction

Self-regulated learning is regarded as one of the most important skills needed for life-long learning. Zimmerman (1989, p. 4) describes self-regulated learning as a process in which learners “are metacognitively, motivationally, and behaviorally active participants in their own learning process.” Hence, self-regulated learning is a complex process which involves numerous dimensions of human information processing (Azevedo, 2008, 2009; Pintrich, 2000; Schraw, 2007; Veenman, van Hout-Wolters, & Afflerbach, 2006; Zimmerman, 2008). Accordingly, in order to solve challenging tasks in problem situations, learners not only have to perform cognitive activities, e.g., activating existing knowledge structures or organizing new information (Seel, Ifenthaler, & Pirnay-Dummer, 2009), they also have to set specific goals, plan their activities, monitor their performance during the problem-solving process, and evaluate the efficiency of their actions (Wirth & Leutner, 2008).

Moreover, the facilitation of self-regulated learning is a balancing act between necessary *external support* and desired *internal regulation* (Koedinger & Aleven, 2007; Simons, 1992). From an instructional point of view, there are two vital ways to externally support self-regulated learning within problem-solving processes. *Direct* external support, in terms of direct instruction, aims at facilitating explicit problem-solving strategies and skills as well as their application and transfer to different domains. Hence, direct instruction could include detailed scaffolds (step-by-step instruction) on how to solve a specific phenomenon in question (Collins, Brown, & Newman, 1989). *Indirect* external support provides learning aids which induce and facilitate already existing problem-solving strategies and skills. Accordingly, if learners already possess comprehensive problem-solving strategies but fail to use this knowledge in a specific situation, it seems reasonable to motivate them to apply their existing strategic knowledge effectively (Lin & Lehmann, 1999). A possible instructional method for indirectly guiding and supporting the regulation of learners’ problem-solving processes is prompting (Wirth, 2009). In general, prompts are presented as simple questions (e.g., “What will be your first step when solving the problem?”), incomplete sentences (e.g., “To approach the solution to the problem step by step, I have to …”), explicit execution instructions (e.g., “First, draw the most important concepts and link them.”), or pictures and graphics for a specific learning situation (Bannert, 2009). Accordingly, well-designed and embedded prompts direct learners to perform a specific desired activity which is contextualized within a particular problem-solving situation (see Davis, 2003; Davis & Linn, 2000; Lin & Lehmann, 1999). According to Davis (2003), prompts can be categorized into *generic* and *directed* prompts. While the generic prompt only asks learners to stop and reflect about their current problem-solving activities, the directed prompt also provides them with an expert model of reflective thinking in the problem-solving process.

From a methodological point of view, we argue that it is essential to identify economic, fast, reliable, and valid techniques to assess and analyze these complex problem-solving processes. Especially within experimental setting where huge sets of data need to be processed, standard methodologies (e.g., paper and pencil tests) may have disadvantages with regard to analysis economy. Therefore, we developed an automated assessment and analysis technology, HIMATT (Highly Integrated Model Assessment Technology and Tools; Pirnay-Dummer, Ifenthaler, & Spector, 2010), which combines qualitative and quantitative research methods and provides bridges between them.

In our current research we are investigating effective instructional interventions for self-regulated learning within problem-solving processes (e.g. Ifenthaler, 2009; Ifenthaler, Masduki, & Seel, 2011). Hence, the present study was conducted to explore and evaluate different types of prompts for self-regulated learning in a problem-solving scenario. Furthermore, we introduce an automated assessment and analysis methodology for determining the quality of learning outcomes.

Cognitive processes and problem solving

A central assumption of cognitive psychology is that mental representations enable individuals to understand and explain experience and events, process information, and solve problems (Johnson-Laird, 1989). More specifically, Rumelhart, Smolensky, McClelland, and Hinton (1986) argue that these internal functions of the human mind are dependent on two interacting modules or sets of units: *Schemata* and *mental models*. In this context, schemata and mental models are theoretical constructs which specify different functions of human information processing. The resulting cognitive architecture corresponds to a great extent to Piaget's epistemology (1943, 1976) and its basic mechanisms of assimilation and accommodation.

Accordingly, assimilation is dependent on the availability and activation of schemata, which allow new information to be integrated immediately into pre-existing cognitive structures. As soon as a schema can be activated, it runs automatically and regulates information processing. If a schema does not fit immediately into the requirements of a new problem-solving task it can be adjusted to meet them by means of accretion, tuning, or reorganization (Seel, et al., 2009). Accordingly, if a schema for any problem type is available, it is promptly mapped onto the problem to be solved (Jonassen, 2000). If assimilation is not successful, accommodation must take place in order to reorganize or restructure an individual's knowledge. However, when no schema is available at all or when its reorganization fails, the human mind switches to the construction of a mental model, which is defined as a dynamic ad hoc representation of a phenomenon or problem that aims at creating subjective plausibility through the simplification or envisioning of the situation, analogical reasoning, or mental simulation.

We further argue that a learner constructs a mental model by integrating relevant bits of domain-specific knowledge into a coherent structure step by step in order to meet the requirements of a phenomenon to be explained or a problem to be solved. From an instructional point of view, providing direct or indirect external support within this step-by-step process could be an effective way to guide learners through problem-solving processes and facilitate their self-regulated learning in the long run. Winne (2001) provides an in-depth discussion on the above introduced concepts.

The role of metacognition and reflection in problem solving

Various researchers have highlighted the importance of metacognition for the adjustment and the regulation of learning and problem-solving activities (e.g., Boekaerts, 1999; Mayer, 1998; Schmidt-Weigand, Hänze, & Wodzinski, 2009; Zimmerman & Schunk, 2001). According to Pintrich (2000), metacognition is defined as a superordinate ability to direct and regulate cognitive, motivational, and behavioral learning and problem-solving processes in order to achieve a specific goal. Generally, researchers distinguish between two major components of metacognition, namely *knowledge of cognition* and *regulation of cognition*. Knowledge of cognition includes declarative knowledge about the self as a learner and problem-solving strategies, procedural knowledge about how to use these strategies, and conditional knowledge about when and why to use them – this metacognitive knowledge is also referred to as metacognitive awareness. Regulation of cognition, on the other hand, refers to components which facilitate the control and regulation of learning. These skills involve abilities such as planning, self-monitoring, and self-evaluation (Schraw & Dennison, 1994).

But how do learners transfer their knowledge of effective problem solving to regulate their problem-solving activities? In general, the key link between knowledge about and the regulation of one's own problem-solving activities is assumed to be *reflective thinking* (see Ertmer & Newby, 1996). If learners manage to generate information about the efficiency of their problem-solving strategies and successfully implement these findings in the ongoing problem-solving process, they are able to control and regulate their cognitive activities. Thus, metacognition refers to the ability to reflect on, understand, and control one's learning and problem-solving activities (Simons, 1993). Accordingly, we have to distinguish between three different levels of learner-orientated reflective thinking: (1) a problem-based reflection of the learning content, (2) a behavior-oriented reflection of one's own problem-solving activities, and (3) the learner's identity-based reflection of his or her own learning ability. While the superordinate level of reflection requires the progressive verification of existing beliefs and established practices of one's own learning, the behavior-oriented reflection takes place in the wake of experience (see Jenert, 2008). Furthermore, according to Wirth (2009, p. 91), "teaching learning regulation means to regulate the learner's learning regulation." This leads to the question of how to support learners' reflection through instruction.

Supporting learners' reflection via prompting

The instructional goal of teaching self-regulated problem solving is a highly demanding task (Wirth, 2009). It requires supporting learners in the acquisition and application of strategic knowledge for effective problem solving. The self-regulated learner possesses a set of problem-solving strategies and most importantly the ability to transfer and to apply this knowledge to different problem situations. In the course of their development from novice to expert, learners need guidance to learn how to regulate their problem-solving activities. Accordingly, the type of instructional aid depends on the state of the learner (grade of self-regulation). Novice learners (in terms of their self-regulation abilities) may need stronger guidance whereas expert learners do need less or no guidance at all. Hence, this decrease of strength in guidance could be described as fading of guidance (Collins, et al., 1989). On the other hand, learners also need a certain extent of autonomy to self-regulate their problem-solving activities in terms of learning by doing. The problem of accomplishing a balance between support and autonomy is referred to as the "assistance dilemma" (Koedinger & Aleven, 2007, p. 239). Additionally, in order to provide an optimal balance between external assistance and the facilitation of autonomous learning, it is necessary to distinguish between *ability deficiency* and *production deficiency* (Veenman, Kerseboom, & Imthorn, 2000). Learners with an ability deficiency suffer from a lack of metacognitive knowledge and skills. Accordingly, teachers have to convey problem-solving strategies to the learners and provide them with opportunities to exercise and reflect on their knowledge. In the case of a production deficiency, learners actually possess the knowledge and skills to regulate their problem-solving processes. However, they fail to use the inert knowledge and skills in specific problem-solving situations. In such cases, instructional support can be reduced to the activation of knowledge and skills in order to not restrict the learners in their autonomy.

Prompting is an instructional method for guiding and supporting the regulation of the learner's problem solving processes. Prompts are presented as simple questions (e.g., "What will be your first step when solving the problem?"), incomplete sentences (e.g., "To approach the solution to the problem step by step, I have to …"), execution instructions (e.g., "First, draw the most important concepts and link them."), or pictures and graphics (Bannert, 2007, 2009). The main goal of the method is to focus the learner's attention on specific aspects of his or her own problem-solving process. By activating learners and motivating them to think about the efficiency of their strategies, one can increase their awareness for mostly unconsidered problem-solving activities. Therefore, they reflect on their own thoughts and are able to monitor, control, and regulate their strategic procedure in a specific situation (see Chi, Bassok, Lewis, Reimann, & Glaser, 1989; Chi, De Leeuw, Chiu, & Lavancher, 1994; Davis, 2003; Davis & Linn, 2000; Ertmer & Newby, 1996; Ge & Land, 2004; Lin & Lehmann, 1999). The best point in time to present a prompt depends on the intention of the specific intervention. Learners should receive the prompt just in time, i.e. at the moment in which they require external support. Otherwise, these short interventions might result in cognitive overload (Thillmann, Künsting, Wirth, & Leutner, 2009). In general, a distinction is made between presentation before, during, or after a learning sequence. If the prompt is intended to activate the learners to monitor their problem-solving activities, presentation during a learning sequence is reasonable. If the intention is to induce the learners to assess certain problem-solving activities, presentation after the sequence is appropriate. Presenting the prompt before a problem-solving sequence is expedient when one wishes to inspire the learners to generate an approach to the problem-solving situation (Davis, 2003). Another crucial aspect is how metacognitive prompts can be designed and embedded to provide an optimal scaffold to the learners. Davis (2003) investigated the efficiency of reflective prompts and differentiates between generic and directed prompts. Her primary interest was to explore whether learners merely need to be prompted to reflect or need more guidance in order to reflect productively.

Accordingly, the presentation of generic prompts would seem to be more effective, because the learner's autonomy is not undermined. The directed prompt, on the other hand, additionally asks learners to process more information, because it introduces a new expert model for reflection (see Davis, 2003).

To sum up, prompting is an instructional method that guides learners during problem-solving processes. Well-designed and embedded prompts may direct learners to perform a specific desired activity, which is contextualized within a particular problem-solving situation. Accordingly, more empirical evidence is needed to investigate which type of prompting leads to a better performance (generic vs. directed; see Davis, 2003).

New ways of assessment and analysis

Cognitive and educational researchers use theoretical constructs, e.g. metacognition, mental models, schemata, etc., to explain complex cognitive structures and procedures for learning, reasoning, and problem solving (Seel, et al., 2009). However, these internal cognitive structures and functions are not directly observable, which leads to biased assessment and analysis. Accordingly, the assessment and analysis of internal cognitive structures and functions requires that they be externalized. Therefore, we argue that it is essential to identify economic, fast, reliable, and valid techniques to elicit and analyze these cognitive structures (see Ifenthaler, 2008, 2010b). Appropriate standard methodologies include standardized questionnaires and interviews (Zimmerman, 2008), think-aloud protocols (Ericsson & Simon, 1993), the assessment of log files or click streams (Chung & Baker, 2003; Dummer & Ifenthaler, 2005; Veenman, Wilhelm, & Beishuizen, 2004), and eye-tracking measures (Mikkilä-Erdmann, Penttinen, Anto, & Olkinuora, 2008) as well as mind tools (Jonassen & Cho, 2008). However, the possibilities of externalization are limited to a few sets of sign and symbol systems (Seel, 1999b) – characterized as *graphical-* and *language-based approaches* (Ifenthaler, 2010b). A widely accepted application is concept, causal, or knowledge maps which are automatically scored and compared to an expert's solution (Herl, Baker, & Niemi, 1996; Ifenthaler, 2010a).

However, current discussion about the above-described methodological options suggests that it will be necessary to find new assessment and analysis alternatives (Ifenthaler, 2008; Seel, 1999a; Veenman, 2007; Veenman, et al., 2006). As not every available methodology is suitable for this research, we have introduced our own web-based assessment and analysis platform, HIMATT (Highly Integrated Model Assessment Technology and Tools; Pirnay-Dummer, et al., 2010).

HIMATT is a combined toolset which was developed to convey the benefits of various methodological approaches in a single environment and which can be used by researchers with only little prior training (Pirnay-Dummer & Ifenthaler, 2010). Methodologically, the tools integrated into HIMATT touch the boundaries of qualitative and quantitative research methods and provide bridges between them. First of all, text can be analyzed very quickly without loosening the associative strength of natural language. Furthermore, concept maps can be annotated by experts and compared to other solutions. The automated analysis function produces measures which range from surface-oriented structural comparisons to integrated semantic similarity measures. There are four *structural* (surface, graphical, structural, and gamma matching) and three *semantic* (concept, propositional, and balanced propositional matching) measures available (see the Method section for a detailed description of them). All of the data, regardless of how it is assessed, can be analyzed quantitatively with the same comparison functions for all built-in tools without further manual effort or recoding. Additionally, HIMATT generates standardized images of text and graphical representations (Pirnay-Dummer & Ifenthaler, 2010; Pirnay-Dummer, et al., 2010).

Research questions and hypotheses

The central research objective of this study is to identify the efficiency of different types of prompts (generic vs. directed) for activating learners to reflect on their ongoing problem-solving process. Based on prior research (Davis, 2003; Ge & Land, 2004), we hypothesized that learners who receive generic prompts during the problem-solving process will perform better than those who receive directed prompts. Accordingly, a generic prompt provides learners necessary support and allows them a certain extent of autonomy to self-regulate their problem-solving activities (Koedinger & Aleven, 2007). Hence, we assume that learners who receive generic prompts will perform better with regard to their domain-specific understanding (Hypothesis 1). If learners do not already possess the required self-regulative knowledge and skills, directed prompts would be more effective. Additionally, we assume that the problem representations (in the form of a concept map) of learners with generic prompts will be structurally (Hypothesis 2) and semantically (Hypothesis 3) more similar to an expert's solution than those of learners who have received directed prompts.

Additionally, previous research studies have found contradictory results concerning learners' metacognitive processes and deductive reasoning skills in association with learning outcomes when working with concept maps in problem solving scenarios (e.g. Hilbert & Renkl, 2008; Ifenthaler, Pirnay-Dummer, & Seel, 2007; O'Donnell, Dansereau, & Hall, 2002; Veenman, et al., 2004). We assume that learners with higher metacognitive awareness will outperform those with lower metacognitive awareness with regard to their learning outcomes (Hypothesis 4a). Additionally, we assume that better deductive reasoning skills will have a positive effect on the learning outcomes (Hypothesis 4b).

Method

Participants

Ninety-eight students (68 female and 30 male) from a European university participated in the study. Their average age was 21.9 years ($SD = 3.5$). They were all enrolled in an introductory course on research methods and had studied for an average of 2.4 semesters ($SD = 3.1$).

Design

Participants were randomly assigned to the three experimental conditions. The three experimental conditions were related to the three forms of reflective thinking prompt: *generic prompt* (GP; $n_1 = 32$), *direct prompt* (DP; $n_2 = 40$), and *control group* (CG; $n_3 = 26$). Participants in the GP group received general instructions for planning and reflecting on their ongoing problem-solving activities (see materials for details). For participants in the DP group, we provided nine sentences which referred to *planning* (1–3), *monitoring* (4–6), and *evaluation* (7–9) of the ongoing problem-solving activities (see materials for details). The CG did not receive a reflective thinking prompt. ANOVA was used to test for study experience differences (number of semesters studied) among the three experimental groups. The experimental groups did not differ with regard to the semesters studied, $F(2, 95) = 0.42$, $p > .05$.

Materials

Problem scenario

A German-language article on the human immune system and the consequences of virus infections with 1,120 words was used as learning content. The problem was to identify differences between an influenza and HIV infection. Specifically, the problem task consisted of the following two questions: (1) What happens to the immune system during an initial infection with the influenza virus? (2) What effect does an HIV infection have on the immune system in contrast to an influenza infection? Additionally, learners were asked to graphically represent their understanding of these complex biological processes (questions one and two) in form of a concept map. Also, an expert solution (based on the article) in the form of a concept map was generated which functioned as a reference model for later analysis.

Domain specific knowledge test

The knowledge test included 13 multiple-choice questions with four possible solutions each (1 correct, 3 incorrect). First, 20 questions were developed on the basis of the article on the human immune system and the consequences of virus infections. Second, in a pilot study ($N = 10$ participants), we tested the average difficulty level to account for ceiling effects. Finally, we excluded seven questions because they were not appropriate for our experimental study. In our experiment we administered two versions (in which the 13 multiple-choice questions appeared in a different order) of the domain-specific knowledge test (pre- and posttest). It took about eight minutes to complete the test.

Metacognitive awareness inventory

The participants' metacognitive awareness was assessed with the *Metacognitive Awareness Inventory* (Schraw & Dennison, 1994). Each of the 52 items of the inventory was answered on a scale from 1 to 100 (Cronbach's alpha = .90). Two dimensions of metacognitive awareness were addressed: (1) *knowledge of cognition*, which includes knowledge about personal skills, learning strategies, and the efficiency of these strategies, and (2) *regulation of cognition*, which includes planning and initiating of learning, implementation of strategies, monitoring and control of learning, and the evaluation of personal learning efficiency.

Deductive reasoning inventory

A subscale of the ASK (Analyse des Schlussfolgernden und Kreativen Denkens; i.e. inventory for deductive reasoning and creative thinking) was used to test the participants' deductive reasoning (Schuler & Hell, 2005). The subscale included questions on the *interpretation of information* (21 items), *drawing conclusions* (32 items), and *facts and opinions* (27 items). Schuler and Hell (2005) report good reliability scores for the ASK (Cronbach's alpha = .72; test-retest reliability = .78).

Experience with concept mapping test

The participants' experience with concept mapping was tested with a questionnaire including eight items (Ifenthaler, 2009; Cronbach's alpha = .87). The questions were answered on a five-point Likert scale (1 = totally disagree; 2 = disagree; 3 = partially agree; 4 = agree; 5 = totally agree). Items included in the test, e.g. "I use concept maps to structure learning content", "The construction of a concept map raises no difficulties", or "I use computer software for constructing concept maps" (translated from German).

Reflective thinking prompts

Two versions of prompts were developed in order to stimulate the participants to reflect on their problem-solving activities. (1) The *generic prompt* („stop and reflect") included the following advice: "Use the next 15 minutes for reflection. Reflect critically on the course and outcome of your problem-solving process. Amend and improve your concept map if necessary. Feel free to use all materials provided! (translated from German)." (2) The *direct prompt* included the following advice: „Use the next 15 minutes for reflection. Reflect critically on the course and outcome of your problem-solving process. Feel free to use all materials provided! The guidelines provided below may be used as an aid. Please complete the list item by item by completing each sentence on its own in your mind. 1. The requirements/goals of the problem included...; 2. The basic conditions which had to be taking in account to complete this problem were...; 3. In order to find the best solution to the problem, I...; 4. In order to understand the context and main ideas of the text, I...; 5. In order to come a bit closer to the solution with each step, I...; 6. In order to create an optimal concept map of the text, I...; 7. I believe I solved the problem well, because...; 8. I could solve the problem better next time if I...; 9. In order to improve my explanation model I will now... (translated from German)".

HIMATT concept mapping tool

The concept mapping tool, which is part of the HIMATT (Pirnay-Dummer, et al., 2010) environment, was used to assess the participants' understanding of the problem scenario. The intuitive web-based tool allows participants to create concept maps with only little training (Pirnay-Dummer & Ifenthaler, 2010). Once created, all concept maps are automatically stored on the HIMATT database for further analysis.

Procedure

First, the participants were randomly assigned to the three experimental conditions (GP, DP, CG). Then they completed a *demographic data survey* (three minutes), the *metacognitive awareness inventory* (ten minutes), the *deductive reasoning inventory* (33 minutes), and the *experience with concept mapping test* (five minutes). Next, the participants were given an introduction to concept maps and were shown how to use the HIMATT environment (ten minutes). After a short relaxation phase (five minutes), they answered the 13 multiple choice questions of the domain-specific knowledge test on the immune system and the consequences of virus infections (pretest; eight minutes). Then they received the article on the immune system and the consequences of virus infections and were introduced into the problem scenario. In total, all participants spent 25 minutes on the problem scenario. Additionally, participants in the experimental condition GP and DP received their reflective thinking prompt after 15 minutes working on the problem scenario. The CG did not receive a reflective thinking prompt. They were allowed to take notes with paper and pencil. After another short relaxation phase (five minutes), the participants logged into the HIMATT environment and constructed a concept map on their understanding of the problem scenario (ten minutes). Finally, the participants answered the 13 multiple choice questions of the posttest on declarative knowledge (eight minutes).

Data analysis

In order to analyze the participants' understanding of the problem scenario, we used the seven measures implemented in HIMATT (see Ifenthaler, 2010b; Pirnay-Dummer, et al., 2010). Accordingly, each of the participants' concept maps was compared automatically against the reference map (expert solution based on the article). Table 1 describes the seven measures of HIMATT, which include four structural measures and three semantic measures (Ifenthaler, 2010a, 2010b; Pirnay-Dummer & Ifenthaler, 2010; Pirnay-Dummer, et al., 2010).
HIMATT uses specific automated comparison algorithms to calculate similarities between a given pair of frequencies f_1 (e.g. expert solution) and f_2 (e.g. participant solution). The similarity s is generally derived by

$$s = 1 - \frac{|f_1 - f_2|}{\max(f_1, f_2)}$$

which results in a measure of $0 \leq s \leq 1$, where $s = 0$ is complete exclusion and $s = 1$ is identity. The other measures collect sets of properties. In this case, the Tversky similarity (Tversky, 1977) applies for the given sets A (e.g. expert solution) and B (e.g. participant solution):

$$s = \frac{f(A \cap B)}{f(A \cap B) + \alpha \cdot f(A - B) + \beta \cdot f(B - A)}$$

α and β are weights for the difference quantities which separate A and B. They are usually equal ($\alpha = \beta = 0.5$) when the sources of data are equal. However, they can be used to balance different sources systematically, e.g. comparing a learner's concept map which was constructed within five minutes to an expert's concept map, which may be an illustration of the result of a conference or of a whole book (see Pirnay-Dummer & Ifenthaler, 2010). The Tversky similarity also results in a measure of $0 \leq s \leq 1$, where $s = 0$ is complete exclusion and $s = 1$ is identity.

Reliability scores exist for the single measures integrated into HIMATT. They range from $r = .79$ to $r = .94$ and are tested for the semantic and structural measures separately and across different knowledge domains (Pirnay-Dummer, et al., 2010). Validity scores are also reported separately for the structural and semantic measures. Convergent validity lies between $r = .71$ and $r = .91$ for semantic comparison measures and between $r = .48$ and $r = .79$ for structural comparison measures (Pirnay-Dummer, et al., 2010).

Table 1. Description of the seven HIMATT measures

Measure [abbreviation] and type	*Short description*
Surface matching [SFM] *Structural indicator*	The surface matching (Ifenthaler, 2010a) compares the number of vertices within two graphs. It is a simple and easy way to calculate values for surface complexity.
Graphical matching [GRM] *Structural indicator*	The graphical matching (Ifenthaler, 2010a) compares the diameters of the spanning trees of the graphs, which is an indicator for the range of conceptual knowledge. It corresponds to structural matching as it is also a measure for structural complexity only.

Structural matching [STM] *Structural indicator*	The structural matching (Pirnay-Dummer & Ifenthaler, 2010) compares the complete structures of two graphs without regard to their content. This measure is necessary for all hypotheses which make assumptions about general features of structure (e.g. assumptions which state that expert knowledge is structured differently from novice knowledge).
Gamma matching [GAM] *Structural indicator*	The gamma or density of vertices (Pirnay-Dummer & Ifenthaler, 2010) describes the quotient of terms per vertex within a graph. Since both graphs which connect every term with each other term (everything with everything) and graphs which only connect pairs of terms can be considered weak models, a medium density is expected for most good working models.
Concept matching [CCM] *Semantic indicator*	Concept matching (Pirnay-Dummer & Ifenthaler, 2010) compares the sets of concepts (vertices) within a graph to determine the use of terms. This measure is especially important for different groups which operate in the same domain (e.g. use the same textbook). It determines differences in language use between the models.
Propositional matching [PPM] *Semantic indicator*	The propositional matching (Ifenthaler, 2010a) value compares only fully identical propositions between two graphs. It is a good measure for quantifying semantic similarity between two graphs.
Balanced propositional matching [BPM] *Semantic indicator*	The balanced propositional matching (Pirnay-Dummer & Ifenthaler, 2010) is the quotient of propositional matching and concept matching. In specific cases (e.g., when focusing on complex causal relationships), balanced propositional matching could be preferred over propositional matching.

Results

Initial data checks showed that the distributions of ratings and scores satisfied the assumptions underlying the analysis procedures. All effects were assessed at the .05 level. As effect size measures, we used Cohen's d (small effect: $d < .50$, medium effect $.50 \leq d \leq .80$, strong effect $d > .80$) and partial η^2 (small effect: $\eta^2 < .06$, medium effect $.06 \leq \eta^2 \leq .13$, strong effect $\eta^2 > .13$).

More than half of the participants (58%) did not use concept maps to structure their own learning materials before our experiment. Only 5% of the participants used concept mapping software to create their own concept maps beforehand. On the other hand, over 60% of the participants answered that they did not find it difficult to create a concept map. Consequently, there was no significant difference in the learning outcome as measured by the domain-specific knowledge posttest between participants who used concept mapping software before the experiment and those who did not use concept mapping software at all, $t(96) = .105$, *ns*.

Domain-specific knowledge

On the domain-specific knowledge test (pre- and posttest), participants could score a maximum of 13 correct answers. In the pretest they scored an average of $M = 4.38$ correct answers ($SD = 1.71$) and in the posttest $M = 6.71$ correct answers ($SD = 2.49$). The increase in correct answers was significant, $t(97) = 9.611$, $p < .001$, $d = 1.068$. ANOVA was used to test for knowledge gain differences among the three experimental groups. The experimental groups did not differ with regard to the results in the pretest, $F(2, 95) = 2.14$, $p > .05$. However, the increase in correct answers differed significantly across the three experimental groups, $F(2, 95) = 8.21$, $p = .001$, $\eta^2 = .147$. Tukey HSD post-hoc comparisons of the three groups indicate that the generic prompt group ($M = 3.66$, $SD = 2.40$, *95% CI* [2.79, 4.52]) gained significantly more correct answers than the directed prompt group ($M = 1.68$, $SD = 2.14$, *95% CI* [.99, 2.36]), $p = .001$, and the control group ($M = 1.73$, $SD = 2.20$, *95% CI* [.84, 2.62]), $p = .005$. Comparisons between the directed prompt group and the control group were not statistically significant at $p < .05$. Accordingly, the results support the hypothesis that participants who receive generic prompts outperform those in other groups with regard to their domain-specific understanding.

HIMATT structural measures

The participants' understanding of the problem scenario as illustrated by concept maps was analyzed automatically with the HIMATT tool. The four structural measures reported in Table 2 show the average similarity between the participants' solution and the referent solution (expert concept map). Four separate ANOVAs (for HIMATT measures SFM, GRM, STM, GAM) with Tukey HSD post-hoc comparisons were computed to test for differences between the three experimental groups.

ANOVA revealed a significant difference between participants in the three experimental groups for the HIMATT measure STM, $F(2, 95) = 7.77$, $p = .001$, $\eta^2 = .141$. Tukey HSD post-hoc comparisons of the three groups indicate that the complete structure (STM) of the generic prompt group's concept maps ($M = .84$, $SD = .14$, *95% CI* [.79, .89]) was significantly more similar to the expert solution than that of the directed prompt group's maps ($M = .70$, $SD = .14$, *95% CI* [.66, .75]), $p = .001$. Additionally, the complete structure (STM) of the control group's concept maps ($M = .80$, $SD = .19$, *95% CI* [.73, .88]) was significantly more similar to the expert solution than that of the directed prompt group's maps, $p = .026$. Comparisons between the directed prompt group and the control group were not statistically significant at $p < .05$.

For the HIMATT measure GAM, ANOVA revealed a significant difference between the three experimental groups, $F(2, 95) = 5.49$, $p = .006$, $\eta^2 = .104$. Tukey HSD post-hoc comparisons of the three groups indicate that the density of vertices (GAM) of the generic prompt group's concept maps ($M = .83$, $SD = .10$, *95% CI* [.79, .87]) was significantly more similar to the expert solution than that of the directed prompt group's maps ($M = .70$, $SD = .19$, *95% CI* [.64, .76]), $p = .004$. All other comparisons between groups were not statistically significant at $p < .05$.
ANOVAs for the HIMATT measures SFM and GRM revealed no significant differences between the experimental groups. Accordingly, the results support the hypothesis that participants who receive generic prompts outperform participants in other groups with regard to the HIMATT measures STM and GAM.

Table 2. Means (SD) HIMATT structural measures for the three experimental groups ($N = 98$)

	GP ($n_1 = 32$)	DP ($n_2 = 40$)	CG ($n_3 = 26$)
Surface matching [SFM]	.73 (.19)	.60 (.25)	.68 (.28)
Graphical matching [GRM]	.77 (.18)	.72 (.21)	.71 (.21)
Structural matching [STM]	.84 (.14)	.70 (.14)	.80 (.19)
Gamma matching [GAM]	.83 (.10)	.70 (.19)	.74 (.19)

Note. HIMATT similarity measures between participant's solution and expert's solution (0 = no similarity; 1 = total similarity); GP = generic prompt, DP = directed prompt, CG = control group

HIMATT semantic measures

Additional HIMATT analysis for the semantic measures of the participants' understanding of the problem scenario as expressed by concept maps was computed. The three semantic measures reported in Table 3 show the average similarity between the participants' solution and the referent solution (expert concept map). Three separate ANOVAs (for HIMATT measures CCM, PPM, BPM) with Tukey HSD post-hoc comparisons were computed to test for differences between the three experimental groups.

ANOVA revealed a significant difference between participants in the three experimental groups for the HIMATT measure CCM, $F(2, 95) = 7.40$, $p = .001$, $\eta^2 = .135$. Tukey HSD post-hoc comparisons of the three groups indicate that the semantic correctness of single concepts used in the concept maps (CCM) of the generic prompt group ($M = .43$, $SD = .19$, *95% CI* [.37, .50]) was significantly more similar to the expert solution than in those of the directed prompt group ($M = .30$, $SD = .14$, *95% CI* [.26, .34]), $p = .001$, and the control group ($M = .31$, $SD = .15$, *95% CI* [.25, .37]), $p = .011$. Comparisons between the directed prompt group and the control group were not statistically significant at $p < .05$.

ANOVA revealed a significant difference between participants in the three experimental groups for the HIMATT measure PPM, $F(2, 95) = 10.80$, $p < .001$, $\eta^2 = .185$. Tukey HSD post-hoc comparisons of the three groups indicate that the semantic correctness of propositions (concept-link-concept) used in the concept maps (PPM) of the generic prompt group ($M = .17$, $SD = .16$, *95% CI* [.11, .23]) was significantly more similar to the expert solution than in those of the directed prompt group (*M* = .06, *SD* = .06, *95% CI* [.04, .08]), $p < .001$, and the control group (*M* = .07, *SD* = .08, *95% CI* [.04, .10]), $p = .002$. Comparisons between the directed prompt group and the control group were not statistically significant at $p < .05$.

Table 3. Means (SD) HIMATT semantic measures for the three experimental groups ($N = 98$)

	GP ($n_1 = 32$)	DP ($n_2 = 40$)	CG ($n_3 = 26$)
Concept matching [CCM]	.43 (.19)	.30 (.14)	.31 (.15)
Propositional matching [PPM]	.17 (.16)	.06 (.06)	.07 (.08)
Balanced propositional matching [BPM]	.33 (.22)	.16 (.16)	.17 (.17)

Note. HIMATT similarity measures between participant's solution and expert's solution (0 = no similarity; 1 = total similarity); GP = generic prompt, DP = directed prompt, CG = control group

ANOVA revealed a significant effect between participants in the three experimental groups for the HIMATT measure BPM, $F(2, 95) = 8.97$, $p < .001$, $\eta^2 = .159$. Tukey HSD post-hoc comparisons of the three groups indicate that the quotient of the semantic correctness of propositions (concept-link-concept) and single concepts used in the concept maps (BPM) of the generic prompt group (*M* = .43, *SD* = .19, *95% CI* [.25, .41]) was significantly more similar to the expert solution than in those of the directed prompt group (*M* = .30, *SD* = .14, *95% CI* [.11, .21]), $p < .001$, and the control group (*M* = .31, *SD* = .15, *95% CI* [.11, .24]), $p = .004$. Comparisons between the directed prompt group and the control group were not statistically significant at $p < .05$. Accordingly, the results support the hypothesis that participants who receive generic prompts outperform participants in other groups with regard to the HIMATT measures CCM, PPM, and BPM.

Correlational analyses

Correlations were calculated between metacognitive awareness, deductive reasoning and the seven HIMATT measures as well as for the domain-specific knowledge of the posttest (see Table 4).

Table 4. Correlations between metacognitive awareness, deductive reasoning and HIMATT measures, domain specific knowledge (post-test)

	HIMATT structural measures				HIMATT semantic measures			
	Surface matching [SFM]	Graphical matching [GRM]	Structural matching [STM]	Gamma matching [GAM]	Concept matching [CCM]	Propositional matching [PPM]	Balanced propositional matching [BPM]	Domain specific knowledge
Metacognitive awareness								
Knowledge of cognition	.050	.002	.163	.104	.109	.114	.178	-.001
Regulation of cognition	.088	.050	.049	-.051	.037	.124	.177	.076
Deductive reasoning								
Interpretation of information	-.014	-.016	.072	-.030	.104	.068	.055	.351**
Drawing conclusions	.018	.029	.013	.070	-.118	-.055	-.034	.413**
Facts and opinions	.109	.099	.066	-.061	-.072	-.042	.022	.297**

Note. * $p < .05$; ** $p < .01$

Positive deductive reasoning abilities were related to better domain-specific knowledge. Accordingly, interpretation of information correlated significantly with the learning outcomes as measured by the domain-specific knowledge test, $r = .351$, $p < .01$. Apparently the learners' ability to interpret available information was associated positively with the domain-specific knowledge. Additionally, drawing conclusions correlated significantly with the learning outcomes, $r = .416$, $p < .01$. Hence, the learners' logical reasoning from given information was strongly associated with the domain-specific knowledge. Furthermore, "facts and opinions" correlated significantly with the learning outcomes, $r = .297$, $p < .01$. Accordingly, the learners' ability to differentiate between facts and opinions was positively associated with the domain-specific knowledge.

However, no correlations were found between metacognitive awareness and the domain-specific knowledge. Finally, no correlations were found between the HIMATT measures and metacognitive awareness or deductive reasoning (see Table 4).

Discussion

The facilitation of self-regulated learning is a balancing act between external support and internal regulation. An instructional method for guiding and supporting the regulation of learners' problem-solving processes is prompting. Prompts are presented as simple questions, incomplete sentences, explicit execution instructions, or pictures and graphics for a specific learning situation. Prompts are categorized in generic and directed forms. Generic prompts ask learners to stop and reflect about their current activities. Directed prompts additionally provide learners expert models of reflective thinking.

The aim of the present study was to explore the efficiency of different types of prompts for reflection in a self-regulated problem-solving situation. It was assumed that well-designed and embedded prompts may direct learners to perform successfully within a particular self-regulated problem-solving situation (Davis, 2003; Thillmann, et al., 2009). The problem was to identify differences between an influenza and HIV infection as well as their effects to the human immune system. In order to assess the participants' understanding of the problem scenario, we asked them to create a concept map on their subjectively plausible understanding of the phenomenon in question. Three experimental conditions with different reflective thinking prompts were realized. Participants in the generic prompt group (GP) received general instructions for planning and reflecting on their ongoing problem-solving activities. For participants in the direct prompt group (DP), we provided nine sentences which referred to planning, monitoring, and evaluation of the ongoing problem-solving activities. Participants in the control group (CG) did not receive a reflective thinking prompt.

In order to analyze the elicitation of the participants' understanding of the problem scenario, we introduced our own web-based platform HIMATT (Pirnay-Dummer & Ifenthaler, 2010; Pirnay-Dummer, et al., 2010). Within HIMATT, participants' concept maps can be automatically compared to a referent map created by an expert based on the problem scenario. The HIMATT analysis function produces measures which range from surface-oriented structural comparisons to integrated semantic similarity measures. Four structural measures (surface [SFM], graphical [GRM], structural [STM], and gamma [GAM]) and three semantic measures (concept [CCM], propositional [PPM], balanced propositional [BPM]) were used to answer our research questions.

Major findings of the present study are that participants in the generic prompt group outperformed other learners with regard to their (1) domain-specific knowledge gain as well as their (2) structural and (3) semantic understanding of the problem scenario.

First, findings on domain-specific knowledge suggest that generic prompts (e.g., What will be your first step when solving the problem?) are most effective in self-regulated learning environments. Generic prompts guide learners to use a specific set of problem-solving strategies and at the same time give them a certain extent of autonomy to self-regulate their problem-solving activities (Koedinger & Aleven, 2007). In contrast, direct prompts seem to prevent learners from solving a problem autonomously. However, we believe that direct prompts could be helpful for novices who do not yet possess the necessary problem-solving skills. Hence, further empirical investigations are necessary to answer these assumptions.

Second, generic prompts also had a positive effect on the structural similarity of learners' understanding of the problem scenario with regard to the expert solution. Compared to the expert solution, GP learners' solutions represented more strongly connected knowledge, which could indicate a deeper subjective understanding of the underlying subject matter (HIMATT measures STM, GAM). However, the number of concepts and links (SFM) and the overall complexity of the problem representations were not influenced by the different prompts. We believe that an effect towards complexity will occur in longer perspectives requiring an in-depth analysis of the learning-dependent change (Ifenthaler, et al., 2011; Ifenthaler & Seel, 2005).

Third, findings for the semantic HIMATT measures (CCM, PPM, BPM) are in line with the above-discussed results. Solutions of GP learners are semantically more similar to the expert solution than those of other learners. However, the overall similarity of the learners' problem representation to the expert representation is low. Hence, the learners of the present study are far from being experts and should be given more time and resources to improve their overall performance. Accordingly, we believe that further studies are needed to better understand the underlying cognitive processes of learning-dependent progression from novice to expert and, as a consequence, to provide more effective instructional materials.

Furthermore, correlational analysis showed that metacognitive awareness and deductive reasoning skills were not associated with the problem scenario representation as expressed by concept maps. These results complement previous research studies which have found similar results (e.g. Hilbert & Renkl, 2008; Ifenthaler, et al., 2007; O'Donnell, et al., 2002). However, we found significant correlations between domain-specific knowledge and deductive reasoning skills. Accordingly, deductive reasoning skills have positive effects on the declarative learning outcomes. One final consideration based on our findings is that when we train novices to become experts, we often think about training general abilities to efficiently facilitate the process. While this works well for training abilities themselves, these methods may have limits when we train experts who have to decide and act within complex domains (Chi & Glaser, 1985; Ifenthaler, 2009).

Reviewing the results of our experimental investigation, we suggest that a generic prompt which includes general instructions for planning and reflecting on their ongoing problem-solving activities are most effective for learners which already have a solid set of skills (in our case students at a university). In contrast, if learners do not have a specific set of problem-solving skills, directed prompts may be more effective. Accordingly, future studies should focus on the effectiveness of different prompts for different types of learners (novices, advanced learners, expert learners).

The present research is limited to the single problem scenario on differences between an influenza and HIV infection as well as their effects to the human immune system. The limited time and resources for solving the problem may have also had an influence on our results. Also, further empirical investigations should focus on the "best" point in time when to present a prompt (Thillmann, et al., 2009). In addition, the present research is limited by our use of concept maps to elicit the problem scenario. However, such graphical representations are a widely accepted method for illustrating the meaning of locally discussed information (Eliaa, Gagatsisa, & Demetriou, 2007; Hardy & Stadelhofer, 2006; Ruiz-Primo, Schultz, Li, & Shavelson, 2001). In order to improve the external validity of our research, we suggest applying additional methodologies such as think-aloud protocols (Ericsson & Simon, 1993), standardized questionnaires and interviews (Zimmerman, 2008), and log files or click streams (Chung & Baker, 2003; Veenman, et al., 2004) within multimedia learning environments. Especially thinking aloud protocols applied during the reflection phase could give more insights into the metacognitive procedures induced by different types of prompts. Lastly, the timing of the prompts should be investigated in future studies (Thillmann, et al., 2009). Accordingly, future studies will include not only prompts for reflecting on the problem-solving process but also reflection prompts provided before the learners enter the problem scenario.

Conclusions

To sum up, since cognitive and educational researchers are not able to measure internal cognitive structures and functions directly, studies like ours will always be biased. A major bias includes the limited possibilities for externalizing learners' internal cognitive structures (Ifenthaler, 2008, 2010b). However, we are adamant in our belief that it is essential to identify economic, fast, reliable, and valid methodologies to elicit and analyze these cognitive structures and functions (Zimmerman, 2008). In conclusion, new ways of assessment and analysis could make more precise results available, which may in turn lead to superior instructional interventions in the future.

References

Azevedo, R. (2008). The role of self-regulation in learning about science with hypermedia. In D. Robinson & G. Schraw (Eds.), *Recent innovations in educational technology that facilitate student learning* (pp. 127-156). Charlotte, NC: Information Age Publishing.

Azevedo, R. (2009). Theoretical, conceptual, methodological, and instructional issues in research on metacognition and self-regulated learning: A discussion. *Metacognition and Learning, 4*(1), 87-95.

Bannert, M. (2007). *Metakognition beim Lernen mit Hypermedia. Erfassung, Beschreibung, und Vermittlung wirksamer metakognitiver Lernstrategien und Regulationsaktivitäten*. Münster: Waxmann.

Bannert, M. (2009). Promoting self-regulated learning through prompts. *Zeitschrift für Pädagogische Psychologie, 23*(2), 139-145.

Boekaerts, M. (1999). Self-regulated learning: where we are today. *International Journal of Educational Research, 31*(6), 445-457.

Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 13*(2), 145-182.

Chi, M. T. H., De Leeuw, N., Chiu, M.-H., & Lavancher, C. (1994). Eliciting self-explanations improves understanding *Cognitive Science, 18*(3), 439-477.

Chi, M. T. H., & Glaser, R. (1985). Problem solving ability. In R. J. Sternberg (Ed.), *Human abilities: An information processing approach* (pp. 227-257). San Francisco, CA: W. H. Freeman & Co.

Chung, G. K. W. K., & Baker, E. L. (2003). An exploratory study to examine the feasibility of measuring problem-solving processes using a click-through interface. *Journal of Technology, Learning and Assessment, 2*(2), Available from http://www.jtla.org.

Collins, A., Brown, J. S., & Newman, S. E. (1989). Cognitive apprenticship: Teaching the crafts of reading, writing, and mathematics. In L. B. Resnick (Ed.), *Knowing, learning, and instruction* (pp. 453-494). Hillsdale, NJ: Lawrence Erlbaum.

Davis, E. (2003). Prompting middle school science students for productive reflection: Generic and directed prompts. *Journal of the Learning Sciences, 12*(1), 91-142.

Davis, E., & Linn, M. C. (2000). Scaffolding students' knowledge integration: Prompts for reflection in KIE. *International Journal of Science Education, 22*, 819-837.

Dummer, P., & Ifenthaler, D. (2005). Planning and assessing navigation in model-centered learning environments. Why learners often do not follow the path laid out for them. In G. Chiazzese, M. Allegra, A. Chifari & S. Ottaviano (Eds.), *Methods and technologies for learning* (pp. 327-334). Sothhampton: WIT Press.

Eliaa, I., Gagatsisa, A., & Demetriou, A. (2007). The effects of different modes of representation on the solution of one-step additive problems. *Learning and Instruction, 17*(6), 658-672.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data*. Cambridge, MA: MIT Press.

Ertmer, P. A., & Newby, T. J. (1996). The expert lerner: Strategic, self-regulated, and reflective. *Instructional Science, 24*(1), 1-24.

Ge, X., & Land, S. M. (2004). A conceptual framework for scaffolding ill-structured problem-solving processes using question prompts and peer interactions. *Educational Technology Research and Development, 52*(2), 5-22.

Hardy, I., & Stadelhofer, B. (2006). Concept Maps wirkungsvoll als Strukturierungshilfen einsetzen. Welche Rolle spielt die Selbstkonstruktion? *Zeitschrift für Pädagogische Psychologie, 20*(3), 175-187.

Herl, H. E., Baker, E. L., & Niemi, D. (1996). Construct validation of an approach to modeling cognitive structure of U.S. history knowledge. *Journal of Educational Research, 89*(4), 206-218.

Hilbert, T. S., & Renkl, A. (2008). Concept mapping as a follow-up strategy to learning from texts: what characterizes good and poor mappers? *Instructional Science, 36*, 53-73.

Ifenthaler, D. (2008). Practical solutions for the diagnosis of progressing mental models. In D. Ifenthaler, P. Pirnay-Dummer & J. M. Spector (Eds.), *Understanding models for learning and instruction. Essays in honor of Norbert M. Seel* (pp. 43-61). New York: Springer.

Ifenthaler, D. (2009). Model-based feedback for improving expertise and expert performance. *Technology, Instruction, Cognition and Learning, 7*(2), 83-101.

Ifenthaler, D. (2010a). Relational, structural, and semantic analysis of graphical representations and concept maps. *Educational Technology Research and Development, 58*(1), 81-97. doi: 10.1007/s11423-008-9087-4

Ifenthaler, D. (2010b). Scope of graphical indices in educational diagnostics. In D. Ifenthaler, P. Pirnay-Dummer & N. M. Seel (Eds.), *Computer-based diagnostics and systematic analysis of knowledge* (pp. 213-234). New York: Springer.

Ifenthaler, D., Masduki, I., & Seel, N. M. (2011). The mystery of cognitive structure and how we can detect it. Tracking the development of cognitive structures over time. *Instructional Science, 39*(1), 41-61. doi: 10.1007/s11251-009-9097-6

Ifenthaler, D., Pirnay-Dummer, P., & Seel, N. M. (2007). The role of cognitive learning strategies and intellectual abilities in mental model building processes. *Technology, Instruction, Cognition and Learning, 5*(4), 353-366.

Ifenthaler, D., & Seel, N. M. (2005). The measurement of change: Learning-dependent progression of mental models. *Technology, Instruction, Cognition and Learning, 2*(4), 317-336.

Jenert, T. (2008). Ganzheitliche Reflexion auf dem Weg zu Selbstorganisiertem Lernen. *Bildungsforschung, 2*(5), 1-18.

Johnson-Laird, P. N. (1989). Mental models. In M. I. Posner (Ed.), *Foundations of cognitive science* (pp. 469-499). Cambridge, MA: MIT Press.

Jonassen, D. H. (2000). Toward a design theory of problem solving. *Educational Technology Research & Development, 48*(4), 63-85. doi: 10.1007/BF02300500

Jonassen, D. H., & Cho, Y. H. (2008). Externalizing mental models with mindtools. In D. Ifenthaler, P. Pirnay-Dummer & J. M. Spector (Eds.), *Understanding models for learning and instruction. Essays in honor of Norbert M. Seel* (pp. 145-160). New York: Springer.

Koedinger, K. R., & Aleven, V. (2007). Exploring the assistance dilemma in experiments with cognitive tutors. *Educational Psychology Review, 19*(3), 239-264.

Lin, X., & Lehmann, J. D. (1999). Supporting learning of variable control in a computer-based biology environment: Effects of prompting college students to reflect on their own thinking. *Journal of Research in Science Teaching, 36*, 837-858.

Mayer, R. E. (1998). Cognitive, metacognitive, and motivational aspects of learning. *Instructional Science, 26*(1-2), 49-63. doi: 10.1023/A:1003088013286

Mikkilä-Erdmann, M., Penttinen, M., Anto, E., & Olkinuora, E. (2008). Constructing mental models during learning from science text. Eye tracking methodology meets conceptual change. In D. Ifenthaler, P. Pirnay-Dummer & J. M. Spector (Eds.), *Understanding models for learning and instruction. Essays in honor of Norbert M. Seel* (pp. 63-79). New York: Springer.

O'Donnell, A. M., Dansereau, D. F., & Hall, R. H. (2002). Knowledge maps as scaffolds for cognitive processing. *Educational Psychology Review, 14*, 71-86.

Piaget, J. (1943). *Le developpement mental de l'enfant*. Zürich: Rascher.

Piaget, J. (1976). *Die Äquilibration der kognitiven Strukturen*. Stuttgart: Klett.

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. R. Pintrich & M. Zeidner (Eds.), *Handbook of self-regulated learning* (pp. 451-502). San Diego, CA: Academic Press.

Pirnay-Dummer, P., & Ifenthaler, D. (2010). Automated knowledge visualization and assessment. In D. Ifenthaler, P. Pirnay-Dummer & N. M. Seel (Eds.), *Computer-based diagnostics and systematic analysis of knowledge* (pp. 77-115). New York: Springer.

Pirnay-Dummer, P., Ifenthaler, D., & Spector, J. M. (2010). Highly integrated model assessment technology and tools. *Educational Technology Research and Development, 58*(1), 3-18. doi: 10.1007/s11423-009-9119-8

Ruiz-Primo, M. A., Schultz, S. E., Li, M., & Shavelson, R. J. (2001). Comparison of the reliability and validity of scores from two concept-mapping techniques. *Journal of Research in Science Teaching, 38*(2), 260-278.

Rumelhart, D. E., Smolensky, P., McClelland, J. L., & Hinton, G. E. (1986). Schemata and sequential thought processes in PDP models. In J. L. McClelland & D. E. Rumelhart (Eds.), *Parallel distributed processing. Explorations in the microstructure of cognition. Volume 2: Psychological and biological models* (pp. 7-57). Cambridge, MA: MIT Press.

Schmidt-Weigand, F., Hänze, M., & Wodzinski, R. (2009). Complex problem solving and worked examples. *Zeitschrift für Pädagogische Psychologie, 23*(2), 129-138.

Schraw, G. (2007). The use of computer-based environments for understanding and improving self-regulation. *Metacognition and Learning, 2*(2-3), 169-176.

Schraw, G., & Dennison, R. S. (1994). Asessing metacoginitive awareness. *Contemporary Educational Psychology, 19*(4), 460-475.

Schuler, H., & Hell, B. (2005). *Analyse des Schlussfolgernden und Kreativen Denkens (ASK)*. Bern: Verlag Hans Huber.

Seel, N. M. (1999a). Educational diagnosis of mental models: Assessment problems and technology-based solutions. *Journal of Structural Learning and Intelligent Systems, 14*(2), 153-185.

Seel, N. M. (1999b). Educational semiotics: School learning reconsidered. *Journal of Structural Learning and Intelligent Systems, 14*(1), 11-28.

Seel, N. M., Ifenthaler, D., & Pirnay-Dummer, P. (2009). Mental models and problem solving: Technological solutions for measurement and assessment of the development of expertise. In P. Blumschein, W. Hung, D. H. Jonassen & J. Strobel (Eds.), *Model-based approaches to learning: Using systems models and simulations to improve understanding and problem solving in complex domains* (pp. 17-40). Rotterdam: Sense Publishers.

Simons, P. R. J. (1992). Lernen selbstständig zu lernen - ein Rahmenmodell. In H. Mandl & H. F. Friedrich (Eds.), *Lern- und Denkstrategien. Analyse und Intervention* (pp. 251-264). Göttigen: Hogrefe.

Simons, P. R. J. (1993). Constructive learning: The role of the learner. In T. M. Duffy, J. Lowyck & D. H. Jonassen (Eds.), *Designing environments for constructive learning* (pp. 291-313). New York: Springer.

Thillmann, H., Künsting, J., Wirth, J., & Leutner, D. (2009). Is it merely a question of "What" o prompt or also "When" to prompt? *Zeitschrift für Pädagogische Psychologie, 23*(2), 105-115.

Tversky, A. (1977). Features of similarity. *Psychological Review, 84*, 327-352.

Veenman, M. V. J. (2007). The assessment and instruction of self-regulation in computer-based environments: a discussion. *Metacognition and Learning, 2*(2-3), 177-183.

Veenman, M. V. J., Kerseboom, L., & Imthorn, C. (2000). Test anxiety and metacognitive skillfulness: Availability versus production deficiencies *Anxiety, Stress and Coping, 13*(4), 391-412.

Veenman, M. V. J., van Hout-Wolters, B. H. A. M., & Afflerbach, P. (2006). Metacognition and learning: conceptual and methodological considerations. *Metacognition and Learning, 1*(1), 3-14.

Veenman, M. V. J., Wilhelm, P., & Beishuizen, J. J. (2004). The relation between intellectual and metacognitive skills from a developmental perspective. *Learning and Instruction, 14*(1), 89-109.

Winne, P. H. (2001). Self-regulated learning viewed from models of information processing. In B. J. Zimmerman & D. Schunk (Eds.), *Self-regulated learning and academic achievement. Theoretical perspectives* (pp. 153-190). Mahawah, NJ: Lawrence Erlbaum Associates.

Wirth, J. (2009). Prompting self-regulated learning through prompts. *Zeitschrift für Pädagogische Psychologie, 23*(2), 91-94.

Wirth, J., & Leutner, D. (2008). Self-regulated learning as a competence. Implications of theoretical models for assessment methods. *Zeitschrift für Psychologie, 216*, 102-110.

Zimmerman, B. J. (1989). Models of self-regulated learning and academic achievement. In B. J. Zimmerman & D. H. Schunk (Eds.), *Self-regulated learning and academic achievement. Theory, research and practice* (pp. 1-25). New York: Springer.

Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal, 45*(1), 166-183.

Zimmerman, B. J., & Schunk, D. (2001). Theories of self-regulated learning and academic achievement: An overview and analysis. In B. J. Zimmerman & D. Schunk (Eds.), *Self-regulated learning and academic achievement. Theoretical perspectives* (pp. 1-37). Mahawah, NJ: Lawrence Erlbaum Associates.