TECHNISCHE UNIVERSITÄT DARMSTADT

GRADUIERTENKOLLEG E-LEARNING

Dissertation zur Erlangung des akademischen Grades
Doctor rerum naturalium (Dr. rer. nat.)
am Fachbereich Humanwissenschaften
der Technischen Universität Darmstadt (D17)
Institut für Psychologie
Graduiertenkolleg Qualitätsverbesserung im E-Learning durch rückgekoppelte Prozesse

Improving the Quality of E-Learning by Enhancing Self-Regulated Learning.

A Synthesis of Research on Self-Regulated Learning and an Implementation of a Scaffolding Concept.

vorgelegt von
Dipl.-Psych. Bastian Frithjof Benz,
geboren in Frankfurt am Main

Darmstadt, 2010

1. Referent: Prof. Dr. Bernhard Schmitz
2. Referent: Prof. Dr. Bruno Rüttinger
3. Referent: Prof. Dr. Ralf Steinmetz

Einreichungsdatum: 25.03.2010
Disputationsdatum: 08.06.2010

Acknowledgments

Zunächst möchte ich meine Dankbarkeit gegenüber Bernhard Schmitz ausdrücken, der mir die Gelegenheit zu dieser Promotion gegeben hat. Herr Schmitz, ich bin Ihnen sehr dankbar für die erfahrene Unterstützung und Förderung!

Mein Dank gilt auch den aktuellen und ehemaligen Mitarbeitern der AG Schmitz, zu denen ich trotz der inhaltlichen Distanz immer ein gutes Verhältnis hatte!

Ebenfalls bedanken möchte ich mich bei meiner studentischen Hilfskraft Thomas Lösch, der mir aufgrund seiner akribischen und engagierten Arbeitsweise eine große Hilfe war!

Ein Riesendank gilt allen Mitgliedern des Graduiertenkollegs *Qualitätsverbesserung im E-Learning durch rückgekoppelte Prozesse*. Die vielen interdisziplinären Gespräche waren nicht nur sehr hilfreich für diese Arbeit, sondern haben insbesondere zu meiner persönlichen Weiterentwicklung und Horizonterweiterung beigetragen. Ich freue mich, dass ich die Gelegenheit hatte, Teil dieses GKs zu sein!

Insbesondere möchte ich meinen Projektpartnern Doreen Böhnstedt und Philipp Scholl danken, ohne die diese Promotion nicht in dieser Form hätte verwirklicht werden können. Ich danke Euch für die gute Zusammenarbeit und das freundschaftliche Verhältnis!

Des Weiteren möchte ich meinen GK-Mitstreitern Nina Hollender, Cristian Hofmann, Antje Müller und Martin Leidl für das freundschaftliche und unterstützende Verhältnis danken!

Mein Dank gilt auch Christoph Rensing, meinem Zweitbetreuer im Graduiertenkolleg!

Bruno Rüttinger und Ralf Steinmetz danke ich, dass sie sich für die Begutachtung dieser Arbeit zur Verfügung gestellt haben!

I would also like to thank Phil Winne for giving me the opportunity to visit his lab and for taking the time to make very helpful comments and suggestions on my work! Furthermore, I would like to thank Mingming Zhou and Rylan Egan for enjoyable and enlightening discussions!

Auch möchte ich mich bei Maria Bannert, Susanne Narciss und Dietrich Albert für die netten und aufschlussreichen Gespräche bedanken! I would also like to thank Marcel Veenman and Roger Azevedo for being so open-minded and supportive!

I further thank Jane Thompson for doing a wonderful job proofreading this dissertation!

Und zum Schluss möchte ich noch all jenen danken, die mir die Kraft gegeben haben, diese Arbeit zu vollenden und alle Hürden zu überspringen. Ein unendlich großer Dank gilt meiner Freundin Angelika, die viele Entbehrungen ertragen musste und immer für mich da war! Außerdem danke ich von ganzem Herzen meinen Eltern Gerlinde und Volker, meinen Brüdern und ihren Partnerinnen Tobias und Maggi sowie Manuel und Larissa und meiner Oma Hildegard! Nicht zuletzt gilt ein großer Dank Familie Ackermann!

Summary

Until this day, e-learning scenarios have not been able to meet the great expectations that have been put on them. From a psychological point of view, one of the reasons for previously disappointing results is the high degree of freedom that e-learning scenarios confront individuals with. The concept of self-regulated learning (SRL) provides a framework for improving the quality of learning in environments of high self-determination. Following the assumption that the deployment of SRL processes during learning is an indicator for a learning process of high quality, which in turn entails a learning outcome of high quality, scientists have created a variety of SRL interventions. In line with this research, the main goal of this dissertation was to improve the quality of e-learning, applying learning on the World Wide Web (WWW) as a specific scenario, by enhancing individuals' deployment of SRL processes.

As a first study, synthesizing past SRL research, a meta-analysis was conducted to evaluate the relevance of SRL for the quality of learning, and to quantify the impact of SRL interventions on academic achievement. Another aim was to provide guidance for future SRL research by identifying features of SRL interventions that have proven to be effective as well as features that have lacked effectiveness and escaped attention in the past. Putting special effort into identifying unpublished dissertations in order to avoid a publication bias, a pool of 39 studies that reported 44 independent treatments was established. Applying a random effects model, a weighted mean effect of SRL interventions on academic performance of $\Delta_{Glass} = .45$ was determined. One-way moderator analyses identified peer-review status, research design, instance of delivery, and age of participants as influential variables. In further analyses, when combining heterogeneous categories of moderators to establish homogeneous subgroups, it was found that treatments focusing on the metacognitive and cognitive layers, strategy instructions, and interventions within a mathematical domain were most effective for young learners between the ages of 9 and 14. With regard to computer-based interventions, nonsignificant effects on performance were revealed.

Aiming to improve the quality of e-learning, the author simultaneously developed a concept for providing SRL support. Following an indirect approach of assistance, the core of the concept was to optimize the learning environment by providing learners with scaffolds that served as tools that could be used to complete a learning task and that simultaneously induced the deployment of six metacognitive processes in the three cyclical phases of learning. Aiming to provide learners with more intensive guidance, an extended concept also included administering additional prompts. The sketched scaffolding concept was realized by implementing the Firefox extension *E-Learning knoWledge Management System* (ELWMS).

The second study of this dissertation served to evaluate whether working with ELWMS enhanced the quality of the learning process and the learning outcome. Randomly assigned to four conditions, N = 64 participants learned on Wikipedia about *Classical Antiquity* for 45 min. The two experimental groups were equipped with the standard version of ELWMS, and were either free to apply the scaffolds of their own accord or received two additional prompts, whereas the two control groups worked with the standard version of Firefox. The quality of the learning process was evaluated by a self-developed context-specific SRL self-report questionnaire. In addition, participants' overt actions were assessed by generating log data and by conducting qualitative analyses of screen recordings. The quality of the learning outcome was evaluated by a self-developed achievement test, applied as a pretest and a posttest, and by determining the quality of the created structure. Results revealed positive effects of the scaffolding concept on the quality of the learning process, and ambiguous effects of additional prompting.

The third study of this dissertation served to provide further insight into the effectiveness of the sketched scaffolding concept with an elaborated study design. Equipped with the second generation of ELWMS, learners of experimental groups were either free to apply the scaffolds of their own accord or received additional intensive prompting of six processes that were considered to enhance achievement. To be able to generate comparable log data for all conditions, the control group worked with a downgraded version of ELWMS that did not provide metacognitive support. Before learning on Wikipedia about *Classical Antiquity* for 45 min, all N = 108 participants had to complete a web-based training on ELWMS. The quality of the learning process was evaluated by applying a revised context-specific SRL self-report questionnaire that was synchronized with an advanced and optimized method of collecting log data. In further qualitative analyses, logged overt actions were assigned a rating of relevance. The quality of the learning outcome was evaluated by a revised achievement test, applied as a pretest and a posttest, and by determining the quality of the created structure. In line with the second study, results revealed positive effects of the scaffolding concept on the quality of the learning process. This effect was enhanced by additional intensive prompting.

In whole, this dissertation presents evidence that enhancing SRL improves the quality of learning and provides a path for creating effective SRL interventions. By applying a complex multimethod approach, it further proposes a promising concept for inducing SRL processes in e-learning.

Zusammenfassung

E-Learning Szenarien sind in der Vergangenheit oft hinter den großen Erwartungen, die in sie gesetzte wurden, zurück geblieben. Aus psychologischer Sicht lassen sich diese enttäuschenden Ergebnisse durch die große Freiheit, der Individuen beim E-Learning begegnen, erklären. Das Konzept des selbstregulierten Lernens (SRL) bietet einen Ansatz zur Verbesserung der Lernqualität in Szenarien, die den Lernenden ein hohes Maß an Eigenverantwortung übertragen. Es wird postuliert, dass die Ausführung von SRL Prozessen während des Lernens ein Indikator für einen qualitativ hochwertigen Lernprozess darstellt, der wiederum ein qualitativ hochwertiges Lernergebnis nach sich zieht. Dieser Annahme folgend haben Wissenschaftler eine große Bandbreite an SRL Interventionen entwickelt. Anknüpfend an diese Forschung war es das Hauptziel dieser Dissertation, eine Qualitätsverbesserung im E-Learning durch die Förderung von Selbstregulationsprozessen zu erreichen. Das Lernen im World Wide Web (WWW) diente dabei als Anwendungsszenario.

In der ersten Studie dieser Dissertation erfolgte eine Synthese der bestehenden SRL Forschung. Ziel war zum einen die Evaluation der Relevanz des SRL Ansatzes für die Lernqualität. Dabei sollte der Effekt von SRL Interventionen auf akademische Leistung quantifiziert werden. Zum anderen stand die Erarbeitung einer Leitlinie für die zukünftige SRL Forschung im Mittelpunkt. Zu diesem Zweck wurden Studieneigenschaften identifiziert, die sich in der Vergangenheit als effektiv bzw. ineffektiv erwiesen haben oder keine Beachtung fanden. Bei der Generierung des Studienpools, der aus 39 Studien und 44 unabhängigen Treatments bestand, wurde besonderes Augenmerk auf die Identifikation von unpublizierten Dissertationen zur Vermeidung einer Publikationsverzerrung (publication bias) gelegt. Die Integration der Studieneffekte erfolgte unter Verwendung eines Modells zufallsvariabler Effekte. Es konnte eine gewichtete mittlere Effektstärke von SRL Interventionen auf akademische Leistung von $\Delta_{Glass} = .45$ ermittelt werden. Mithilfe einfaktorieller Moderatorenanalysen wurden Peer-Review Status, Studiendesign, Vermittlungsinstanz und Alter der Teilnehmer/innen als einflussreiche Variablen identifiziert. Mit dem Ziel der Bildung homogener Subgruppen wurden des Weiteren heterogene Moderatorkategorien mit weiteren Moderatoren kombiniert. Dabei konnten Treatments, die sowohl eine metakognitive als auch eine kognitive Förderung realisieren, Strategievermittlungen und Interventionen im mathematischen Kontext als besonders effektiv identifiziert werden, wenn sie an jungen Lernern zwischen 9 und 14 Jahren durchgeführt wurden. Für computerbasierte Interventionen wurden keine signifikanten Effekte gefunden.

Zur Verbesserung der Qualität im E-Learning entwickelte der Autor parallel ein Konzept, das entsprechend eines indirekten Unterstützungsansatzes auf die Optimierung der Lernumgebung abzielte. Kern des Konzepts war die Bereitstellung von Lernerunterstützungen (scaffolds), die sowohl als Werkzeug zur Bearbeitung einer Lernaufgabe verwendet werden können als auch die Ausführung von 6 metakognitiven Prozessen in den drei zyklischen Phasen des Lernens induzieren. Mit dem Ziel der Applikation einer intensiveren Unterstützung sah ein erweitertes Konzept außerdem die Darbietung zusätzlicher Prompts vor. Das beschriebene Unterstützungskonzept wurde durch die Entwicklung der Firefoxerweiterung *E-Learning knoWledge Management System* (ELWMS) realisiert.

Die zweite Studie dieser Dissertation diente zur Untersuchung des Einflusses der entwickelten ELWMS Software auf die Qualität des Lernprozesses und des Lernergebnisses. N = 64 Probanden lernten 45 Minuten auf Wikipedia zum Thema *Antike* und waren dabei vier Versuchsgruppen zufällig zugeordnet. Die Experimentalgruppen waren mit ELWMS ausgestattet, wobei sie entweder frei über die Verwendung der Scaffolds verfügen konnten oder durch zwei Prompts zusätzliche Unterstützung erfuhren. Die Kontrollgruppen arbeiteten mit der Standardversion des Firefox. Die Qualität des Lernprozesses wurde mithilfe eines selbstentwickelten kontextspezifischen SRL Fragebogens erhoben. Zusätzlich wurden die beobachtbaren Aktionen der Versuchspersonen durch die Generierung von Logdaten und durch qualitative Analysen der Monitoraufzeichnungen untersucht. Die Qualität des Lernergebnisses wurde durch einen selbstentwickelten Leistungstest, der sowohl vor als auch nach der Lernphase appliziert wurde, und durch die Ermittlung der Qualität der erstellten Struktur erhoben. Die Ergebnisse zeigten positive Effekte des vorgeschlagenen Unterstützungskonzepts auf die Qualität des Lernprozesses. Effekte des zusätzlichen Promptings waren uneindeutig.

Ziel der dritten Studie dieser Dissertation war es, mithilfe eines überarbeiteten Studiendesigns einen tieferen Einblick in die Effektivität des beschriebenen Unterstützungskonzepts zu erlangen. Die Experimentalgruppen waren mit der zweiten Generation der ELWMS Software ausgestattet, wobei sie entweder frei über die Verwendung der Scaffolds verfügen konnten oder zusätzlich ein intensives Prompting zur Ausführung sechs leistungsfördernder Prozesse erhielten. Mit dem Ziel der Generierung vergleichbarer Logdaten in allen Versuchsbedingungen wurde die Kontrollgruppe mit einer Basisvariante der ELWMS Software, die keine metakognitive Unterstützung bereitstellte, ausgestattet. Bevor die N = 108 Probanden 45 Minuten auf Wikipedia zum Thema *Antike* lernten, absolvierten sie ein web-basiertes Training zur Verwendung der ELWMS Software. Die Qualität des Lernprozesses

wurde durch einen überarbeiteten kontextspezifischen SRL Fragebogen erhoben, der mit der weiterentwickelten Methodik zur Erhebung von Logdaten synchronisiert war. In weiteren Analysen wurden die geloggten Aktionen einer qualitativen Analyse unterzogen. Die Qualität des Lernergebnisses wurde durch einen überarbeiteten Leistungstest sowie durch die Ermittlung der Qualität der erstellten Struktur erhoben. Wie in der zweiten Studie wurden positive Effekte des vorgeschlagenen Unterstützungskonzepts auf die Qualität des Lernprozesses gefunden. Dieser Effekt wurde durch intensives zusätzliches Prompting verstärkt.

Diese Dissertation belegt, dass die Förderung von SRL die Qualität des Lernens erhöht und stellt eine Leitlinie zur Entwicklung effektiver SRL Interventionen bereit. Unter Verwendung eines aufwendigen Multimethodenansatzes schlägt sie außerdem ein vielversprechendes Konzept zur Induktion von SRL Prozessen im E-Learning vor.

Contents

List of Tables

Part 1: Synopsis

Part 2: Original Articles

Study 1: Self-Regulated Learning and Academic Success. Do Self-Regulated Learning Interventions Enhance Performance? A Meta-Analysis.

Study 2: Improving the Quality of E-Learning. Scaffolding Self-Regulated Learning on the World Wide Web.

Study 3: Improving the Quality of Learning on the World Wide Web by Scaffolding Self-Regulated Learning.

List of Figures

Part 1: Synopsis

Part 2: Original Articles

Study 1: Self-Regulated Learning and Academic Success. Do Self-Regulated Learning Interventions Enhance Performance? A Meta-Analysis.

Study 2: Improving the Quality of E-Learning. Scaffolding Self-Regulated Learning on the World Wide Web.

Study 3: Improving the Quality of Learning on the World Wide Web by Scaffolding Self-Regulated Learning.

List of Original Publications

This dissertation is based on three original articles, referred to in the text by their numerals:

Study 1: Benz, B. F., & Schmitz, B. (2009). *Self-regulated learning and academic success: Do self-regulated learning interventions enhance performance? A meta-analysis.* Manuscript submitted for publication.

Study 2: Benz, B. F., Scholl, P., Boehnstedt, D., Schmitz, B., Rensing, C., & Steinmetz, R. (2010). *Improving the Quality of E-Learning. Scaffolding Self-Regulated Learning on the World Wide Web.* Manuscript submitted for publication.

Study 3: Benz, B. F., Scholl, P., Boehnstedt, D., Schmitz, B., Rensing, C., & Steinmetz, R. (2010). *Improving the Quality of Learning on the World Wide Web by Scaffolding Self-Regulated Learning.* Manuscript submitted for publication.

Improving the Quality of E-Learning by Enhancing Self-Regulated Learning

A Synthesis of Research on Self-Regulated Learning and an Implementation of a Scaffolding Concept.

Introduction

The rapid technological development of the past decades has led educational scientists to enrich learning environments with technology. However, research has shown that such e-learning scenarios often result in an unsatisfying quality of learning (Dillon & Gabbard, 1998; Shapiro & Niederhauser, 2004). From a psychological point of view, one of the reasons for these findings lies in the high degree of freedom that individuals are confronted with during e-learning.

The concept of self-regulated learning (SRL) proposes that individuals' deployment of SRL processes during learning is an indicator for a learning process of high quality, which in turn entails a learning outcome of high quality (Zimmerman & Schunk, 2001). Following this assumption, scientists have created a variety of SRL interventions to enhance individuals' deployment of SRL processes during learning, and as a consequence, to enhance the learners' achievement in situations of high self-determination.

The main goal of this dissertation was to improve the quality of e-learning, applying learning on the World Wide Web (WWW) as a specific scenario, by enhancing individuals' deployment of SRL processes. On the one hand, past research on SRL was synthesized in order to evaluate the relevance of SRL for the quality of learning and to quantify the impact of SRL interventions on academic achievement. Further, to guide future SRL research, features of SRL interventions that have proven to be effective, ineffective, or that have escaped attention in past research were identified. On the other hand, to improve the quality of the learning process and the learning outcome in e-learning, the concept of optimizing the learning environment by providing scaffolds that offer functions for completing a learning task and that simultaneously induce the deployment of six metacognitive processes in three cyclical phases of learning was evaluated. A further goal was to investigate whether providing prompts in addition to the sketched concept of support would further enhance the quality of e-learning.

This PhD thesis consists of two parts. In the first part, after having derived the research questions (section 1), the three studies that were conducted to answer the questions are briefly sketched (section 2, 3, and 4). The results of the dissertation are then discussed (section 5), and future perspectives are pointed out (section 6). The second part of this PhD thesis presents three original publications (Studies 1-3) that have been submitted for publication in scientific peer-reviewed journals.

Part 1: Synopsis

1. Derivation of Research Questions

Accompanied by the rapid technological development of the past 2 decades, there has been a movement to enrich learning environments with technology. However, research shows that such e-learning scenarios have not met the great expectations that have been put on them with regard to quality of learning (Dillon & Gabbard, 1998; Shapiro & Niederhauser, 2004). From a psychological point of view, one of the reasons for these disappointing results is that individuals experience less guidance and higher degrees of freedom in technology-enriched learning environments. As a consequence, in e-learning settings, like they have been commonly implemented, the skills and strategies that a learner is in possession of are major predictors of the quality of learning. If an individual is not able to handle the responsibility that he or she is equipped with, there is a good chance that e-learning will result in poor quality.

The concept of SRL provides a framework for enhancing the quality of learning in environments that provide individuals with high degrees of freedom. It is assumed that the deployment of SRL processes during learning is an indicator for a learning process of high quality, which in turn entails a learning outcome of high quality. From a social cognitive theoretical perspective, SRL is defined as learners' self-generated thoughts, feelings, and actions that are systematically oriented toward the attainment of their learning goals (Zimmerman & Schunk, 2001). More specifically, it has been suggested that individuals have to regulate motivational/emotional, cognitive, and metacognitive processes (Boekaerts, 1999) in the preaction, action, and postaction phases of learning (Schmitz & Wiese, 2006; Zimmerman, 2000) to be able to achieve their goals (see also Pintrich, 2000; Winne & Hadwin, 2008).

Following the assumption that the deployment of SRL processes during learning helps individuals to master situations of high self-determination, scientists have created SRL interventions to enhance learners' academic achievement. However, at the moment, there does not exist a framework that provides researchers with guidance when deciding *what* SRL processes to foster, and *how*, *when*, and *to whom* to administer support, as well as how to evaluate the effectiveness of SRL interventions. As a result, a great variety of SRL treatments can be perceived. Focusing on various age groups, researchers have aimed to enhance single or multiple processes on the motivational/emotional, the cognitive, or the metacognitive layers of SRL (Boekaerts, 1999). They have fostered micro-level learning in order to enhance SRL during the implementation of an elementary task, and they have fostered mid-level learning in order to help learners to manage their daily study routines (Alexander, 1997). Some researchers have followed direct training approaches, carrying out strategy instructions to equip learners who suffer from a mediation deficiency (Reese, 1962) with SRL strategies. Other researchers have pursued indirect approaches, providing process support to induce SRL processes during the implementation of a task, focusing on learners who suffer from a production deficiency (Flavell, 1970), and hence are already in possession of SRL strategies, but do not manage to apply them (Friedrich & Mandl, 1992). Further, SRL interventions have been delivered to participants by humans, by computers, or by paper. They have been carried out over different periods of time and in different domains of learning. And finally, researchers have evaluated the effects of SRL interventions on performance measures of different complexity, in laboratories or real classrooms, and by applying experimental or quasi-experimental designs. Considering this great variety of methods, it is not a surprise that contradictory results regarding the effectiveness of SRL treatments have been found (e.g., Campillo, 2006; Kramarski & Mizrachi, 2006; Mosley, 2006).

Aiming to improve the quality of e-learning by enhancing learners' involvement in SRL, the author designed an SRL intervention himself. As a specific e-learning scenario, web-based learning was applied for two reasons. On the one hand, the WWW, which nowadays is used as a resource for learning in various settings (United Nations [UN], 2008), is of great relevance for modern life. On the other hand, the WWW is a nonlinear and unstructured environment (Jonassen, 1996) that provides learners' with an enormous degree of freedom, and thereby requires learners to be highly self-regulated in order to learn successfully.

To enhance learners' involvement in SRL during web-based learning, and thereby to improve the quality of their learning process and their learning outcome, the following approach was pursued. Focusing on the metacognitive layer (Boekaerts, 1999) and taking a process view of SRL (Pintrich, 2000; Schmitz & Wiese, 2006; Zimmerman, 2000), six metacognitive processes were derived. More specifically, it was assumed that the employment of goal setting and planning in the preaction phase, self-monitoring and process-regulation in the action phase, and reflection and modification in the postaction phase would improve the quality of web-based learning. Focusing on learners who are in possession of SRL strategies, but who do not manage to apply the strategies because they are suffering from a production deficiency (Flavell, 1970), an indirect approach of assistance was followed (Friedrich & Mandl, 1992). It was aimed at optimizing the browser, which constitutes the window through which the WWW is seen, and thereby the WWW as a learning environment itself by inducing the deployment of the six metacognitive processes during web-based learning. In contrast to previous research on hypermedia learning, which has mainly focused on adding metacognitive support to the learning environment (e.g., Land & Zembal-Saul, 2003), an integrated approach was followed. Referring to the concept of *scaffolding* (Vygotsky, 1978) and applying the computer as the instance of delivery, the goal was to provide learners with tools that would serve to complete learning tasks on the WWW and that would simultaneously induce the metacognitive processes.

In SRL research, goal setting and planning, self-monitoring and process-regulation, and reflection and modification have been considered most beneficial when carried out during a specific phase of learning (Pintrich, 2000; Schmitz & Wiese, 2006; Zimmerman, 2000). However, it has been suggested that learners may not be able to decide if, how, and when to apply scaffolds of their own accord (Aleven, Stahl, Schworm, Fischer, & Wallace, 2003; Oliver & Hannafin, 2000). As those results have been based on scaffolds that were added to the functions of a learning environment, it was unclear whether the findings would also hold for the integrated approach that was pursued in this dissertation. With reference to studies that had reported beneficial effects of prompting (e.g., Bannert, 2006; Horz, Winter, & Fries, 2009; Kramarski, & Zeichner, 2001; Schwonke, Hauser, Nuckles, & Renkl, 2006), to further enhance the quality of the learning process and the learning outcome, an extended approach additionally included the administration of prompts to provide learners with more intensive guidance.

In sum, following the assumption that the deployment of SRL processes enhances the quality of learning in environments of high self-determination, scientists have created SRL interventions to improve learners' academic achievement. However, at the moment there does not exist a framework to provide researchers with guidance when designing these treatments. As a consequence, a huge variety of SRL interventions with contradictory effects on performance can be perceived. With the goal of enhancing the quality of the learning process and the learning outcome in e-learning, applying web-based learning as a specific scenario, the author conducted an SRL intervention himself. Following an indirect approach of assistance, the WWW was optimized as a learning environment by integrating scaffolds that combined functionality and metacognitive support in the web browser. Further, to provide learners with more intensive guidance, the approach of administering additional prompts was pursued.

Based on these issues, the present PhD thesis focused on the following research questions:

1. Evaluating the relevance of SRL for learning by synthesizing past SRL research, have SRL interventions positively affected learners' academic achievement, and how can this effect be quantified?
2. To guide future SRL research, which features of SRL interventions have proven to be effective, which have proven to be ineffective, and which have escaped attention in past research?
3. In e-learning, using web-based learning as a specific scenario, does an SRL intervention that is specified by the following three characteristics improve the quality of the learning process and the learning outcome?
   - Following an indirect approach of assistance, the learning environment is optimized to induce the deployment of SRL processes.
   - Following an integrated approach of scaffolding, learners are provided with tools that offer functions to complete a learning task and that simultaneously induce SRL processes.
   - Following a holistic concept of SRL support, the six metacognitive processes of goal setting, planning, self-monitoring, process-support, reflection, and modification are induced in the three cyclical phases of learning.
4. In e-learning, using web-based learning as a specific scenario, does prompting that is administered in addition to the sketched scaffolding approach improve the quality of the learning process and the quality of the learning outcome?

To answer these research questions, the author conducted three studies, which are briefly sketched in the following section.

2. Study 1: Self-Regulated Learning and Academic Success. Do Self-Regulated Learning Interventions Enhance Performance? A Meta-Analysis.

To answer the first and the second research questions of this PhD thesis, the author conducted a meta-analysis synthesizing past SRL research. On the one hand, evaluating the relevance of SRL for learning, the author aimed to investigate whether SRL interventions have positively affected learners' academic achievement, and how the effect could be quantified. On the other hand, to guide future research, an additional goal of the study was to identify features of SRL interventions that have proven to be effective, as well as to identify properties of interventions that have escaped attention in past research.

2.1 Method

To establish a pool of SRL intervention studies, with respect to the independent variable, a broad approach was followed by including a wide variety of different kinds of SRL treatments. Generally, each study that was published from 1990 to March 2007, in English or German, and that had used an intervention referring to a common SRL model (e.g., Boekaerts, 1999; Pintrich, 2000; Schmitz & Wiese, 2006; Winne & Hadwin, 2008; Zimmerman, 2000) in a control-group design, with a minimum of 10 participants per condition, was potentially suitable for integration. However, to be integrated, accounting for the central relevance of metacognition for SRL, a treatment had to contain a metacognitive component. In addition, studies utilizing participants with learning disabilities, learning difficulties, or special needs were excluded. In contrast to the broad approach that was followed with reference to the independent variable, regarding the dependent variable, studies had to assess a measure of academic performance.

With the goal of identifying SRL interventions that met integration criteria, the databases PsycInfo, ERIC, and Psyndex were searched with the keywords self-reg*, selfreg*, and selbstreg* (White, 1994). Special effort was placed on the acquisition of unpublished dissertations in order to avoid a publication bias (Rothstein, Sutton, & Borenstein, 2005). Out of 2,407 abstracts that were screened, 154 papers were viewed in full text, and 38 met integration criteria. One dissertation reported two independent studies, resulting in 39 integrated studies. Relevant data were extracted from each study and coded by two independent researchers with a mean interrater reliability of .97 (Orwin, 1994). Study effect sizes were calculated contrasting post measures of experimental and control groups in relation to the standard deviation of the control group: $\Delta_{Glass} = (\bar{x}_1 - \bar{x}_2) / s_{KG}$ (Glass, McGaw, & Smith, 1981). The variance of the study effect sizes was determined using a formula provided by Hedges and Olkin (1985).

To avoid dependencies between effect sizes, on the side of the independent variable, only as many experimental groups were selected from a study as there were control groups (Hedges & Olkin, 1985), beginning with the treatment that had been rated to consist of the largest number of SRL components (Boekaerts, 1999). Out of the 39 studies that met the integration criteria, three studies were identified that used more than one independent treatment, resulting in a total of 44 treatments. On the side of the dependent variable, following Hedges and Olkin (1985), only the most complex measure of academic achievement was taken into account, which resulted in 44 independent effect sizes. To calculate the weighted mean effect size, single effect sizes were integrated by applying a random effects model in which the component of variance (τ^2) was calculated using a formula provided by Hedges and Vevea (1998). A homogeneity test was conducted by computing the weighted squared discrepancy of the study effect sizes from the weighted mean effect size of all studies (Hedges, 1982) in order to investigate the homogeneity of the sample of interventions, and thereby, the generalizability of the weighted mean effect size. To analyze proposed moderators, categorical one-way moderator analyses were conducted. Further, following an exploratory approach, an attempt was made to resolve heterogeneity within categories of moderators by crossing them with other moderators.

2.2 Results

Summarizing 17 years of SRL research based on 4,047 learners, it could be stated that scientists have managed to create SRL treatments that significantly affect performance by $\Delta_{Glass} = .45$, $p < .01$. In other words, receiving an SRL treatment, regardless of the specific constitution of the treatment, on average enhances the quality of the learning outcome of participants in experimental groups by almost half of the control group's standard deviation. Accordingly, this meta-analysis underlines the relevance of SRL for learning. However, the test for homogeneity, which turned out to be just significant, $\chi^2(43) = 59.3$, $p = .05$, implied heterogeneity within the sample of independent treatments. Accordingly, the weighted mean effect size should not be taken as an estimate of the population parameter, but should serve rather as a descriptive result (Shadish & Haddock, 1994).

In order to analyze the heterogeneity of the sample for systematic patterns, and thereby to establish homogeneous groups of interventions (Hedges & Olkin, 1985), categorical moderator analyses were conducted using a random effects model. To examine the model fit, as well as to test for meaningful differences between categories, three tests were conducted to calculate homogeneity (a) within each category, (b) over all categories, and (c) between categories (Hedges & Olkin, 1985).

Hypothesis-driven one-way moderator analyses revealed a satisfying model fit for the variables *review status*, *research design*, *instance of delivery of intervention*, and *age of participants*. Treatments from peer-reviewed studies showed a mean effect on academic achievement of $\Delta_{Glass} = .82$, $p < .01$, whereas the effect of treatments from non-peer-reviewed studies was $\Delta_{Glass} = .23$, $p < .05$. In contrast to experimental interventions that did not significantly affect academic achievement, quasi-experimental interventions showed an effect of $\Delta_{Glass} = .74$, $p < .01$. Also, in contrast to interventions delivered by teachers, which revealed an effect of $\Delta_{Glass} = .85$, $p < .01$, and interventions delivered by researchers, which showed an effect of $\Delta_{Glass} =.55$, $p < .01$ on academic achievement, treatments delivered by computers, paper, or humans and paper did not have significant effects. SRL treatments that focused on learners between the ages of 9 to 13 showed an effect of $\Delta_{Glass} = .81$, $p < .01$, whereas older learners between the ages of 19 to 37, $\Delta_{Glass} = .33$, $p < .05$ benefited less, and adolescent learners between the age of 14 to 18 did not profit at all.

To further investigate variables for which hypothesis-driven one-way moderator analyses had not revealed a satisfying model fit, an exploratory approach was followed. Analyzing dependencies between moderators, an attempt was made to resolve heterogeneity within categories of moderators by crossing them with categories of other moderators, whereas homogeneous categories were left untouched. Satisfying model fits could be achieved for crossing the variables *SRL layer, type of support,* and *domain of learning* with *age of participants.*

Splitting the group of treatments that supported processes on the metacognitive and cognitive layers into the three age groups, treatments that focused on learners between the ages of 9 to 13 years showed a very high effect on academic achievement of $\Delta_{Glass} = 1.30$, $p < .01$, whereas the other groups did not show significant effects on achievement scores. For all age groups, metacognitive treatments showed an effect of $\Delta_{Glass} = .55$, $p < .05$, and metacognitive, cognitive, and motivational treatments of $\Delta_{Glass} = .43$, $p < .01$ on academic achievement, whereas metacognitive and motivational treatments were not effective.

Interventions that conducted strategy instructions on young learners between the ages of 9 to 13 years did have an effect on performance of $\Delta_{Glass} = 1.21$, $p < .01$, whereas adults between the ages of 19 to 37 years did not significantly benefit from this type of support. As for process support on adult learners between the ages of 19 to 37, an effect of $\Delta_{Glass} = .45$, $p < .05$ was found, and for adolescent learners between the ages of 14 to 18, an effect of $\Delta_{Glass} = .08$, $p > .05$ was found; thus, the pattern for process-support seemed to point in the opposite direction. Interventions that combined a strategy instruction with a process support did not show significant effects on performance.

Treatments conducted on 9 to 13 year-old learners in a mathematical context showed a very high effect on academic achievement of $\Delta_{Glass} = 1.09$, $p < .01$, whereas treatments conducted in the same context were not effective for learners between the ages of 14 to 18. Treatments conducted on 9 to 13 year-old learners in a language context significantly affected academic achievement by $\Delta_{Glass} = 0.72$, $p < .01$. This was also the case for SRL treatments in a science context across all age groups, $\Delta_{Glass} = 0.49$, $p < .01$. Treatments that took place in other contexts were not effective.

For SRL level of intervention, duration of intervention, and measure of academic achievement, no model fit could be found. Hence, the results should not be generalized, but they indicate the main characteristics of the sample of studies (Shadish & Haddock, 1994). Interventions that focused on the micro level of elementary tasks had a large effect of $\Delta_{Glass} = .56$, p < .01 on academic achievement, in contrast to the lack of effectiveness found for interventions that focused on the mid level of daily study routines, as well as on the micro and mid levels combined. Interventions that lasted less than 1 hour and interventions that took place on a single day were not effective, whereas interventions of longer duration had significant effects on achievement. With regard to the learning outcome, significant effects of SRL interventions could be found on grades and undefined achievement measures, $\Delta_{Glass} = .44$, p < .01, as well as multimedia-based comprehension, $\Delta_{Glass} = .92$, p < .01, but not on problem solving, multimedia-based knowledge, and writing quality.

2.3 Conclusion

With respect to the first research question of this PhD thesis, it can be stated that SRL interventions have positively affected academic performance by $\Delta_{Glass} = .45$. Accordingly, synthesizing SRL research provided evidence that the enhancement of SRL processes during learning improves the quality of the learning outcome. With respect to the second research question, when considering only one moderator, two indicators for the specific design of SRL treatments could be found. Treatments conducted with young learners and interventions delivered by teachers were highly effective. In addition, peer review status and research design were influential variables. When analyzing combined effects, treatments that focused on the metacognitive and cognitive layers, strategy instructions, and treatments conducted within mathematical learning environments turned out to be most effective for young learners.

3. Study 2: Improving the Quality of E-Learning. Scaffolding Self-Regulated Learning on the World Wide Web.

Aiming to improve the quality of the learning process and the learning outcome in e-learning, using web-based learning as the specific scenario, and thereby to answer the third and the fourth research questions of this PhD thesis, the author carried out an SRL intervention. Based on the sketched scaffolding approach, the WWW was optimized as a learning environment by embedding a sidebar called *E-Learning knoWledge Management System* (ELWMS) on the left-hand side of the Firefox web browser. ELWMS provides integrated scaffolds that serve as tools to complete a learning task on the WWW, and that upon application induce the deployment of the six metacognitive processes of goal setting and planning, self-monitoring and process-regulation, and reflection and modification. In the standard version of ELWMS, the scaffolds were offered in a nonembedded way (Clarebout & Elen, 2006), leaving the decision of if, how, and when to apply them during web-based learning to the learners. To investigate whether learners would profit from additional prompting, an extended version of ELWMS was created that supplemented the standard version of ELWMS by two invasive prompts that directed learners to become involved in goal-setting and planning processes in the preaction phase and in reflection processes in the postaction phase.

3.1 Method

The study was carried out in one session that lasted 110 min. Bachelor of Science Psychology students (*N* = 64; mean age: 23.1 years) were randomly assigned to work with either the standard version of ELWMS, the extended version of ELWMS, Firefox and pen and paper, or just Firefox. After a 5-min introduction into either ELWMS or Firefox, in a 45-min learning period, participants were required to conduct a micro-level task, learning information about the topic of *Classical Antiquity* on Wikipedia. Besides the common demographic variables, metacognitive skills, measured by adapted scales from the Motivated Strategies for Learning Questionnaire (MSLQ; Pintrich, Smith, Garcia, & McKeachie, 1993) and the Volitional Components Questionnaire II (VCQII; Kuhl & Fuhrmann, 1998), and computer literacy, measured by a self-developed questionnaire, were assessed with a pretest. To evaluate the quality of the learning process, for all conditions, *offline* self-reports on the processes that learners had carried out during task implementation were assessed by a questionnaire that was developed by the author based on MSLQ and VCQ II scales. In addition, *online* log data were collected, tracking overt actions during task implementation. Detailed video analyses, which served to establish quantitative and qualitative metacognitive scales, could only be conducted for experimental groups, as the two control groups working with Firefox performed actions of a different nature. This multi-method approach allowed for validating *offline* and *online* measures. The quality of the learning outcome was assessed for all conditions by gain in factual knowledge on the topic of *Classical Antiquity* by applying an achievement pre- and posttest. For the two experimental conditions, the quality of the goal-resource structure that the learner had created during task implementation was assessed.

3.2 Results

With regard to the quality of the learning processes, participants of both experimental groups who had received indirect scaffolding of SRL during the implementation of the web-based learning task deployed more SRL processes, more metacognitive processes, and specifically more process-regulation processes than participants in control conditions. In addition, experimental groups experienced more positive emotions. Also, the two experimental groups deployed more planning and reflection processes, and visited a smaller number of web pages than the group that worked just with Firefox, whereas they were more motivated than the group that worked with Firefox and pen and paper. With regard to the quality of the learning outcome, as all groups gained equal amounts of factual knowledge on the topic of *Classical Antiquity*, differences in the performances on the achievement test were not found.

Comparing the two experimental conditions on the quality of the learning process, the group that was prompted in an invasive and directive way to deploy goal setting, planning, and reflection carried out fewer self-monitoring processes in general, fewer self-monitoring processes with relevance to the achievement test, and browsed fewer web pages than the group that was free to apply the scaffolds of their own accord. However, learners who were prompted carried out more reflection processes, and more reflection processes with relevance to the achievement test. Regarding the quality of the learning outcome, individuals who did not receive prompts created a goal-resource structure of higher quality, whereas on the achievement posttest, no differences between groups were observed.

3.3 Conclusion

With respect to the third research question of this PhD thesis, this intervention study indicated that optimizing the WWW as a learning environment—by offering scaffolds that serve as tools to complete a learning task and simultaneously induce metacognitive processes—enhanced the quality of the learning process. However, as learners who worked with ELWMS did not employ more goal-setting and self-monitoring processes, they were not able to attain a higher gain in factual knowledge than learners who were not scaffolded. With respect to the fourth research question of this PhD thesis, the results of this study indicated that scaffolds that combine functionality and metacognitive support do not necessarily have to be supplemented by invasive and directive prompts. As learners suppressed self-monitoring processes that were not prompted, they did not manage to create a structure of higher quality. As a consequence, when following the invasive and directive reflection prompt, they were missing the basis by which to actually gain more factual knowledge than learners who were free to apply the scaffolds of their own accord.

4. Study 3: Improving the Quality of Learning on the World Wide Web by Scaffolding Self-Regulated Learning.

The third study was designed to investigate the third and the fourth research questions of this PhD thesis more deeply. With a revised version of the ELWMS software and an elaborated study design, the general objective was to enhance learners' deployment of the metacognitive processes that had not been supported effectively in the second study. Therefore, the goal was to enhance the quality of the learning process, and consequently, to enhance the quality of the learning outcome. Further, great effort was put into optimizing the instruments that were used to assess the dependent variables.

Referring to the third research question, the second study had provided evidence that integrating scaffolds that embody functionality and provide metacognitive process support into the Firefox web browser constituted a powerful concept for enhancing the quality of learning on the WWW. However, learners who had received indirect support did not employ more goal-setting and self-monitoring processes, which hindered them from accumulating greater factual knowledge than learners who were not scaffolded. To enhance the effectiveness of the ELWMS goal-setting support—which, in the second study was assumed to be ineffective due to learners' erroneous implementation of prerequisite processes—in the third study, learners received feedback on the knowledge gaps that they had experienced on the achievement pretest. In addition, besides optimizing its usability and appearance, ELWMS was equipped with a goal-activation function to enhance the employment of self-monitoring processes during web-based learning.

Referring to the fourth research question, the second study provided evidence that learners profited from the invasive and directive reflection prompt, but failed to deploy self-monitoring processes. To provide learners with more intensive prompting, for the third study, the extended version of ELWMS was revised. Reflecting on the results of the second study, nine processes that were considered to enhance achievement in web-based learning were identified. To help learners to apply the scaffolds as intended, and thereby to promote the deployment of six achievement-enhancing processes, learners were prompted in an invasive and directive way to set goals in the preaction phase, in a noninvasive and directive way to set relevant goals, to activate the current goal, to check the relevance of resources, and to check the goal-resource fit, and in an invasive and directive way, to prepare for the posttest in the postaction phase.

4.1 Method

To evaluate the quality of the learning process of web-based learning, a synchronized multi-method approach was followed by creating corresponding instruments that were aligned to assess identical processes. On the one hand, based on the SRL posttest that had been used in the second study, an *offline* questionnaire was developed to assess participants' reports on the deployment of the achievement-enhancing processes and SRL processes. This questionnaire was applied as a pretest, embedded in a web-based learning scenario, and as a posttest, asking participants to indicate whether they had deployed the processes during the implementation of the task. On the other hand, objective *online* data were used to assess whether participants had carried out the processes of interest. In contrast to the second study, in which most of the *online* data had been established by conducting time-consuming video analyses, in the third study, *online* data were automatically generated by an improved methodology of collecting log data. In this regard, ELWMS was equipped with a function to view single goals and resources, which allowed for logging the process of self-monitoring more thoroughly. Further, in the second study, as a drawback of equipping control groups with the Firefox browser to implement real-world WWW learning approaches, overt actions of the control and experimental groups were not comparable. In the third study, to be able to raise comparable log data for all conditions, a downgraded control version of ELWMS was created to simulate standard software for web-based learning. Instead of being equipped with a goal-setting function, participants in control groups were provided with a folder function. Accordingly, the control version of ELWMS allowed for performing overt actions that were comparable to the ones of the standard version, but did not provide metacognitive support. To be able to establish scales for the achievement-enhancing processes, and quantitative and qualitative metacognitive scales in further analyses, log data were analyzed in a qualitative way by assigning logged actions a relevance rating with respect to the achievement test. Pursuing this synchronized multi-method approach of collecting data on the quality of the learning process allowed for extensive analyses of the validity of measures.

To evaluate the quality of the learning outcome of web-based learning on two different levels, a multi-method approach was also pursued. On the one hand, a value for the quality of the goal/folder-resource structure that a participant had created throughout learning on the WWW was established on the basis of automatically generated log data. As the design of the study had been harmonized with the method of collecting data, in contrast to the second study, it was possible to establish this value for all groups. On the other hand, gains in factual knowledge were evaluated in a pretest/posttest design. The achievement test was comprised of a shortened and revised version of the multiple-choice test that had been developed for the second study. In addition, each question that assessed factual knowledge was supplemented by a question that assessed participants' certainty of their answer.

The study was conducted in two sessions. Students from Technische Universitaet Darmstadt (*N* = 108; mean age: 23.5 years) were randomly assigned to work with either the standard version of ELWMS, the extended version of ELWMS, or the control version of ELWMS. On the first day, participants filled out the demographic and psychometric pretests and completed an extensive computer-based training on either the standard or the control version of ELWMS, being obligated to meet a predefined criterion. On the consecutive day, after having completed the achievement pretest and having received feedback on their knowledge gaps, participants conducted a micro-level task, learning on Wikipedia for 45 min about the topic of *Classical Antiquity*. After the learning period, the achievement posttest and the self-report posttest were administered.

4.2 Results

With regard to the quality of the learning process, participants of both experimental groups, who had received indirect scaffolding of SRL during the implementation of the web-based learning task, deployed more metacognitive processes per se and more metacognitive processes with respect to their knowledge gaps. More specifically, in the preaction phase, ELWMS effectively supported the deployment of goal-setting and planning processes per se, and goal-setting and planning processes with relevance to the achievement test. With respect to the achievement enhancing processes, learners who received indirect scaffolding more often approached the web-based learning task in a goal-oriented way, and created more relevant goals. In the postaction phase, ELWMS was effective in supporting reflection processes per se. On the basis of participants' self-reports, no differences between groups were found. With regard to the quality of the learning outcome, as all groups had equally profited from the learning period, differences in the quality of the goal/folder research structure, as well as in the performance on the achievement test, were not found. However, further analyses revealed that structures of experimental groups contained more relevant goals.

Comparing the two experimental conditions on the quality of web-based learning revealed a similar pattern. The group that received intensive prompting of the achievement-enhancing processes in addition to the scaffolds deployed more metacognitive processes per se and more metacognitive processes with respect to their knowledge gaps. More specifically, in the preaction phase, the two goal-setting prompts effectively supported learners on the level of the achievement-enhancing processes to pursue a goal-oriented approach and to set more relevant goals, and on the level of the metacognitive processes to carry out more goal-setting and planning processes per se and more goal-setting and planning processes with relevance to the achievement test. However, in the postaction phase, the reflection prompt enhanced the deployment of not only reflection processes per se, but also of reflection processes with relevance to the achievement test. Again, participants self-reports did not reveal differences between groups. Regarding the quality of the learning outcome, learners who received additional prompts did not establish a goal/folder resource structure of higher quality, nor did they perform better on the achievement posttest. However, the structures of participants again contained more relevant goals.

In an additional analysis, the relevance of the achievement-enhancing processes and the metacognitive processes was further evaluated by determining their predictive value for a gain in factual knowledge. On the level of metacognitive processes, on the basis of quantitative log data and self-reports, *goal setting* was a significant predictor of performance gain, whereas on the basis of qualitative log data, *goal setting* and *self-monitoring* had a significant positive impact and *planning* had a significant negative impact on performance gain. On the level of the achievement-enhancing processes, performance gain could be predicted by log data, with *goal orientation during action* having a significant negative impact and *importing relevant resources* showing a positive impact, but not by self-reports.

4.3 Conclusion

Regarding the third research question of this PhD thesis, this intervention study provided additional evidence that optimizing the WWW as a learning environment—by offering scaffolds that serve as tools to complete a learning task and simultaneously induce metacognitive processes—enhances the quality of the learning process. More specifically, results showed that ELWMS effectively supported learners to approach web-based learning in a high quality way by deploying goal-setting and planning processes. However, as ELWMS did not effectively support learners to engage in self-monitoring and process-regulation, and thereby to continue on the high-quality path, learners failed to accumulate information that related to their knowledge gaps. As a consequence, even though ELWMS effectively supported reflection processes, learners missed the opportunity to actually enhance the quality of their learning outcome. With regard to the fourth research question of this PhD thesis, the results of this study indicated also that supplementing scaffolds that combine functionality and metacognitive support by the prompting of the achievement-enhancing processes promotes the quality of the learning process. More specifically, an invasive and directive prompt as well as a noninvasive and directive prompt fostered the processes of goal setting and planning in the preaction phase, and an invasive and directive prompt enhanced the process of reflection in the postaction phase. However, supplementing the scaffolds by intensive, noninvasive, and directive prompting in the action phase did not result in an enhancement of the processes of self-monitoring and process-regulation. Considering that evidence was found that intensive prompting has the potential to enhance the quality of the learning process during web-based learning, a lack of effect on the quality of the learning outcome could be due to the ineffectiveness of the prompts that were designed to foster the achievement-enhancing processes during the action phase.

5. Discussion

From a psychological point of view, one of the reasons that e-learning scenarios have revealed disappointing results, regarding the quality of learning, is that, along with a technological enrichment of learning environments, individuals experience less guidance and higher degrees of freedom. As a consequence, the skills and strategies that a learner is in possession of become major predictors of the quality of learning. The concept of SRL proposes that individuals' deployment of SRL processes during learning is an indicator for a learning process of high quality, which in turn provides a learning outcome of high quality (Zimmerman & Schunk, 2001). Following this assumption, scientists have created a variety of SRL interventions to enhance individuals' deployment of SRL processes during learning, and, by this means, their achievement in situations of high self-determination.

Aiming to gain insight into the effectiveness of SRL interventions, the author conducted a meta-analysis. With respect to the independent variable, a broad approach was followed by including a wide variety of different kinds of SRL treatments, whereas, regarding the dependent variable, the focus was on academic performance. When creating the pool of studies, special effort was put into identifying unpublished dissertations in order to avoid a publication bias (Rothstein et al., 2005). Also, special effort was put into dissolving heterogeneity within the sample of studies by conducting moderator analyses and by combining effects.

Regarding the first research question of this PhD thesis, the meta-analysis verified assumptions about the relevance of SRL for learning. It was shown that learners who had received an SRL intervention outperformed learners who had not received an SRL intervention by about half a standard deviation. Accordingly, the first result of this dissertation was that enhancing learners' deployment of SRL improves the quality of learning.

Regarding the second research question of this PhD thesis, which was also covered by the meta-analysis, a path toward a framework for creating effective SRL was provided. It was shown that SRL interventions conducted on young learners, interventions delivered by teachers, and interventions that applied a quasi-experimental design greatly affected academic performance. In addition, treatments that focused on the metacognitive and cognitive layers, strategy instructions, and treatments conducted within mathematical learning environments turned out to be highly effective for young learners. Hence, the second result of this dissertation provides future SRL research with guidance for how to establish effective interventions. It also provides information about which features of SRL interventions require optimization and further research.

Aiming to improve the quality of the learning process and the learning outcome in e-learning, using web-based learning as a specific scenario, the author designed and implemented an SRL intervention himself. In doing so, special attention was paid to the application of sophisticated methods and instruments. The quality of the learning process was evaluated by a self-developed self-report questionnaire that assessed SRL in a specific web-based learning context. In addition, screen recordings were conducted, and a methodology for collecting log data was created. Further, extensive qualitative analyses were carried out by assigning raised overt actions a rating of relevance with respect to the achievement test. The quality of the learning outcome was evaluated by the quality of the established structure and an achievement test that was developed based closely on the information from Wikipedia.

Regarding the third research question of this PhD thesis, Studies 2 and 3 provided evidence that optimizing the WWW as a learning environment—by offering scaffolds that serve as tools to complete a learning task and that simultaneously induce metacognitive processes—enhances the quality of the learning process. More specifically, the two studies showed that ELWMS, the manifestation of this scaffolding concept, is highly effective in helping learners to approach web-based learning in a high-quality way, as well as to employ reflection processes toward the end of a learning period. However, the fact that in the second study, only process-regulation could be fostered in the action phase, whereas in the third study, neither self-monitoring nor process-regulation could be fostered, indicates the difficulty of effectively assisting individuals during the ongoing learning process. It is assumed that this lack of the effectiveness of support during the action phase is responsible for the lack of effect on gains in factual knowledge. Accordingly, the third result of this dissertation was that the sketched scaffolding approach improved the quality of the learning process; thus, it has the potential to provide an improvement in the quality of the learning outcome.

Regarding the fourth research question of this PhD thesis, Studies 2 and 3 provided evidence that supplementing scaffolds that combine functionality and metacognitive support by prompts enhances the quality of the learning process. More specifically, the two studies showed that learners will follow invasive and directive prompts when the prompts are administered in the preaction and postaction phases. However, the second study indicated that invasive and directive prompts may have negative effects, as self-monitoring processes, which had not been prompted, were suppressed. In the third study, the intensive, noninvasive, and directive prompting of metacognitive processes during the action phase was not effective.

Both results, again, underline the difficulty of effectively assisting individuals during the ongoing learning process and serve as an explanation for the lack of effect on the quality of the learning outcome. Hence, the fourth result of this dissertation was that supplementing the sketched scaffolding approach by additional prompts improves the quality of the learning process, and has the potential to provide an improvement in the quality of the learning outcome.

Applying the results of the meta-analysis to the SRL intervention that the author conducted himself, several aspects have to be discussed. First of all, according to the findings of the meta-analysis, young learners seem to suffer from mediation deficiencies, and hence benefit most from strategy instructions, whereas older learners seem to suffer from production deficiencies, and therefore benefit most from support that helps them to apply the strategies they are already equipped with. Accordingly, with respect to the SRL intervention of the author, it was appropriate to pursue an indirect approach of assistance by administering process support to adult learners. In addition, as the meta-analysis revealed that interventions focusing on the metacognitive and cognitive layers of SRL were effective for young learners, but not for adults, it was appropriate to administer metacognitive support that had proven to be effective for all age groups. Also, descriptive results of the meta-analysis indicated that the enhancement of SRL is a promising approach for improving the quality of micro-level learning. However, in line with the disappointing findings that had been found for e-learning (Dillon & Gabbard, 1998; Shapiro & Niederhauser, 2004), SRL interventions that were delivered by computers did not affect academic achievement. This was also the case for the second and the third studies of this dissertation, which did not reveal positive effects on the quality of the learning outcome, even though it appears to have been an adequate approach for providing adults with metacognitive process support for micro-level learning.

However, further factors may be responsible for the fact that in the second and the third studies, no effects on the quality of the learning outcome could be found. The meta-analysis provides evidence that studies that are carried out in a quasi-experimental design are highly effective, whereas experimental studies lack effectiveness. In addition, moderators for which no satisfying model fit could be reached suggest that interventions that are carried out in less than 1 hour or on one single day do not seem to suffice to improve academic achievement. Further, the domain of learning as well as the measure of academic achievement might have prevented the detection of effects.

6. Future Perspectives

Having verified that SRL plays a key role for learners' performance, this PhD thesis suggests that future research should continue conducting SRL interventions to enhance the quality of individuals' learning process, and thereby the quality of their learning outcome. This PhD thesis also provides a path toward a framework for creating effective SRL interventions. Without a doubt, one of the major findings is the great influence of the age of participants on the effect of SRL treatments.

With regard to improving the quality of e-learning, it has to be stated that the scaffolding concept that has been implemented in the second and the third studies constitutes a powerful approach for enhancing the quality of learning. Based on this concept, in this dissertation, software was successfully created to provide effective support in the preaction and the postaction phases of learning. It will be the challenge of future research to implement effective support in the action phase in order to attain an improvement in the quality of the learning outcome.

Further, this dissertation provides evidence that prompting, which is provided in addition to scaffolds that are based on the sketched concept, is a promising approach to further enhance the quality of learning. In this dissertation, the author managed to create prompts that provided effective support in the preaction and the postaction phases of learning. With regard to the type of prompting, invasive and directive prompts seemed to be more effective than noninvasive and directive prompts. It will be the task of future research to create prompts that effectively support metacognitive processes in the action phase to attain an improvement in the quality of the learning outcome. Concerning this matter, the advantages and disadvantages of using different types of prompts that vary in intensity should be investigated.

Besides enhancing learners' employment of SRL processes during the ongoing learning process, this dissertation indicates further approaches that can be used to improve the quality of the learning outcome in e-learning. As the meta-analysis revealed that SRL support delivered by humans was highly effective, it would be a very interesting strategy to utilize a human tutor to administer scaffolds that are based on the sketched concept of support in an e-learning setting. Further, as the meta-analysis revealed that quasi-experimental designs had larger effects, the implementation of field studies with greater external validity might be a promising approach. In laboratory settings, many variables are kept constant to be able to isolate effects, which might entail that participants experience fewer degrees of freedom than in natural settings. As a consequence, in laboratory settings, SRL might not be as relevant as in natural settings, and smaller effects on performance might be found. Further, the implementation of longer interventions providing learners with more time, the application of another domain of learning, and the assessment of more complex achievement measures might reveal larger effects on the quality of the learning outcome.

However, features of SRL interventions that in the meta-analysis were identified to have small or no effects on academic achievement should not be abandoned from SRL research. By contrast, it is the challenge of future studies to find ways to make those features more effective. In particular, the disappointing results of SRL interventions that were delivered by computers should not discourage scientists from investigating this very young area of research. In turn, they should be confident that more elaborated concepts of support, along with further technological development, will help to exploit the great potential that computers offer as the instance of delivery, a process that should render computers more effective.

Further, besides dealing with the question of *what* SRL processes to foster, and *how*, *when*, and *to whom* to administer support, it will be the task of future research to reflect on the methods and instruments that are applied to evaluate SRL. In the second and the third studies of this PhD thesis, a synchronized multi-method approach was pursued, evaluating the deployment of a specific process through *online* and *offline* methods. More specifically, using the computer as the instance of delivery allowed the author to conduct screen recordings and to automatically generate log data. In addition, a self-report questionnaire was applied. Both studies provided evidence for convergent and divergent validity, when correlating *offline* and *online* measures. Those results indicate that each method has specific advantages and shortcomings. Accordingly, when evaluating learning quality, SRL researchers should always pursue synchronized multi-method approaches, collecting *online* and *offline* measures, when putting together the puzzle of the “actual learning process.”

With regard to *offline* self-report measures, even though it has been stated many times that SRL is considered to be context specific (Boekaerts, 1999; Zimmerman & Schunk, 2001), current questionnaires like the MSLQ (Pintrich et al., 1993) and the VCQ II (Kuhl & Fuhrmann, 1998) assess SRL in a general way. To be able to evaluate SRL in a web-based learning context, for the second and the third studies of this PhD thesis, the author had to develop an *offline* self-report test himself. It is overdue for SRL research to enter the domain of evaluating SRL with context-specific instruments.

With regard to *online* measures of learning, a precondition for comparing conditions on the basis of overt actions is that participants of all groups are equipped with the same functions when working on a task. In this regard, in the third study of this PhD thesis, the design of the study was harmonized with the method of collecting data by equipping the control group with a downgraded version of ELWMS. However, by doing so, the author created a very strong control condition, which might have entailed a loss of effect on dependent variables. It is the challenge of future research to find ways to generate comparable *online* data for all groups without establishing conditions that are too much alike. A solution to this problem might be to focus on *online* methods that are not dependent on overt actions like thinking-aloud protocols. Nevertheless, the application of more than one *online* method is always desirable. However, with respect to the time-consuming qualitative analyses of overt actions, the results of the second and third studies of this PhD thesis indicate that relying on only quantitative data is not too much of a trade-off.

References

Aleven, V., Stahl, E., Schworm, S., Fischer, F., & Wallace, R. M. (2003). Help seeking and help design in interactive learning environments. *Review of Educational Research, 73(2)*, 277-320. doi:10.3102/00346543073003277

Alexander, P. A. (1997). Mapping the multidimensional nature of domain learning: The interplay of cognitive, motivational and strategic forces. In M. L. Maehr & P. R. Pintrich (Eds.), *Advances in Motivation and Achievement* (Vol. 10, pp. 213-250). Greenwich, CT: JAI Press.

Bannert, M. (2006). Effects of reflection prompts when learning with hypermedia. *Journal of Educational Computing Research, 35(4)*, 359-378. doi:10.2190/94V6-R58H-3367-G388

Boekaerts, M. (1999). Self-regulated learning: Where we are today. *International Journal of Educational Research*, 31(6), 445-457. doi:10.1016/S0883-0355(99)00014-2

Campillo, M. (2006). Acquisition and transfer of a writing revision strategy: A self-regulatory analysis. *Dissertation Abstracts International*, 67, 3-A. (UMI No. Aai3213136)

Clarebout, G., & Elen, J. (2006). Tool use in computer-based learning environments: Towards a research framework. *Computers in Human Behavior, 22(3)*, 389–411. doi:10.1037/h0046016

Dillon, A., & Gabbard, R. (1998). Hypermedia as educational technology: A review of the quantitative research literature on learner comprehension, control, and style. *Review of Educational Research, 68(3)*, 322-349. doi:10.2307/1170600

Flavell, J. H. (1970). Developmental studies of mediated memory. In H. W. Reese & L. P. Lipsitt (Eds.), *Advances in child development and behaviour* (Vol. 5, pp. 181-211). New York: Academic Press.

Friedrich, H. F., & Mandl, H. (1992). Lern- und Denkstrategien – ein Problemaufriss [Learning and thinking strategies: A sketch of the problem]. In H. Mandl & H. F. Friedrich (Eds.), *Lern- und Denkstrategien. Analyse und Intervention* (pp. 3-54). Goettingen, Germany: Hogrefe.

Glass, G. V., McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research.* Beverly Hills, CA: Sage.

Hedges, L. V. (1982). Estimation of effect size from a series of independent experiments, *Psychological Bulletin, 92(2)*, 490-499. doi: 10.1037/0033-2909.92.2.490

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. New York: Academic Press.

Hedges, L. V., & Vevea, J. L. (1998). Fixed- and random-effects models in meta-analysis. *Psychological Methods, 3(4)*, 486-504. doi: 10.1037/1082-989X.3.4.486

Horz, H., Winter, C., Fries, S. (2009). Differential benefits of situated instructional prompts. *Computers in Human Behavior, 25(4)*, 818-828. doi:10.1016/j.chb.2008.07.001

Jonassen, D. (1996). *Computers as mind tools for schools.* Columbus, OH: Merril.

Kramarski, B., & Mizrachi, N. (2006). Online discussion and self-regulated learning: Effects of instructional methods on mathematical literacy. *Journal of Educational Research, 99(4)*, 218-230. doi: 10.3200/JOER.99.4.218-231

Kramarski, B., & Zeichner, O. (2001). Using technology to enhance mathematical reasoning: effects of feedback and self-regulation learning. *Educational Media International, 38(2-3)*, 77-82. doi: 10.1080/09523980110041458

Kuhl, J., & Fuhrmann, A. (1998). Decomposing self-regulation and self-control: The volitional component inventory. In J. Heckhausen & C. S. Dweck (Eds.), *Motivation and self-regulation across the life span* (pp. 15–49). New York: Cambridge University Press.

Land, S. & Zembal-Saul, C. (2003). Scaffolding reflection and articulation of scientific explanations in a data-rich, project-based learning environment: An investigation of progress portfolio. *Educational Technology Research & Development, 51(4)*, 65-84. doi:10.1007/BF02504544

Mosley, J. D. (2006). How does the implementation of self-regulated learning strategies affect the academic achievement of eighth-grade algebra students. *Dissertations Abstracts International, 67*, 1-A. (UMI No. Aai3206366)

Oliver, K., & Hannafin, M.J. (2000). Student management of web-based hypermedia resources during open-ended problem solving. *Journal of Educational Research, 94(2)*, 75-93. doi: 10.1080/00220670009598746

Orwin, R. G. (1994). Evaluating coding decisions. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis (pp. 139-162)*. New York: Russell Sage Foundation.

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In P. R. Pintrich, M. Boekaerts & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 452-502). San Diego, CA: Academic Press.

Pintrich, P. R., Smith, D. A., Garcia, T., & McKeachie, W. (1993). Reliability and predictive validity of the Motivated Strategies for Learning Questionnaire (MSLQ). *Educational and Psychological Measurement, 53(3)*, 801-813. doi:10.1177/0013164493053003024

Reese, H. W. (1962). Verbal mediation as a function of age level. *Psychological Bulletin*, *59(6)*, 502-509. doi:10.1037/h0040739

Rothstein, H. R., Sutton A. J., & Borenstein, M. (2005). *Publication Bias in Meta-Analysis*. West Sussex: Wiley.

Schmitz, B., & Wiese, B.S. (2006). New perspectives for the evaluation of training sessions in self-regulated learning: Time-series analyses of diary data. *Contemporary Educational Psychology, 31(1)*, 64-96. doi:10.1016/j.cedpsych.205.02.002

Schwonke, R., Hauser, S., Nuckles, M., & Renkl, A. (2006). Enhancing computer-supported writing of learning protocols by adaptive prompts. *Computers in Human Behavior, 22(1),* 77-92. doi:10.1016/j.chb.2005.01.002

Shadish, W. R., & Haddock, C. K. (1994). Combining estimates of effect size. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis (pp. 261-281)*. New York: Russell Sage Foundation.

Shapiro, A., & Niederhauser, D. (2004). Learning from hypertext: Research issues and findings. In D. H. Jonassen (Ed.). *Handbook of Research for Education Communications and Technology* (2nd ed., pp. 605-620). Mahwah, NJ: Lawrence Erlbaum Associates.

United Nations. (2008). *The global information society: A statistical view*. Chile: United Nations.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* Cambridge, MA: Harvard University Press.

White, H. D. (1994). Scientific communication and literature retrieval, In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis (pp. 41-55)*. New York: Russell Sage Foundation.

Winne, P., & Hadwin, A. (2008). The weave of motivation and self-regulated learning. In D. Schunk & B. J. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and applications* (pp. 297–314). New York: Taylor & Francis.

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich & M. Zeidner (Eds.), *Handbook of self-regulation (pp. 13-39).* San Diego, CA: Academic Press.

Zimmerman, B. J., & Schunk, D. J. (Eds.). (2001). *Self-regulated learning and academic achievement: Theoretical perspectives*. New York: Lawrence Erlbaum Associates.

Part 2: Original Articles

Study 1: Benz, B. F., & Schmitz, B. (2009). *Self-regulated learning and academic success: Do self-regulated learning interventions enhance performance? A meta-analysis.* Manuscript submitted for publication.

Abstract

This meta-analysis investigated the effect of treatments that were aimed at fostering self-regulated learning (SRL) on academic achievement, and thereby evaluated the relevance of SRL for learning. With regard to the independent variable, a very broad approach was followed by integrating a great variety of interventions. This heterogeneity was addressed by performing categorical moderator analyses, which also allowed for the identification of fruitful properties of SRL interventions, and thereby provided a path toward a framework for creating SRL treatments. Paying great attention to avoid mechanisms of dependency, 44 independent treatments out of 39 studies were integrated, applying a random effects model. A weighted mean effect of SRL interventions on academic performance of $\Delta_{Glass} = .45$ was found. One-way moderator analyses identified peer-review status, research design, instance of delivery, and age of participants as influential variables. Combining predictors revealed that treatments focusing on the metacognitive and cognitive layers, strategy instructions, and interventions within a mathematical domain were most effective for young learners between the ages of 9 and 14. Further, study features that have lacked attention in SRL research were identified.

Keywords: meta-analysis, self-regulated learning, academic achievement, intervention

1. Introduction

From a social cognitive theoretical perspective, SRL is defined as students' self-generated thoughts, feelings and actions that are systematically oriented toward the attainment of their learning goals (Zimmerman & Schunk, 2001). Whenever learning is not guided, but takes place with a certain degree of freedom, the quality of the learning process, as well as the learning outcome, strongly depends on a learner's application of SRL strategies. Following this assumption, scientists have created SRL interventions in order to help learners to master situations of high self-determination. A broad variety of SRL treatments can be perceived, as they differ in terms of the SRL layer, SRL level, type of support, instance of delivery, duration, domain of learning, and age of participants, among other qualities. However, with the reporting of positive (e.g., Kramarski & Mizrachi, 2006a), no (e.g., Mosley, 2006), or even negative (e.g., Campillo, 2006) effects on academic achievement, diverse results have been found by studies that evaluated the impact of SRL interventions on academic achievement. Is this variability in effectiveness due to invalid assumptions about the relevance of SRL for learning? Have researchers managed to create SRL interventions that positively affect academic performance? Which properties of SRL interventions have proven to be effective?

Referring to meta-analyses by Hattie, Biggs, and Purdie (1996), Dignath, Buettner, and Langfeldt (2008), and Dignath and Buettner (2008), we investigated these questions on an aggregated level. In contrast to existing meta-analyses, we followed a very broad approach with respect to the independent variable, and thus incorporated a great variety of existing SRL interventions. We also paid special attention to prevent dependencies between study effect sizes. With respect to the dependent variable, in order to avoid mixing predictors and criteria, we focused on hard measures of academic achievement, and excluded soft measures such as strategy use, motivation, and emotion. The conceptualization of SRL in this meta-analysis was based on the models by Boekaerts (1999), Pintrich (2000), Schmitz and Wiese (2006), Winne and Hadwin (2008), and Zimmerman (2000). Following a prototypical procedure for synthesizing research presented by Cooper (1982), after sketching several aspects of SRL, we will discuss study retrieval, coding of studies, and data analysis, followed by the presentation and discussion of results.

2. Self-Regulated Learning

2.1 Relevance of Self-Regulated Learning

Three levels of learning (Alexander, 1997) that are characterized by different peculiarities of learning problems and appropriate SRL strategies will be sketched in the following section. On the level of elementary tasks (micro level), learners are confronted with basic tasks from several domains such as solving a math problem, dealing with a text, or searching for information on the World Wide Web. Those tasks provide different degrees of freedom. If a person is not equipped with strategies to handle this latitude, poor learning processes and outcomes are often the consequence. When searching for information in an unstructured environment like the World Wide Web, for example, learners tend to lose focus, get lost in hyperspace, or misjudge the trustworthiness of sources of information. In order to perform successfully, SRL encourages learners to identify adequate approaches to the task, to adapt to current requirements, to overcome emerging obstacles, and to optimize strategies based on conclusions from the previous learning process. International comparison studies on school education, like the Program for International Student Assessment (PISA) (Organisation for Economic Co-operation and Development, 2001) and the Third International Mathematics and Science Study (TIMSS) (Beaton et al., 1996), have pointed out the importance of the ability to self-regulate one's learning for successful performance in math, science, and language assessments.

Besides the level of elementary tasks, SRL plays an important role in managing one's daily study routine (mid level). Within formal educational systems—like schools and universities, but also vocational learning settings—learners are commonly dominated by upcoming deadlines. However, the process of preparation usually strongly relies on the learner's own responsibility. While working on one task, the learner generally is exposed to further tasks from other origins, several distracting stimuli, and a constant lack of time. If not equipped with adequate SRL strategies, helplessness due to excessive demands, stress, negative emotions, and a lack of motivation are oftentimes experienced. In order to help learners to cope with those hassles, SRL induces learners to, for example, get involved in efficient time management, focus on priorities, create productive and undisturbed learning environments, and to develop positive study habits.

On a third level (macro level), SRL plays an important role for non-formal education, as well as life and career management. The current period of post-industrialization confronts human beings with many challenges in terms of a rapidly changing and enormously growing knowledge base. After learners have left formal educational systems, they are obliged to improve their knowledge, skills, and personalities as their own responsibilities. To organize their careers or to simply remain able to participate in modern everyday life, learners have to constantly identify their weak spots, to organize and to carry out the closure of those gaps, as well as to apply their newly acquired competences to their daily routines. Again, it is the goal of SRL to provide members of modern societies with essential strategies to be able to master those challenges of life-long learning.

2.2 Models of Self-Regulated Learning

Accompanied by a growing awareness about the importance of SRL, there has been quite a lot of activity regarding the creation of SRL models in the past 2 decades. Most SRL models show some degree of relation, but also focus on different aspects. Boekaerts (1999) describes three systems of self-regulated learning: a metacognitive, a cognitive, and a motivational system that a learner has to regulate in order to perform successfully. Zimmerman (2000) takes a process view of SRL by subdividing the learning process into a forethought, a performance or volitional control and a self-reflection phase. Schmitz and Wiese (2006) adopt the distinction of these three phases of learning, but focus on states by allocating learning processes within the preaction, action, and postaction phase. Winne and Hadwin (2008) take a process view as well, and define students' activities in terms of five features that describe how a learner COPES with a task. A combination of a process view and the systems proposed by Boekaerts is suggested by Pintrich (2000).

Due to its relevance for this paper, the Boekaerts (1999) model is described in detail (see Figure 1) in the following. As mentioned above, the model postulates three systems that a learner has to be able to regulate in order to learn successfully when equipped with freedom of action. The outer layer of the model is concerned with the motivational system, which can be described metaphorically as the engine of learning. In order to even start learning, but also to stay on task, to overcome obstacles, and negative motivational states a learner has to be able to regulate his motivation by deploying effective strategies. Learning also depends on applied cognitive strategies, like calculating a math task or reading a text. Cognitive strategies, which are located at the mid-layer of the Boekaerts model, can be described as the basic armamentarium a learner has to posses to be able to solve the task. The inner layer of the model is concerned with regulating metacognition. In order to organize the learning process the learner defines learning goals and plans how to achieve them, monitors his learning, and engages in regulation during the learning process if a discrepancy from an earlier defined standard is perceived. He finally defines intentions for modifying future learning processes based on reflections on the previous learning process. Boekaerts has pointed out that for a successful performance, all three systems have to be regulated. A failure in one system cannot be compensated for by the others.

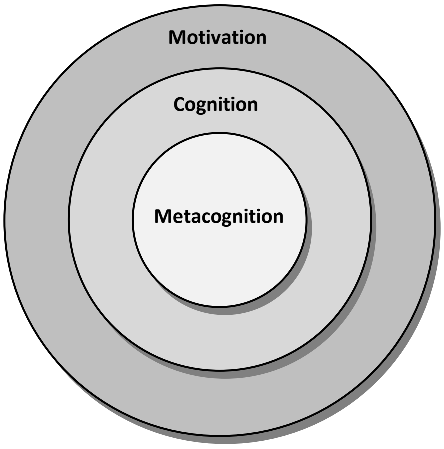


Figure 1. Model of self-regulated learning by Boekaerts (1999).

Obviously, cognitive strategies are highly domain specific. A learner, who is able to solve a math task, does not necessarily possess effective strategies for dealing with a text. Further, it has been argued that metacognitive strategies might also be domain specific (Veenman, Van Hout-Wolters, & Afflerbach, 2006). Being able to plan, monitor, and reflect on a math task does not automatically mean that a student would be able to perform the same processes when working with a text. The same pattern might be the case for motivational strategies. In addition to a domain specification of SRL strategies Benz, Polushkina, Schmitz, and Bruder (2007) have suggested a level specification of SRL competence, referring to the levels that have been sketched in the previous section. Simply because a learner is able to monitor, and to regulate his actions during the solution of a math task, does not necessarily imply that he would be able to organize his daily study routine or to become successfully involved in life-long learning. However, more research has to be conducted in order to investigate level specification of SRL, as well as its possible dependency on age or development of SRL competence.

2.3 Investigating SRL

As sketched in the previous sections, SRL research has focused on two major matters of interest. On the one hand, theorists have investigated SRL and its interdependencies with other variables. On the other hand, by exploring ways to equip learners with SRL strategies, as well as to support SRL strategy application, a more practical approach has been followed. When examining existing research, a remarkably high number of correlational studies dealing with the SRL construct can be found. Scientists administer questionnaires, like the motivated strategies for learning questionnaire (MSLQ) by Pintrich and DeGroot (1990), in order to assess SRL variables and to further analyze their relationship to other variables, like performance measures or self-reports on the learning process (e.g., Colorado, 2006). However, correlational studies, even though they are relatively easy to conduct, bring along the commonly known weakness of non-causality. Hence, to further expand our understanding of SRL, as well as its interdependencies with other variables (e.g., academic performance), the implementation of intervention studies with elaborated methodological designs (Campbell & Stanley, 1966) is inevitable. However, when evaluating SRL within experimental or quasi-experimental designs in primary research, SRL is always confounded with the design of the specific treatment. Hence, if a dependent variable is not affected by an SRL intervention, it remains unknown whether this was caused by an ineffective treatment, by invalid assumptions about SRL, or both. Following this argumentation, by synthesizing the effects of interventions that were aimed at fostering SRL, meta-analyses simultaneously provide an evaluation of the relevance of SRL for learning by isolating its effect on other variables.

Furthermore, the existence of evaluated and effective SRL treatments is essential for two reasons. First, a basis has to be provided to put researchers into the position to examine SRL through intervention studies. Second, due to the relevance of SRL in everyday life, strategies have to be developed to support learners in dealing with learning environments of high latitude. Regarding the importance of this matter, a striking lack of guidance for the creation of effective SRL treatments can be perceived. At the moment there does not exist a framework, which could help SRL researchers when making design decisions. Figure 2 is the basis for our moderators that refer to the constitution of SRL treatments. It summarizes design decisions that have to be made when developing SRL interventions. Based on the SRL construct, researchers decide what they aim to foster, as well as how, when, and to whom support is provided, and how they intend to evaluate the intervention. Considering the freedom that designers of interventions are confronted with, a huge variety of SRL treatments, as well as contradictory conclusions about effectiveness, is not a surprise. Hence, besides evaluating the SRL construct and its interdependencies with other variables, it is the task of meta-analyses to provide a path on an aggregated level leading toward a framework for the creation of effective SRL interventions.

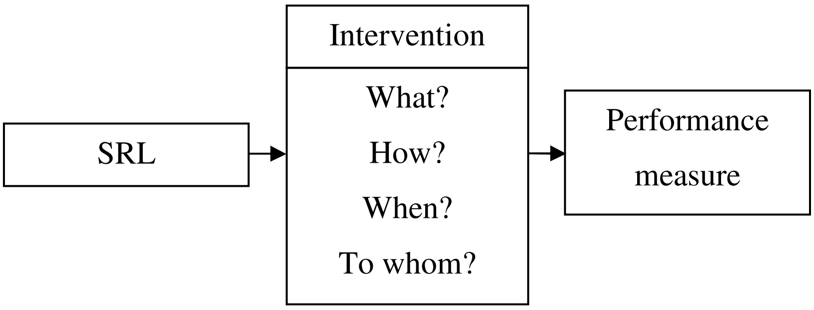


Figure 2. Essential decisions when creating SRL interventions.

2.4 Meta-Analytic Research on SRL

A first approach for investigating these matters on an aggregated level was conducted by Hattie et al. (1996). In a meta-analysis, they integrated 51 interventions on study skills that were aimed at improving students' use of task-related skills, self-management skills, and motivational and affective elements in order to investigate their effects on performance, study skills, and affect. However, as the meta-analysis by Hattie et al. was based on primary research from 1982-1992, our meta-analysis accounts for a distinct set of studies and new developments in research. Further, Hattie et al. chose a somewhat different scope with respect to the independent variable by applying the study skills concept instead of SRL, and, as they did not include interventions that took place within a regular teaching context, by using distinct integration criteria. As a consequence, Hattie et al. applied different sets of search terms for identifying relevant studies. The majority of their interventions were implemented within universities and were mostly based on atypical students (low, high, and underachievers) who voluntarily chose to participate. Since we sought to account for all variations of SRL interventions based on typical learners, a different pattern of studies emerged. With regard to the dependent variable, however, Hattie et al. chose a very broad approach by examining effects on three variables. In order to avoid putting together "apples and oranges" (Bangert-Drowns, 1986), but rather, to focus on a specific research question, in our meta-analysis the dependent variables were narrowed down to measures of academic performance.

Further, including several effect sizes based on one study within the same analysis, Hattie et al. followed another approach of integration by ignoring issues of dependency. In our analysis, we paid great attention to this matter, and thus, only one effect size was calculated per study. The meta-analysis by Hattie et al. revealed a very promising result: The effect of study skills interventions on performance was determined to be $\bar{g}_{Hedges}$ = .57. Also, some specific design recommendations for the creation of interventions could be drawn. Among other results, Hattie et al. found interventions for primary and secondary school students to be more effective than interventions for university students and adults with regard to performance. In general, interventions were most effective if they were taught in relation to content rather than in an all-purpose package, if they were researcher-directed instead of directed by teachers, and if they lasted either 1-2 days or 4-30 days.

A first approach to explicitly synthesizing SRL interventions was published by Dignath, Buettner, et al. (2008) who, focusing on primary schools, analyzed the effects of 48 interventions on strategy use, motivation, affect, and academic performance. In a revision, their analysis was extended to primary and secondary school students up to the tenth grade, and included 72 studies (Dignath & Buettner, 2008). Even though Dignath, Buettner, et al. and Dignath and Buettner applied the Boekaerts model (Boekaerts, 1999) in order to classify interventions as we did, a different scope can be perceived in their studies with respect to the integrated treatments. In contrast to our approach, Dignath, Buettner, et al. and Dignath and Buettner included interventions that were exclusively aimed at fostering cognitive or motivational strategies. In our view, the metacognitive layer, as well as the interaction between the different systems of SRL, constitutes a main aspect of the nature of the concept; thus, we excluded those one-dimensional interventions so that we would not dilute our sample.

Further, Dignath, Buettner, et al. (2008) and Dignath and Buettner (2008) narrowed their selection of studies to interventions that implemented a direct strategy instruction of SRL within a school context. Experimental laboratory settings, computer-based trainings, non-peer-reviewed studies, and very short interventions of one single session or 1 week were excluded. Again, as we followed a very broad approach that was aimed at accounting for all variations of SRL interventions, we did not restrict integration criteria by age, setting, type of support (strategy instruction or process support), or duration of intervention. Further, we put great effort into obtaining grey non-peer-reviewed literature, and included computer-based treatments.

With regard to the outcome, Dignath, Buettner, et al. (2008) and Dignath and Buettner (2008) followed the same broad approach as Hattie et al. (1996) by examining the effects of SRL interventions on three dependent variables. Hence, again, a different pattern of studies emerged for our meta-analysis. A comparison of the studies included in the meta-analysis by Dignath, Buettner, et al. and Dignath and Buettner to our sample reveals an overlap of only two studies (Fuchs et al., 2003; Souvignier & Mokhlesgerami, 2006).

Further, Dignath, Buettner, et al. (2008) and Dignath and Buettner (2008) also calculated more than one effect size per study. It has already been mentioned that in the current analysis, we applied other procedures. The meta-analysis by Dignath, Buettner, et al. revealed a mean effect of SRL interventions on performance of $\bar{g}_{\text{Hedges}} = .62$. Dignath, Buettner, et al. and Dignath and Buettner further found that treatments that fostered metacognitive and motivational strategies were more effective than treatments that fostered metacognitive and cognitive or just metacognitive strategies. Interventions that took place within a mathematical context were more effective than those in reading or writing contexts. This was also the case for measures of mathematical performance in comparison to reading and writing performance. In line with Hattie et al., they found a superiority of researcher-directed to teacher-directed interventions; however, in contrast to Hattie et al., the effect sizes increased with the number of training sessions.

Summing up 17 years of research, it was the intention of our meta-analysis to investigate whether scientists have managed to create SRL treatments that affect performance, to quantify this effect, and at the same time to evaluate the relevance of SRL for learning. In order to do so, we chose a broad approach by including a wide variety of different kinds of SRL treatments. We thus acknowledged this variety of treatments by applying various moderators in order to subdivide the sample of studies into homogeneous groups. Results of moderator analyses allow conclusions about the impact of specific study features and therefore provide a path toward the development of a framework for SRL interventions. In order to systematically guide future research, moderator analyses also provide a descriptive overview of the occurrence of specific study features in current SRL research. To avoid putting together “apples and oranges” (Bangert-Drowns, 1986), but rather to focus on a specific research question, we narrowed the dependent variables to measures of academic performance.

3. Problem Specification

3.1 Proposed Overall Effect

As previously sketched, SRL is currently considered to be a major competence, necessary for performing well in learning environments with a high degree of freedom. However, as has been pointed out, various studies have emerged in which SRL interventions show positive, no, or even negative gains in academic performance for experimental groups compared to control conditions. It was one of the intentions of this meta-analysis to synthesize research on SRL interventions in order to investigate whether SRL interventions do indeed have a positive influence on academic performance, and to further quantify this impact. Based on SRL theory, as well as the existing meta-analyses by Hattie et al. (1996), Dignath, Buettner, et al. (2008), and Dignath and Buettner (2008), it was hypothesized that in sum, SRL interventions would have a positive impact on academic performance measures. Synthesizing the effects of SRL interventions simultaneously provides an avenue for evaluating the importance of SRL for learning.

3.2 Proposed Moderators

When creating SRL interventions and making design decisions, at the moment there does not exist a framework that researchers can rely on when deciding what to deliver and how, when, and to whom to provide SRL treatments. In order to build a basis for developing a framework for the creation of effective SRL interventions, a second research goal was aimed at identifying fruitful properties of SRL interventions. Also, by identifying the frequency of occurrence of these properties in current SRL research, areas that lack attention can be identified in order to investigate SRL more systematically. In the following section, proposed moderators will be described. Before presenting moderators that are concerned with the constitution of SRL interventions, one formal and one methodological moderator is discussed.

3.2.1 Review Status

It has been stated many times (Smith, 1980) that studies reporting large significant effects are more likely being published than studies reporting small or no effects at all. On the other hand, strict review processes may encourage researchers to work more precisely and might thereby cause greater effects to be found. To investigate what has been called the “file drawer problem" (Rosenthal, 1979) or publication bias (Rothstein, Sutton & Borenstein, 2005), review status was proposed to moderate effect sizes. In line with Hattie et al. (1996), who found a large effect for journal articles, a medium effect for books, and no effect for dissertations or theses, studies that have been run through a review process were hypothesized to report greater effect sizes.

3.2.2 Research Design

Various methodologies have been applied in the implementation of SRL interventions. On the one hand, researchers have conducted experimental studies within laboratory settings by randomly assigning participants to conditions (e.g., Azevedo, Cromley, & Seibert, 2004). On the other hand, a remarkable proportion of SRL research has taken place in real classrooms. In this case researchers commonly apply quasi-experimental designs by randomly assigning whole classes to different conditions (e.g., Bruder, 2006). Nonetheless, *research design* and *location of study* are not always confounded, as in some field studies, researchers have randomly assigned learners to conditions by breaking up classes (e.g., Guertler, 2003). To account for the methodological quality of studies (Glass, 1976), following Hattie et al. (1996), *research design* was proposed to moderate effects. Stating a directed hypothesis, however, is an ambivalent matter. Experimental designs, as they prevent nonequivalent groups, are by nature a more sophisticated approach (Campbell & Stanley, 1966). However, as quasi-experimental designs take place in real classrooms, they are confounded with other variables that might boost effects. For example, there is a good chance that classes receiving a treatment will be more motivated than classes selected for control conditions that are simply following their daily routines. Due to such contradictory arguments and unclear results found by Hattie et al., a non-directed hypothesis was applied for this moderator.

3.2.3 SRL Layer of Intervention

In examining SRL interventions, a striking variability of strategies that researchers have created to foster SRL can be observed. These techniques focus on promoting different aspects of SRL. Using the Boekaerts (1999) model to classify existing SRL interventions, treatments applied on the metacognitive, cognitive, and motivational layers can be perceived. Schwartz (1996), for example, conducted a metacognitive intervention by fostering only goal setting. Integrating an instruction on goal setting and planning within the training of writing techniques, Glaser (2005) promoted the metacognitive and cognitive layers. Kauffman

(2002), however, supported all three systems of SRL by administering a matrix in order to support self-monitoring, note taking, and self-efficacy during the reading of a text.

It was a matter of interest for this meta-analysis to investigate the effectiveness of interventions that focused on different SRL layers. In line with Boekaerts' (1999) proposal that all SRL systems have to be regulated in order to perform successfully, it was hypothesized that the effect of an intervention would increase with the number of layers in focus. Therefore, if a treatment focused on promoting metacognition, its effect on performance should be lower than the influence of an intervention that applied support on the metacognitive and cognitive (or motivational) or even the metacognitive, cognitive, and motivational layers. This assumption was also backed up by Hattie et al. (1996), who found that interventions that were taught in relation to content were more effective than all-purpose packages. Dignath, Buettner, et al. (2008) have reported that interventions that promote the metacognitive and motivational layers have been more effective than interventions that focus on the metacognitive and cognitive layers; thus, we expected to find the same pattern. However, in taking a broader look and taking other moderators into consideration, the pattern might change. For example, short evaluations focusing on promoting all three systems might cause an overload and result in poorer performance than short interventions focusing on only metacognition. Additionally, *SRL level of intervention*, *age of participants*, and the way the SRL strategies are promoted and delivered might interact with the SRL system.

3.2.4 SRL Level of Intervention

As was sketched earlier, SRL plays an important role for learning on different levels (Alexander, 1997). Regarding the need of modern learners to be prepared for the challenges on those levels, researchers have created SRL interventions of different qualities. Approaches aimed at fostering SRL on the micro level support the practicing of elementary tasks, for example, by promoting a writing strategy in combination with self-monitoring and self-evaluation (Campillo, 2006). Interventions focused on the mid level try to improve learners' study routines, for example, by instructing them to develop plans and methods for the upcoming semester (McGovern, 2005). Approaches that simultaneously foster SRL on the micro-, as well as on the mid level also exist (e.g., Guan, 1995).

Concerning the hypothesis for *level of intervention*, the same pattern as described for *SRL layer* was expected. Both micro- and mid-level strategies are believed to be necessary in order to perform successfully. Accordingly, learners who receive an intervention focused on both levels should profit the most. Superior results from interventions focused on either level were not expected since both levels are believed to be crucial. However, the measure of academic achievement must be sensitive to the intervention. Obviously, a performance measure that only focuses on elementary tasks would be sensitive only for an intervention on the mid level in an indirect way, if at all, and vice versa. *Duration of the intervention*, *age of participants*, as well as the way the SRL strategies were promoted and delivered might be additional influential variables.

3.2.5 Type of Support

Learners who do not possess a skill that is needed in order to perform well on a task suffer from a mediation deficiency (Reese, 1962), whereas learners who are equipped with the skill, but do not manage to employ it, suffer from a production deficiency (Flavell, 1970). In SRL research, different concepts of support based on assumptions about learner deficits can be perceived. Some researchers focus on overcoming mediation deficits by providing learners with SRL strategies before actually confronting them with a learning task (e.g., Azevedo & Cromley, 2004). Other researchers focus on helping participants to master production deficits by inducing SRL processes during the actual performance of a task (e.g., Kramarski & Zeichner, 2001). Approaches that account for both deficits primarily equip participants with SRL strategies and further provide process support during the task performance (Souvignier & Mokhlesgerami, 2006). Obviously the appropriate SRL support depends on the deficit of the individual learner. Hence, it is not possible to specify a directed hypothesis favoring either strategy instruction or process support.

However, another framework gives a hint for the value of interventions that focus on both deficits. Following Anderson's (1993) Act-R theory, in order to provide learners with skills, primarily declarative knowledge has to be encoded through observation and instruction. By employing declarative knowledge in the context of a problem-solving activity, it is converted into production rules through a process called knowledge compilation. Further practice then produces smoother, more rapid, and less erroneous execution. Hence, a combination of strategy instruction and process support should cover the needs of all participants and therefore provide large effects. Again, *age of participants*, *instance of delivery*, *duration* and *SRL layer of intervention* might be variables that change the pattern.

3.2.6 Instance of Delivery of Intervention

SRL research has employed several ways of delivering interventions. Most treatments have been administered by humans. In these cases, either researchers have visited real classrooms themselves in order to deliver their interventions (e.g., Boone, 1999), or they have instructed teachers to carry out previously mapped SRL programs, like adjusted curricula, within classes (e.g., Gargallo-Lopez, 2001). Other studies have been carried out within computer-based learning environments and have used computers to deliver the interventions (e.g., Berthold, 2006). A third category of delivery that has been observed has used paper-based interventions, such as the administration of a learning strategy brochure (Xiao, 2006).

In a first investigation, we subsumed interventions carried out by researchers and by teachers in order to compare the effect of a treatment delivered by humans to the delivery by computers and by paper. In order to state a hypothesis in this case, several aspects had to be accounted for. Computers have the great benefit of permitting a one-to-one learner-teacher relation, which is hardly realizable with human tutors. However, (nowadays) technology is not capable of providing adaptive learner support of the same quality as a human tutor. Hence, the great potential of the one-to-one relation between learner and tutor is not fully exploited. Accordingly, for the current analysis, we hypothesized that human tutors would be most effective. However, computer-based learning environments have developed rapidly in recent years, and already offer great possibilities in terms of interactivity and different modes of media. Therefore, computer-based learning environments were expected to outperform interventions based on paper. In a second research question, we distinguished between interventions carried out by researchers and by teachers. Following the results of Hattie et al. (1996), Dignath, Buettner, et al. (2008) and Dignath and Buettner (2008), it was hypothesized that interventions conducted by researchers would be more effective than interventions carried out by real teachers.

3.2.7 Duration of Intervention

Treatments involving SRL have differed greatly in terms of duration. Whereas some researchers conduct very short interventions of less than 1 hour (e.g., Cuevas, 2005), others perform treatments across several weeks (Kramarski & Gutman, 2006), months (Lan, 1996) or even years (e.g., Beck, Guldimann & Zutavern, 1994). Even though Dignath and Buettner (2008) found that effect sizes tended to increase with the number of training sessions, formulating a direct hypothesis is a rather complex matter. Longer treatments in general focus on fostering competencies, whereas shorter interventions by nature can only alter states. Furthermore, shorter interventions are usually characterized by a performance measurement within or directly after the treatment, and therefore minimize a dilution of effect. Longer interventions, however, commonly aim to achieve sustainability, and consequently contain a longer gap between support and assessment of performance. Regarding these interdependencies, a negative curvilinear trend, such as the one reported by Hattie et al. (1996), who found large effect sizes for interventions of 1-2 and 4-30 days, does not seem surprising. An undirected hypothesis was stated for the current study, considering that short and long interventions are different by nature.

3.2.8 Domain of Learning

SRL researchers have carried out interventions in many different domains. Whereas interventions in mathematical (e.g., Fuchs et al., 2003) or language (e.g., Trenk-Hinterberger, 2006) contexts are rather common appearances, treatments are also conducted within accounting (Eide, 1999) and PowerPoint® courses (Keith, 2005). Following SRL theory, the relevance of the application of SRL is not determined by the domain of a learning task, but rather by the degree of freedom the learner is confronted with during task performance. However, Dignath and Buettner (2008) found a superiority of treatments conducted within a mathematical context, followed by treatments within reading and writing contexts and treatments within other contexts.

3.2.9 Age of Participants

SRL researchers have chosen various age groups to use when conducting their interventions. Starting somewhere around the age of primary school students (e.g., Walser, 2001), which applies to the sample of studies that Dignath, Buettner, et al. (2008) and Dignath and Buettner (2008) focused on in their meta-analyses, SRL interventions have also been carried out on adolescents, adult learners, or employees (e.g., Leutner & Leopold, 2003). Following SRL theory (Demetriou, 2000), it can be assumed that treatments focused on young learners encounter rather beneficial conditions. Besides the openness of this age group to new approaches, SRL competence is still rudimentary and in the process of development. Instead of having to rebuild undesirable habits, the intervention can be focused on fostering development in a beneficial direction. Researchers who focus on adolescent learners have to cope with a greater variability of preconditions. Furthermore, adolescent learners represent a population that is very resistant to treatments. Adult learners represent the most heterogeneous sample in terms of SRL competence, but should be more open to accepting new approaches than adolescents. Following this line of argument, it was expected that SRL interventions for young learners would be most effective, whereas treatments focusing on adolescents would be least effective. Even though adult learners are not an easy group in terms of openness and formability, it is expected that researchers may encounter very differential conditions. Therefore, the effectiveness of treatments focusing on adults should range somewhere in the middle.

3.2.10 Measure of Academic Achievement

Examining the field of research on SRL interventions, the application of various measures of performance can be perceived. Oftentimes, specific achievement tests are developed by researchers in order to create sensitive measures for particular treatments (e.g., Azevedo, Cromley, Thomas, Seibert, & Tron, 2003). In other cases scientists rely on existing instruments like class exams, or simply apply semester grades (e.g., Masui & De Corte, 2005). By nature, those measures differ in terms of complexity. Whereas a great deal of research focuses on knowledge reproduction (e.g., Parcel, 2005), more complex measures can also be perceived; these measures confront learners with a high degree of freedom by assessing understanding or transfer (e.g., Duke, 2004). Whereas tests that focus on one level of complexity are considered to be homogeneous, tests that assess several levels of complexity provide heterogeneous measures. Furthermore, due to the duration of intervention, as well as the delay of time between treatment and assessment, some interventions assess alterations in *states*, whereas others focus on sustainability by assessing modifications in *traits*.

Our first intention was to regard all variations within performance measures. Unfortunately, when screening the literature we had to abandon this ambitious intention due to a lack of reported data. Following SRL theory, we decided to focus on the degree of freedom that learners were confronted with during the measurement of academic achievement. A similar approach was used by Hattie et al. (1996) who rated the degree of transfer between training task and outcome task. According to Hattie et al., larger effects can be perceived for measures that provide a small degree of freedom. However, referring to SRL theory, measures that grant a high degree of freedom during task performance require a more sophisticated application of SRL strategies. Regarding that experimental groups are supported in applying SRL, they should outperform control groups more extensively for complex than for rather simple tasks.

3.2.11 Interactions Between Moderators

At a first glance, analyzing the interactions between the moderators presented in the previous section appears very promising. For example, the effect of an intervention on a specific SRL layer might very well depend on the age, or specifically the stage of development, of participants. Further factors of influence entailing a relation of even higher complexity are also conceivable. Referring to the sketched relation between *SRL layer* and *age*, the pattern might change again when simultaneously accounting for *type of support* or *instance of delivery*. Without doubt, analyses of interactions would be very desirable in order to provide researchers with specific design guidelines, as well as to build a fruitful basis toward the development of a framework for the creation of effective SRL interventions. However, meta-analysts depend on existing studies.

Accounting for a restricted number and variety of SRL interventions within this rather young area of research, in this meta-analysis we chose an exploratory approach for analyzing interactions. We decided not to state hypotheses, but to combine categories of moderators in order to resolve possible heterogeneous samples of treatments. We also decided to stay on the level of two-way interactions to avoid extremely small cell sizes or even empty cells.

4. Method

4.1 Study Retrieval

4.1.1 Integration Criteria

To create a relevant study pool, criteria for study integration were defined prior to the literature search. In order to avoid comparing “apples and oranges” (Bangert-Drowns, 1986), this meta-analysis used a broad approach in terms of the features of interventions (IV), but applied narrow criteria on the outcome side by focusing on only performance measures (DV). In the following, integration criteria have been categorized by methodology, constitution, and formal characteristics of the study (Stock, 2001; Durlak & Lipsey, 1991). All variations were coded into descriptive and moderator variables.

To meet methodological requirements a study had to apply an *intervention*. Correlational studies were not integrated because they were not capable of addressing the research questions of this meta-analysis. Furthermore, following the "garbage in garbage out" argument (Mansfield & Busse, 1978), studies had to sustain a certain methodological quality. Interventions only met integration criteria if they had *more than 10 participants per condition* and were implemented in an *experimental* or *quasi-experimental design* with a control group that either did not receive a treatment at all or received no SRL treatment.

Regarding the constitution of the intervention, each treatment that utilized an SRL definition that was related to *common SRL models* (e.g., Boekaerts, 1999; Pintrich, 2000; Schmitz & Wiese, 2006; Winne & Hadwin, 2008; Zimmerman, 2000) was potentially suitable for integration. Any study that implemented a treatment referring to Deci and Ryan's (2002) Self-Determination Theory was excluded. However, referring to the Boekaerts model, in order to be integrated, a treatment had to be implemented on at least the *metacognitive layer*. This was the case because, in our opinion, metacognition in SRL constitutes the main component, and is a precondition for the ability to self-control learning. Interventions that focus on only the cognitive and/or motivational layers lack the meta-levels of planning, monitoring, and reflection, and thereby miss the basic components of the concept. Additionally, integrating only cognitive interventions would have opened the study pool to several treatments with totally different scopes, and thereby would have diluted the focus. Hence, only studies that utilized support on the metacognitive; metacognitive and cognitive; metacognitive and motivational; or metacognitive, cognitive and motivational layers, met integration criteria. There was no restriction concerning the *SRL level*, the *type*, *duration* and *domain of support*, the *learning environment*, the *instance of delivery*, or the *age of participants*. However, in order to avoid suffering from loss of generalization, studies utilizing participants with learning disabilities, learning difficulties, or special needs were not integrated. Furthermore, studies had to assess some kind of measure of *academic performance*. Integration criteria were therefore not met by interventions that reported physical skill performance measures like throwing darts (Zimmerman & Kitsantas, 1997) or treatments that were only evaluated by improvement in SRL, or alteration in emotional and motivational factors or stress.

Referring to formal criteria, in order to counteract the "file drawer problem" (Rosenthal, 1979) or publication bias (Rothstein, Sutton, & Borenstein, 2005), all types of publications and papers, *peer-reviewed* or *non-peer-reviewed,* were integrated. Furthermore, studies of *all origins*, whether published in *English or German*, were included. Considering the rising and growing importance of SRL in the 1990s (Boekaerts, 1999), studies that had been published since 1990 were potentially relevant. A first literature search was conducted in summer 2006. Results were updated in spring 2007, resulting in a *time frame* for this meta-analysis from January, 1990 to March, 2007. Additionally, studies had to report *relevant statistical data*. If this was not the case, researchers were contacted by email. The intervention was not included if means and standard deviations for relevant groups could not be obtained.

4.1.2 Literature Search and Acquisition of Literature

In order to identify a pool of relevant studies (Cooper, 1985), the databases PsycInfo and ERIC, as well as the German database Psyndex were searched with the keywords self-reg*, selfreg*, and the German translation selbstreg* (White, 1994). To define search restrictions, in a first step, the first 100 abstracts that were found on PsycInfo without restrictions were analyzed. Based on these studies, search restrictions to narrow down false hits were defined (Reed & Baxter, 1994). These restrictions had to be adapted to the functions provided by each database. All hits were transferred into EndNote (Version 6), which allowed for deleting doublets. Table 1 provides a detailed overview of the literature search.

As presented in Table 1, 2407 abstracts were screened to identify relevant studies. If it was not possible to decide whether a study met integration criteria based on the abstract, the article was acquired in full text. In this meta-analysis, special effort was placed on the acquisition of unpublished dissertations in order to avoid a publication bias. Out of the 154 papers viewed in full text, 51 met the integration criteria. However, due to reasons such as failing to make contact with the researcher, 13 papers lacked statistical data. Accordingly, this meta-analysis was based on 38 papers (see Table 2). One dissertation reported two independent studies, resulting in 39 integrated studies.

Table 1

Overview of Literature Search

Database	Time of search	Restrictions	Search term	Hits	Duplets I	Total hits database	Duplets II	Total hits/ Abstracts viewed
PsycINFO	March 2007	Classification Code: Human Experimental Psychology (23), Animal Experimental and Comparative Psychology (24), Developmental Psychology (28), Educational Psychology (35), Industrial and Organizational Psychology (36)	self-reg*	1521				
		Methodology: Empirical Study	selfreg*	1				
		Publication Year: 1990-2006	Selbstreg*	10	29	1503		

Database	Time of search	Restrictions	Search term	Hits	Duplets I	Total hits database	Duplets II	Total hits/ Abstracts viewed
Psyndex	March 2007	Classification Code: Human Experimental Psychology (23), Animal Experimental and Comparative Psychology (24), Developmental Psychology (28), Educational Psychology (35), Industrial and Organizational Psychology (36)	self-reg*	304				
		Form-/Inhaltstyp: Empirie (Empirical Study)	selfreg*	0				
		Publication Type: 1990-2006	Selbstreg*	242	172	374		
ERIC	March 2007	Publication Type:"Numerical Quantitative Data" OR Publication Type:"Reports Evaluative" OR Publication Type:"Reports Research"	self-reg*	719				
		Publication Date: 1990-2006	selfreg*	9				
			Selbstreg*	0	8	720	190	2407

Table 2
Identification of Relevant Papers

Total hits / abstracts viewed	2407
Papers viewed in full text	154
Papers feasible for integration	51
Papers with missing values	13
Papers integrated (pool)	38

4.2 Coding Studies

4.2.1 Coded Variables

After having identified a pool of studies for integration, studies were run through a data extraction process (Woodworth, 1994). Some variables were assessed for descriptive purposes; others were proposed as moderators and were coded for further analyses. In the next section the coding process is described; the current section briefly sketches a description of the assessed variables. Categories of proposed moderators were partially based on theoretical assumptions and were otherwise iteratively derived from the sample of studies itself.

4.2.1.1 Review status. Studies were identified as *peer-reviewed* if they had been run through a review process. Hence, journal articles were part of this category, whereas dissertations, reports, and speeches/meeting papers were rated as *non-peer-reviewed*.

4.2.1.2 Research design. Referring to Campbell & Stanley (1966), interventions were identified as *experimental* if they had used a control group design with random assignment of participants to conditions. Studies that utilized predefined groups, like existing classes, for experimental and control groups, were considered *quasi-experimental*.

4.2.1.3 SRL layer. Referring to the Boekaerts (1999) model, interventions were coded by the SRL layer they aimed to support. Since interventions that only aimed to foster cognitive and/or motivational components were not included in this analysis, the resulting categories of treatments supporting the *metacognitive*; *metacognitive and cognitive*; *metacognitive and motivational*; as well as the *metacognitive, cognitive, and motivational* layers were coded. As many different concepts of SRL are applied in the field, we did not rely on classifications specified by authors, but rather examined descriptions of interventions in detail.

4.2.1.4 SRL level of intervention. Level of SRL was coded referring to the description provided earlier in this article. Interventions that focused on elementary tasks were considered to be aimed at fostering SRL on the *micro level*. Treatments that were targeted at promoting effective study routines were classified as *mid-level* interventions. In such cases where a treatment focused on both the *micro and the mid level*, both categories were assigned. Priorities to either level were not taken into consideration.

4.2.1.5 Type of support. Interventions were rated as providing *strategy instruction* if they aimed to equip participants with SRL strategies for overcoming a mediation deficit. As a teacher providing help after an explanation was considered part of the introduction to a new topic, this kind of process support was not excluded from the category of strategy instruction. *Process support*, however, was defined as a systematic encouragement to apply SRL strategies in order to overcome a production deficit (e.g., metacognitive-prompts or the use of a learning journal). Interventions that first equipped participants with SRL strategies and then supported strategy application were rated as *strategy instruction and process support*.

4.2.1.6 Instance of delivery of intervention. Interventions were rated also with regard to the way the treatment was delivered. We coded whether a treatment was delivered by a *human*, differentiating between *researcher* and *teacher* when possible, by a *computer*, or by *paper*. We specifically coded for the delivery of the actual intervention, and not the supervision (e.g., if a researcher welcomed participants and instructed them to learn in a computer-based environment, the category was coded as computer-delivered). An intervention was rated as delivered by paper if paper was used as the main medium for instruction or support, for example, a strategy brochure or a paper-based learning journal. Administering handouts in addition to an intervention that was based on another type of delivery did not meet the criteria of a paper-based intervention. However, to account for complex shapes of delivery in existing research, *human and paper* was coded if a paper-based learning journal was provided in addition to an intervention delivered by a human. Based on the pool of studies, coding was iteratively narrowed down to the named categories. Four treatments applied complex shapes of delivery, and therefore were excluded from this analysis. Accounting for all variations would have resulted in very low cell sizes.

4.2.1.7 Duration of intervention. Coding for duration of an intervention revealed the conflict of two partially independent measures. On the one hand, the duration for which participants were exposed to the actual intervention was considered to have an important impact on the effect of an intervention. On the other hand, in terms of giving learners time to adapt to new aspects, the length of the period over which an intervention was carried out also was regarded as important. Generating a new variable based on hours and length of interventions would have resulted in an equality of short intensive and long less-concentrated approaches. To avoid loss of information, we decided to subdivide duration of intervention into two moderators.

Hours of intervention was defined by the time for which participants were actually exposed to an intervention. Time spent for assessments, such as pre- or posttests, was subtracted if specified. If micro process support was provided during learning, like metacognitive prompts, the whole learning period was taken into consideration. Time employed for process support on the mid level was not included in this variable due to the impossibility of specifying hours of intervention (e.g., when providing a learning journal). Three treatments did not provide information on this variable.

Length of intervention was specified as the time across which an intervention was carried out. It also did not include time spent for assessments. One treatment lacked information on this variable. In order to create equal-sized categories for both variables, after coding continuous variables, distributions of those variables within the sample of studies were analyzed. For *hours of intervention* this resulted in the categories of *0 to 1*, *1.1 to 2*, *6 to 9*, *11 to 16*, and *20+* hours. *Length of intervention* was categorized into *1 day*, *3 to 6 weeks*, and *2 to 7 months*, excluding four treatments that included extreme values.

4.2.1.8 Domain of learning. The categories for domain of learning were generated primarily by extracting relevant information from each of the studies. Three major groups of interventions that were conducted within the fields of *mathematics*, *language*, and *science* could be identified. Nine treatments did not fit into this categorization and, since no further similarities could be found, were rated as *other*.

4.2.1.9 Age of participants. In order to create categories for this variable, primarily mean age was extracted from the studies. When studies reported a range of ages, like 16 to 20 years, the mean (18 years) was calculated. If only grades, like 4th graders or undergraduates, were provided, the mean age of students in this grade considering the country of interest was identified. For three treatments, the age of the participants could not be obtained. In order to create equal-sized categories, after having extracted continuous measures, the distribution of the mean age of participants was viewed. On this basis, the categories *9 to 13*, *14 to 18* and *19 to 37* years were established. A variable representing the type of participants, like elementary school students, college students, or employees, wasn't generated due to the lack of reported data and a potential confounding with age.

4.2.1.10 Measure of academic achievement. In order to categorize the complexity of the measure of academic achievement, the degree of freedom learners were confronted with during performance assessment was analyzed. However, due to the different nature of the tasks, general categories for the level of complexity could not be applied. Hence, we followed an iterative approach by first extracting dependent variables and then establishing categories on the basis of our study pool. Studies that did not define the complexity of the achievement test or just applied grades were rated *grade, undefined*. If problem solving was assessed within a mathematical domain of learning, the study was rated *problem solving*. Studies that assessed the quality of texts written by participants were categorized as *writing quality*. Studies that provided participants with a multimedia environment and later assessed performance were either rated as *knowledge-multimedia-based* or *comprehension-multimedia-based* depending on task complexity. Two treatments did not fit into those iteratively obtained categories.

4.2.2 Data Extraction, Coding Process, and Mean Interrater Reliability

In order to determine values for the descriptive variables and the proposed moderators, relevant data were extracted from each study by two independent researchers (Orwin, 1994). During this process, categories of moderators that were not based on theoretical assumptions were iteratively developed. One paper in the sample reported more than one study (Berthold, 2006). In this case, information that referred to the paper as a whole, like *review status*, was assigned to each study within the paper. Further, three studies reported more than one independent treatment (Berthold, 2006; Eide, 1999; Kramarski & Mizrachi, 2006a). The same procedure was applied by also assigning characteristics of the study, like *research design*, to each treatment. Differing data extractions were discussed between researchers and adjusted in consensus. After the relevant information had been extracted from the studies, in a next step, moderators were coded by two independent researchers applying the categories described in the previous section. In cases where the mean of two groups was reported and the conditions were equal in terms of relevant variables (Mosley, 2006; Parcel, 2005; Schwonke, 2005), the two groups were considered as one, and only one rating was conducted. If a treatment did not fit into the categories specified for moderators, it was coded and later excluded from that specific moderator analysis. Again, varying ratings for one variable were settled in consensus. In sum, 44 treatments out of 39 studies and 38 papers were coded. As presented in Table 3, mean interrater reliability was determined by first computing Cohen's Kappa for each moderator, and then accumulating each moderator's Kappa/number of ratings ratio. The mean interrater reliability of .97 can be considered very satisfying.

Table 3

Computed Mean Interrater Reliability: Accumulated Kappa/Number of Ratings Ratio

No.	Moderator	No. of ratings	Interrater reliability (Cohen's Kappa)
1	Review status	44	.95
2	Methodology	44	1.00
3	Age of participants	44	.94
4	Type of support	44	.97
5	SRL layer of intervention	44	1.00
6	SRL level of intervention	44	1.00
7	Instance of delivery of intervention	44	.94
8	Hours of intervention	44	.92
9	Length of intervention	44	1.00
10	Domain of learning	44	.97
11	Measure of academic achievement	44	1.00
	Sum	484	10.69
	Mean interrater reliability: .97		

4.3 Data Analysis

4.3.1 Study Effect Sizes

The selection of a measure for the study effect size strongly depends on the research question (Rosenthal, 1994). Questions about the relation of two variables generate a sample of studies reporting correlations, which can be used as r-effect sizes. Matters concerning differences among conditions, however, require the application of the d-family of effect sizes, which relies on means and standard deviations. The simplest realization of such a differential question is a cross-sectional analysis, which contrasts two measurements that have been assessed at the same time. Longitudinal intervention studies, however, provide the possibility of contrasting pre- and post-measures between, as well as pre-post measures within groups. Though the meta-analyst depends on data assessed and reported in primary analyses, the selection of a specific d-effect size is often determined. Due to a lack of reported pre measures in our sample, we were not able to use an effect size that controlled for a priori discrepancies in groups (Becker, 1988). Instead, our primary studies determined the application of a derivate of Cohen's d, which contrasts post measures of experimental and control groups: $d_{Cohen} = (\bar{x}_1 - \bar{x}_2) / \sigma$ (Cohen, 1968). The absolute difference of the posttest measures is affected by specific characteristics of the applied instruments, and therefore has to be standardized. Since the standard deviation of the population, as proposed by Cohen, is generally unknown, substitutes have been suggested. If the standard deviation of the experimental group is affected by the intervention (e.g., the training of SRL leads to a larger variety in the sample such that there are students who very much benefit from the intervention and students who do not benefit at all), Glass, McGaw, and Smith (1981) suggest the application of the standard deviation of the control group: $\Delta_{Glass} = (\bar{x}_1 - \bar{x}_2) / s_{KG}$.

Hedges and Olkin (1985) propose the use of a pooled standard deviation, which is based on control and experimental groups. An application of this formula is possible if the posttest standard deviations of the two groups do not differ significantly. To test for homogeneity of variance in the sample, we used a formula provided by Hunter and Schmidt (2004). The result did not indicate that the variances of the control and experimental groups were equal; thus, the application of Δ_{Glass} was necessary for this meta-analysis.

Several authors suggest certain adjustments when calculating study effect sizes (e.g., Hedges & Olkin, 1985; Hunter & Schmidt, 2004). The most common one is the small sample adjustment, which should be applied for study sample sizes of fewer than 20 subjects (Hedges, 1981; Lipsey & Wilson, 2001). To ensure a satisfying quality of studies, we did not integrate any results relying on samples of $N < 20$ beforehand, and therefore did not adjust for small samples. Furthermore, we did not apply any procedures to calculate effect sizes based on F and t values or other statistics. The variance of the study effect sizes was determined using a formula provided by Hedges & Olkin (1985).

4.3.2 Dependencies–Selection of Treatments and Measures of Academic Achievement

There are several mechanisms that can cause dependencies between study effect sizes. Some are shown in Table 4 (Hedges, 1990). In psychological research, studies commonly use more than one experimental and but only one control group, assess various dependent variables and report overall scores and subscales. This is also the case for most of the studies in this meta-analysis, as demonstrated in the following example.

Table 4

Selection of Mechanisms that Cause Dependencies Between Study Effect Sizes

Within studies
- Multiple experimental groups and one control group
- Multiple outcome-measures
- Instruments with an overall score and subscales
Between studies
- Multiple studies in one paper
- Studies of one research group

Theoretically, using an equal number of experimental and control groups allows for contrasting each experimental group to a control group. However, in general, the meta-analyst is confronted with the problem that there are more experimental than control groups assigned within one study. Using a control group more than one time for contrasting experimental groups (Δ_{Glass}) results in dependent study effect sizes. To avoid such dependencies, we selected only as many experimental groups from a study as there were control groups (Hedges and Olkin, 1985). We always began by selecting the treatment that had been identified as consisting of the highest number SRL components (Boekaerts, 1999). The following order was applied: *metacognitive, cognitive, and motivational* > *metacognitive and cognitive / metacognitive and motivational* > *metacognitive*. Due to integration criteria, there were no treatments without a metacognitive component. Treatments with *metacognitive and cognitive / metacognitive and motivational* components were defined as equal; however, there was no study in the sample that simultaneously used both combinations of treatments from this level. It can be criticized that selecting more than one treatment from one study, even if different control groups are used for contrasting, results in dependent effect sizes. We do not dispute that different groups from one study have several qualities in common, a condition that might produce some level of dependency. This argument is comparable to the assumption that two studies from one research group might create dependent study effect sizes. However, we argue that on the continuum of dependency there are certain levels that are acceptable, and others that are not. As long as the sample is comprised of different subjects, independence of effect sizes is a reasonable assumption. The same argument is utilized for studies that were presented in the same paper, as was the case for Berthold (2006). Applying this procedure, out of the 39 studies that met integration criteria, three studies were identified that applied more than one independent treatment, resulting in a total of 44 independent treatments.

Table 5

Example of Studies with Different Treatments (IV) and Different Measures of Academic Achievement

Study	Treatment(s) (IV)	Measures of academic achievement (DV)
Azevedo & Cromley (2004)	Metacognitive, cognitive, motivational[a]	Matching Labeling Mental model[a]
Leutner & Leopold (2003)	Cognitive Metacognitive, cognitive[a]	Knowledge[a] Knowledge gain
Perels, Guertler & Schmitz (2005)	Metacognitive, cognitive, motivational[a] Metacognitive, motivational cognitive	Overall result Overall score[a]

[a] Treatments and measures selected for integration.

A comparable problem arises from the assessment of several dependent variables. The studies that are integrated in this meta-analysis report different measures of academic achievement (Table 5). Calculating a study effect size for each measure again results in dependency. Instead of selecting one dependent measure per study and using it to determine the study effect size, it is possible to use several reported measures to calculate more than one effect size per study, and then to create an average study effect size (Durlak, 2000). However, to compute the mean variance of an effect size, the covariances of the measures must be known, a requirement that is only fulfilled for subscales of well-established instruments. Gleser and Olkin (1994) suggest a complex solution to this problem. Following Hedges and Olkin (1985), in this meta-analysis we chose to integrate only one measure of academic achievement per study, which, based on the 44 independent treatments, resulted in 44 study effect sizes. The selection of the dependent variable was applied by choosing the most complex measure of academic achievement; for example, if knowledge and mental models were assessed (Azevedo et al., 2004), the score of the mental model was integrated; if immediate, near, and far transfers were assessed (Fuchs et al., 2003), far transfer was selected. In cases where there were several measures of equal quality reported, like grades in mathematics and Spanish (Gargallo-Lopez, 2001), we either chose the measure that was also reported in other studies of the sample to create a category with a decent number of cases, or else we used random selection.

Keeping in mind that there were several possibilities for choosing treatments, as well as several possibilities for choosing dependent measures, this meta-analysis thus represents only one possible approach for integrating and analyzing the data. Many different meta-analyses are conceivable based on the pool of studies created. To avoid adjustments of probability of error, only one meta-analysis was conducted.

4.3.3 Weighted Mean Effect Size

When integrating the study effect sizes to one mean effect size, different theories of sampling, which do have a direct influence on the computation, can be applied (Hedges & Vevea, 1998). In the fixed effects model (Hedges, 1994), constant study effect sizes are assumed, derived from the true effect of the study population. Therefore, differences in study effect sizes based on this model are due only to the sampling errors of the studies. Based on this assumption, when integrating, each study effect size is weighted with the inverse of its variance divided by the sum of all inverses of the variances of the studies. Therefore, studies with small variances do have a stronger influence on the mean effect size than studies that differ more.

$$weight = \frac{w_i}{\sum_{i=1}^{k} w_i} = \frac{\frac{1}{s^2_{ES_i}}}{\sum_{i=1}^{k} \frac{1}{s^2_{ES_i}}} \qquad (1)$$

Following the random effects model (Raudenbush, 1994), discrepancies of study effect sizes from the mean effect size are not only due to sampling errors, but also to differences in the true effect of the studies (Hedges & Olkin, 1985). Sampling errors cause deviations of drawn study effect sizes from true study effect sizes; true study effect sizes additionally scatter around the mean study effect size. Therefore, in the random effects model, the variance of each study effect size is based on an additional component of variance τ, which accounts for discrepancies between true study effect sizes and the mean study effect size. If this component is not significant, the random effects model results in a fixed effects model: $s^2_{ESi}* = s^2_{ESi} + \tau^2$. In this meta-analysis, we used a random effects model, applying a formula provided by Hedges and Vevea (1998) to calculate the component of variance.

$$\tau^2 = \left[\frac{Q_T - (k-1)}{\sum_{i=1}^{k} w_i - \frac{\sum_{i=1}^{k} w^2{}_i}{\sum_{i=1}^{k} w_i}} \right] \qquad (2)$$

Another approach, which leads to slightly different results, is reported by Hedges and Olkin (1985) and Shadish and Haddock (1994). We further calculated the variance, the standard error, and the confidence interval of the mean effect size by applying formulas by Hedges and Olkin (1985). Significance was calculated by relating the effect size to its standard error (Hedges, 1994). The fail-safe *n* was determined in order to investigate how the effect would diminish if more SRL interventions with no effect on achievement appeared (Rosenthal, 1991).

4.3.4 Tests for Homogeneity

The weighted mean effect size can only be generalized to the population if the studies that it is based on constitute a homogeneous sample. To test for this assumption, we computed the weighted squared discrepancies between the study effect sizes and the weighted mean effect size of all studies (Hedges, 1982).

$$Q_T = \sum_{i=1}^{k} w_i (ES_i - ES_{\bullet})^2 = \sum_{i=1}^{k} \frac{(ES_i - ES_{\bullet})^2}{s^2_{ES_i}} \quad (3)$$

Assuming that there are no significant differences (H_0) between the studies, Q_T follows a chi-square distribution with $df = k - 1$, where k constitutes the number of studies. A significant result indicates a meaningful discrepancy of at least one study. Due to qualitative and quantitative differences between studies, heterogeneous samples are very common. In this case, moderator analyses to identify several homogeneous groups of studies are to be conducted (Hedges and Olkin, 1985). Outlier analyses (Overton, 1998), followed by procedures to adapt extreme values (e.g., Lipsey and Wilson, 2001), are not feasible, since in a random effects model, outliers per definition do not exist. However, to identify unusual studies in the sample, we calculated adjusted standardized residuals, thus classifying studies with values greater than 2 as outliers (Hedges and Olkin, 1985).

4.3.5 Categorical Moderator Analyses in the Random Effects Model

To analyze our proposed moderators and to divide a heterogeneous sample of treatments into homogeneous groups of studies, we conducted categorical moderator analyses using a random effects model. This procedure is comparable to an ANOVA, applying the study effect sizes as a continuous dependent variable and the categorized study features as the independent variable. Following this approach, we artificially categorized continuous moderator variables. To avoid extremely small cell sizes, we excluded categories that contained fewer than four treatments from the particular moderator analyses. This was also the case for treatments that did not provide sufficient information to be assigned to a specific category. Conducting categorical moderator analyses is comparable to the determination of the overall weighted mean effect size; however, calculations are performed separately for each included category of the moderator.

4.3.6 Tests for Homogeneity for Categorical Moderator Analyses

Comparable to the procedure used for the overall analysis, weighted mean effect sizes of each category of the moderators are only interpretable if they are based on a homogeneous subsample of studies. To examine the model fit, as well as to test for meaningful differences between categories, we conducted three tests for homogeneity, applying procedures by Hedges and Olkin (1985). In a first test, homogeneity within each category of the moderators was examined. Secondly, a more conservative omnibus test was conducted by analyzing the overall homogeneity across all categories of the moderators. To analyze our hypotheses that proposed significant differences between categories, another omnibus test was performed (Hedges, 1994). Whereas nonsignificance of the first two tests accounts for homogeneous groups and therefore allows for interpretation of results, a significant outcome of the third test indicates heterogeneity due to significant group differences.

4.3.7 Combined Effects

In order to analyze dependencies between moderators, we chose an exploratory approach. Whereas homogeneous categories of moderators were left untouched, we tried to resolve heterogeneity within categories by crossing them with categories of other moderators, and therefore splitting treatments up into several groups. The same criterion for integration was applied as for one-way analyses; cells containing fewer than four treatments were excluded from the analysis. If a previously homogeneous group in the new model turned out to be heterogeneous, we also applied the illustrated procedure for this category. We did not perform pure two-way or more complex analyses due to small cell sizes.

5. Results

5.1 Study Effect Sizes

Table 6 presents the pool of integrated studies, information on included treatments, number of participants, means and standard deviations of relevant groups, as well as effect sizes and variances of effect sizes. A distribution of the study effect sizes in combination with 95% confidence intervals is presented in Figure 3 (Light, Singer, & Willet, 1994).

Table 6

Treatments Integrated, Descriptive Statistics and Effect Sizes

				Experimental group			Control group			Study effect	
No. of study	No. of treatment	Reference	Description of selected treatment	*N*	*M*	*SD*	*N*	*M*	*SD*	Δ_{Glass}	s^2
1	1	Azevedo & Cromley (2004)	SRL training and disposition of SRL script	63	10.40	2.20	68	8.70	2.70	.77	.03
2	2	Azevedo, Cromley, Thomas, Seibert, & Tron (2003)	SRL training, disposition of SRL script, and support by human tutor	19	9.60	2.70	17	6.90	1.80	1.00	.13
3	3	Azevedo, Cromley, & Seibert (2004)	Provision of 10 learning goals and support by human tutor	17	10.80	2.20	17	8.50	2.20	1.05	.13
4	4	Beck, Guldimann, & Zutavern (1994)	Teaching of five strategies to foster monitoring and reflection; administration of learning journal	236	47.59	20.59	245	50.85	20.15	-.16	.01
5	5	Berthold (2006)	Self-explanation prompts while working on worked out examples	20	4.55	1.20	20	3.63	1.36	.77	.11

				Experimental group			Control group			Study effect	
No. of study	No. of treatment	Reference	Description of selected treatment	*N*	*M*	*SD*	*N*	*M*	*SD*	Δ_{Glass}	s^2
6	6	Berthold (2006)	Pictorial solutions and self-explanation prompts while working on worked out examples	21	.35	.15	21	.44	.22	-.66	.10
6	7	Berthold (2006)	Arithmetical solutions and self-explanation prompts while working on worked out examples	21	.41	.20	22	.49	.18	-.40	.09
6	8	Berthold (2006)	Pictorial and arithmetical solutions and self-explanation prompts while working on worked out examples	21	.36	.19	21	.48	.22	-.63	.10
6	9	Berthold (2006)	Pictorial and arithmetical solutions, integration help, and self-explanation prompts while working on worked out examples	22	.49	.21	21	.46	.21	.14	.09
7	10	Boone (1999)	Instructions to self-direct the use of study skills strategies	25	79.20	10.70	26	73.20	14.90	.56	.08

				Experimental group			Control group			Study effect	
No. of study	No. of treatment	Reference	Description of selected treatment	*N*	*M*	*SD*	*N*	*M*	*SD*	Δ_{Glass}	s^2
8	11	Bruder (2006)	Training of pupils and parents to foster SRL and problem solving; administration of learning journal	20	2.68	1.35	28	1.82	.83	.64	.09
9	12	Campillo (2006)	Training of writing strategy, self-monitoring, and self-evaluation	17	3.15	.86	16	4.66	1.08	-1.76	.17
10	13	Cuevas (2005)	High level elaboration query after each module of an interactive tutorial	17	.55	.18	17	.51	.12	.22	.12
11	14	Duke (2003)	Explanation of evaluation criteria for writing and training of SRL strategies	82	14.13	4.76	82	11.36	4.36	.58	.03
12	15	Eide (1998)	Integration of SRL strategies in an accounting course	27	79.89	8.33	26	76.75	6.20	.38	.08
12	16	Eide (1998)	Integration of SRL strategies in an accounting course	13	77.52	10.86	15	78.13	9.32	-.06	.14

				Experimental group			Control group			Study effect	
No. of study	No. of treatment	Reference	Description of selected treatment	*N*	*M*	*SD*	*N*	*M*	*SD*	Δ_{Glass}	s^2
13	17	Fuchs, Fuchs, Prentice, Burch, Hamlett, Owen, & Schroeter (2003)	Teaching problem solving and support of goal setting and self-evaluation	137	24.99	7.21	120	18.95	3.77	.84	.02
14	18	Gargallo-Lopez (2001)	Training of affective-motivational, information-processing, and metacognitive strategies	23	7.04	1.06	21	6.53	1.26	.48	.09
15	19	Glaser (2005)	Training of writing in combination with SRL strategies	81	13.47	3.28	38	7.02	2.81	1.97	.06
16	20	Guan (1995)	Training of learning strategies, time management, and stress management	73	72.92	12.07	73	67.40	13.12	.46	.03
17	21	Guertler (2003)	Training of SRL strategies and problem solving; administration of learning journal	21	21.86	9.96	25	22.46	5.92	-.06	.09

				Experimental group			Control group			Study effect	
No. of study	No. of treatment	Reference	Description of selected treatment	*N*	*M*	*SD*	*N*	*M*	*SD*	Δ_{Glass}	s^2
18	22	Heo (1999)	Integration of SRL support in learning material	40	9.18	7.74	40	7.73	6.32	.19	.05
19	23	Kauffman (2001)	Support of note taking, self-monitoring, and self-efficacy	16	12.38	2.16	14	8.71	2.58	1.7	.18
20	24	Keith (2005)	Error management training in combination with instructions for strategic questioning	20	12.40	5.20	18	9.50	4.59	.56	.11
21	25	Kramarski & Gutman (2006)	Integration of the IMPROVE method of metacognitive questioning in mathematics class; support through worksheet	35	.82	.17	30	.79	.16	.18	.06
22	26	Kramarski & Hirsch (2003)	Integration of the IMPROVE method of metacognitive questioning in mathematics class	20	76.10	17.01	23	62.61	18.80	.79	.10

				Experimental group			Control group			Study effect	
No. of study	No. of treatment	Reference	Description of selected treatment	*N*	*M*	*SD*	*N*	*M*	*SD*	Δ_{Glass}	s^2
23	27	Kramarski & Mizrachi (2006)	Integration of the IMPROVE method of metacognitive questioning in mathematics class	20	3.20	.80	23	1.42	.38	2.22	.15
23	28	Kramarski & Mizrachi (2006a)	Integration of the IMPROVE method of metacognitive questioning in mathematics class	22	2.44	.70	21	1.04	.81	2.00	.14
24	29	Kramarski & Mizrachi (2006b)	Integration of the IMPROVE method of metacognitive questioning in mathematics class	21	3.10	.91	22	1.54	.51	1.71	.13
25	30	Kramarski & Zeichner (2001)	Support of metacognitive questioning through the IMPROVE method	102	88.10	14.43	84	67.15	18.25	1.45	.03
26	31	Lan (1996)	Administration of self-monitoring journal in statistics course	25	34.95	2.50	19	32.53	3.66	.97	.10

				Experimental group			Control group			Study effect	
No. of study	No. of treatment	Reference	Description of selected treatment	*N*	*M*	*SD*	*N*	*M*	*SD*	Δ_{Glass}	s^2
27	32	Leutner, Barthel, & Schreiber (2001)	Training of self-motivation strategy and SRL	20	78.75	11.80	26	68.65	12.37	.86	.10
28	33	Leutner & Leopold (2003)	Training and support of a learning strategy in combination with SRL	24	.77	.12	21	.78	.15	-.08	.09
29	34	Masui & De Corte (2005)	Promotion of reflection and attribution in university courses	42	59.50	10.30	44	55.40	10.40	.40	.05
30	35	McGovern (2004)	Writing intervention incorporating process-based mental simulation of SRL techniques	14	61.78	10.01	17	69.11	1.59	-.73	.14
31	36	Mosley (2006)	Teaching and encouraging the use of self-regulated strategies for goal setting and individual evaluation; administration of learning journal	41	84.10	9.58	41	84.05	10.67	.01	.05

				Experimental group			Control group			Study effect	
No. of study	No. of treatment	Reference	Description of selected treatment	*N*	*M*	*SD*	*N*	*M*	*SD*	Δ_{Glass}	s^2
32	37	Parcel (2005)	Metacognitive prompts directly before the two posttests	75	10.84	2.92	72	10.69	2.99	.05	.03
33	38	Perels, Gurtler, & Schmitz (2005)	Training of SRL and problem solving; administration of learning journal	67	15.08	5.49	60	15.04	5.82	.01	.03
34	39	Schwartz (1996)	Goal setting instruction; administration of learning journal	37	78.46	8.25	38	75.24	10.12	.39	.05
35	40	Schwonke (2005)	Adaptive prompts while working on a learning protocol	39	22.59	8.31	20	23.30	6.23	-.09	.08
36	41	Souvignier & Mokhlesgerami (2006)	Training of reading strategies in combination with SRL; support through reading plan	95	7.65	2.08	263	7.04	1.93	.29	.01
37	42	Trenk-Hinterberger (2006)	Training of reading strategies in combination with SRL; support through reading plan	164	11.95	2.57	139	10.98	2.93	.38	.01

				Experimental group			Control group			Study effect	
No. of study	No. of treatment	Reference	Description of selected treatment	*N*	*M*	*SD*	*N*	*M*	*SD*	Δ_{Glass}	s^2
38	43	Walser (2000)	Training of writing strategy in combination with SRL	21	2.29	2.92	20	1.40	1.19	.30	.10
39	44	Xiao (2006)	Administration of a language learning strategy brochure	58	.12	.99	59	-.11	.99	.23	.03

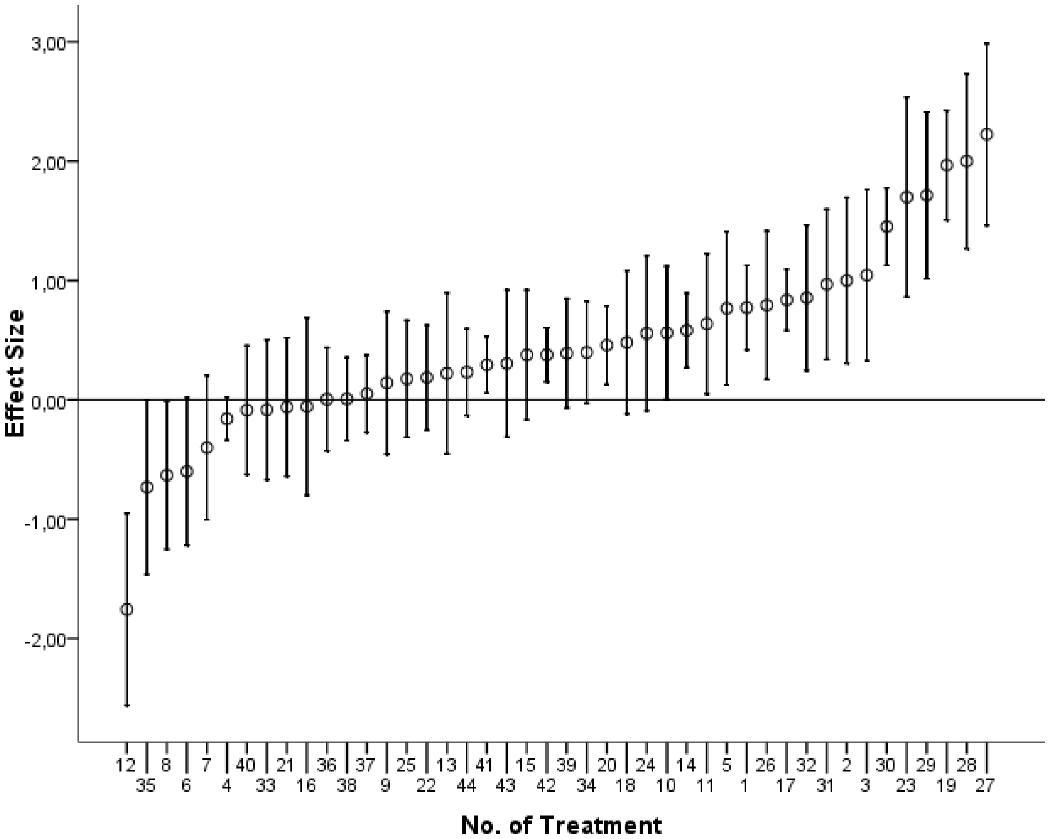


Figure 3. Distribution of study effect sizes and 95% confidence intervals.

5.2 Overall Analysis

As displayed in Table 7, integrating the study effect sizes, which are based on a total of 4047 participants and 44 independent treatments, by applying a random effects model revealed a weighted mean effect size of $\bar{\Delta}_{Glass} = .45$. According to the confidence interval and a large effect size/standard error quotient, the mean effect was highly significant. Its relevance was also underlined by the fail-safe *n*, indicating that the integration of 19 (50) additional studies with no effect would still reveal an effect size of .31 (.21). However, the test for homogeneity, which turned out to be just significant, $\chi^2(43) = 59.3$, $p = .05$, implied heterogeneity within the sample of independent treatments. An inspection of treatments that produced outliers did not reveal any pattern. In sum, we found a significant effect of $\bar{\Delta}_{Glass} = .45$ of SRL treatments on academic achievement; this result is based on a heterogeneous sample and therefore does not serve as an estimate of the population parameter, but is rather presented here as a descriptive result (Shadish & Haddock, 1994).

Table 7

Results of Effect Size Integration Applying a Random Effects Model

No. of integrated papers	38
No. of integrated studies	39
No. of integrated treatments	44
No. of integrated participants	4047
Weighted mean effect size	**.45**
Variance	.0086
Standard error	.0926
Random effect variance	.3002
95 % Confidence interval	[.27, .63]
Test of significance	$z = 4.86$, $p < .01$
Fail-safe N .31 (.21)	19 (50)
Test for homogeneity	$\chi^2(43) = 60.8534$, $p < .05$
Outliers	#12(Δ = -.3.25); #19(Δ = 2.58); #27(Δ = 2.67); #28(Δ = 2.36)

5.3 Categorical Moderator Analysis

As stated in the previous section, even though the weighted mean effect size turned out to be significant, the fit of the overall model did not prove to be satisfactory. In order to analyze the heterogeneity of the sample for systematic patterns, moderator analyses were conducted applying our proposed moderators. Table 8 presents the results, as well as a descriptive overview of the occurrence of specific study features in current SRL research. Again, if the application of a moderator does not result in a good model fit, results cannot serve as estimates of population parameters, but rather serve as an illustration of the mean of the effect sizes integrated in this analysis (Shadish & Haddock, 1994). In the following section, moderators that were able to attain a satisfying model fit will be briefly described.

Table 8

Results of Categorical Moderator Analyses

Moderator / Category	*k*T (*n*P)	ET	Mean Δ	S^2 (*SE*)	95% CI	Q_{wi}	Q_w	Q_{BET}	OT
Review status									
Peer-reviewed	16 (1595)	-	.82**	.020 (.141)	.55, 1.10	$\chi^2(15) = 20.65$, $p > .05$	$\chi^2(42) = 59.22$, $p < .05$	$\chi^2(2) = 249.47$, $p < .01$	27
Non-peer-reviewed	28 (2452)		.23*	.012 (.107)	.02, .44	$\chi^2(27) = 38.57$, $p > .05$			12, 19, 23
Research design									
Experimental	24 (1509)	-	.19	.015 (.124)	-.05, .43	$\chi^2(23) = 27.76$, $p > .05$	$\chi^2(42) = 54.48$, $p > .05$	$\chi^2(2) = 254.21$, $p < .01$	12, 23
Quasi-experimental	20 (2538)		.74**	.017 (.132)	.48, 1.00	$\chi^2(19) = 26.71$, $p > .05$			19, 27
SRL layer of intervention									
Metacognitive	4 (523)	-	.55*	.090 (.301)	-.04, 1.14	$\chi^2(3) = 1.47$, $p > .05$	$\chi^2(40) = 58.54$, $p < .05$	$\chi^2(3) = 250.14$, $p < .01$	-
Metacognitive, cognitive	23 (1909)		.46**	.017 (.132)	.20, .72	$\chi^2(22) = 47.35$, $p < .05$			12, 19, 27, 28

Moderator / Category	*k*T (*n*P)	ET	Mean Δ	S^2 (*SE*)	95% CI	Q_{wi}	Q_w	Q_{BET}	OT
Metacognitive, motivational	4 (214)		.30	.101 (.317)	-.32, .92	$\chi^2(3) = 3.31$, $p > .05$			-
Metacognitive, cognitive, motivational	13 (1401)		.43**	.029 (.172)	.09, .77	$\chi^2(12) = 6.42$, $p > .05$			-
SRL level of intervention									
Micro	29 (2588)		.56**	.012 (.109)	.35, .77	$\chi^2(28) = 57.25$, $p < .01$			8, 12, 19, 27, 28
Mid	4 (232)	-	.18	.087 (.294)	-.40, .76	$\chi^2(3) = 4.00$, $p > .05$	$\chi^2(41) = 63.68$, $p < .05$	$\chi^2(2) = 254.01$, $p < .01$	-
Micro & mid	11 (1227)		.26	.030 (.172)	-.08, .59	$\chi^2(10) = 2.43$, $p > .05$			-
Type of support									
Strategy instruction	18 (1219)		.73**	.019 (.137)	.46, .99	$\chi^2(17) = 36.60$, $p < .01$			12, 19, 27, 28
Process support	15 (1112)	-	.33*	.022 (.149)	.04, .63	$\chi^2(14) = 23.60$, $p > .05$	$\chi^2(41) = 62.36$, $p < .05$	$\chi^2(2) = 246.32$, $p < .01$	23, 30
Strategy instruction & process support	11 (1716)		.18	.027 (.164)	-.14, .50	$\chi^2(10) = 2.16$, $p > .05$			-

Moderator / Category	*k*T (*n*P)	ET	Mean Δ	S^2 (*SE*)	95% CI	Q_{wi}	Q_w	Q_{BET}	OT
Instance of delivery of intervention - A									
Human	16		.76**	.023	.47, 1.06	$\chi^2(15) = 33.40$,			12, 19, 27
	(1182)			(.151)		$p < .05$			
Computer	12	4,	.11	.030	-.23, .45	$\chi^2(11) = 14.13$,			30
	(758)	23,		(.173)		$p > .05$	$\chi^2(36) = 51.84$,	$\chi^2(3) = 205.71$,	
Paper	4	25,	.42	.083	-.14, .99	$\chi^2(3) = 1.08$,	$p < .05$	$p < .01$	-
	(316)	34		(.288)		$p > .05$			
Human & paper	8		.37*	.041	-.03, .76	$\chi^2(7) = 3.22$,			-
	(1129)			(.202)		$p > .05$			
Instance of delivery of intervention - B									
Human (researcher)	7		.55**	.054	09, 1.01	$\chi^2(6) = 17.49$,			12, 19,
	(587)	4,		(.233)		$p < .05$			
Human (teacher)	7	17,	.85**	.059	37, .1.32	$\chi^2(6) = 10.72$,			27
	(295)	23,		(.243)		$p > .05$	$\chi^2(33) = 45.29$,	$\chi^2(4) = 190.75$,	
Computer	12	25,	.11	.032	-.25, .46	$\chi^2(11) = 13.09$,	$p > .05$	$p < .01$	30
	(758)	29,		(.180)		$p > .05$			
Paper	4	34	.42	.090	-.16, .1.01	$\chi^2(3) = 1.00$,			-
	(316)			(.300)		$p > .05$			

Moderator / Category	*k*T (*n*P)	ET	Mean Δ	S^2 (*SE*)	95% CI	Q_{wi}	Q_w	Q_{BET}	OT
Human & paper	8 (1129)		.37*	.044 (.210)	-.04, .78	$\chi^2(7) = 2.99$, $p > .05$			-
Hours of intervention		4, 30, 31					$\chi^2(36) = 57.74$, $p < .05$	$\chi^2(4) = 169.91$, $p < .01$	
0 - 1 hours	11 (545)		.01	.031 (.177)	-.34, .36	$\chi^2(10) = 21.17$, $p < .05$			12, 23
1.1 - 2 hours	7 (528)		.51**	.045 (.213)	.09, .92	$\chi^2(6) = 5.91$, $p > .05$			35
6 - 9 hours	8 (542)		.61**	.040 (.199)	.22, 1.00	$\chi^2(7) = 13.78$, $p > .05$			-
11 - 16 hours	10 (1254)		.59**	.030 (.173)	.26, .93	$\chi^2(9) = 15.22$, $p > .05$			27, 28
20+ hours	5 (467)		.49*	.062 (.248)	.00, .98	$\chi^2(4) = 1.66$, $p > .05$			-
Length of intervention		4, 22, 30, 33,					$\chi^2(36) = 53.91$, $p < .05$	$\chi^2(3) = 162.51$, $p < .01$	
1 day	15 (798)		.24	.023 (.152)	-.06, .54	$\chi^2(14) = 26.88$, $p < .05$			12, 23
3 - 6 weeks	12 (1039)		.78**	.026 (.162)	.47, 1.10	$\chi^2(11) = 23.73$, $p < .05$			19, 27, 28

Moderator / Category	*k*T (*n*P)	ET	Mean Δ	S^2 (*SE*)	95% CI	Q_{wi}	Q_w	Q_{BET}	OT
2 - 7 months	12 (1387)	35	.45**	.026 (.161)	.14, .77	$\chi^2(11) = 3.30$, $p > .05$			-
Domain of learning									
Mathematics	18 (1317)	-	.54**	.020 (.141)	.27, .82	$\chi^2(17) = 32.55$, $p < .05$	$\chi^2(40) = 63.37$, $p < .05$	$\chi^2(3) = 245.32$, $p < .01$	27, 28
Language	8 (1179)		.38*	.042 (.204)	-.02, .78	$\chi^2(7) = 18.16$, $p < .05$			12, 19
Science	9 (697)		.49**	.039 (.197)	.10, .87	$\chi^2(8) = 6.30$, $p > .05$			-
Other	9 (854)		.28	.040 (.201)	-.12, .67	$\chi^2(8) = 6.36$, $p > .05$			23
Age of participants									
9 - 13	14 (1769)	4, 18, 31	.81**	.026 (.161)	.50, 1.13	$\chi^2(13) = 19.74$, $p > .05$	$\chi^2(38) = 50.72$, $p > .05$	$\chi^2(2) = 214.21$, $p < .01$	
14 - 18	11 (879)		.15	.033 (.182)	-.21, .51	$\chi^2(10) = 9.80$, $p > .05$			
19 - 37	16 (830)		.33*	.025 (.158)	.02, .64	$\chi^2(15) = 21.17$, $p > .05$			

Moderator / Category	*k*T (*n*P)	ET	Mean Δ	S^2 (*SE*)	95% CI	Q_{wi}	Q_w	Q_{BET}	OT
Measure of academic achievement									
Grade, undefined	21 (1763)		.44**	.020 (.140)	.17, .72	$\chi^2(20) = 31.17$, $p > .05$			27, 28
Problem solving	7 (666)		.37	.056 (.237)	-.10, .83	$\chi^2(6) = 2.19$, $p > .05$			-
Knowledge multimedia-based	5 (331)	24, 41	.18	.082 (.287)	-.35, .75	$\chi^2(4) = 1.45$, $p > .05$	$\chi^2(37) = 52.67$, $p < .05$	$\chi^2(4) = 255.08$, $p < .01$	-
Comprehension multimedia-based	5 (534)		.92**	.084 (.290)	.35, 1.49	$\chi^2(4) = 2.14$, $p > .05$			-
Writing quality	4 (357)		.40	.103 (.321)	-.23, 1.03	$\chi^2(3) = 15.73$, $p < .01$			12, 19

Note: *k*T = number of treatments; *n*P = number of participants; ET = treatments excluded; Mean Δ = weighted mean effect size; CI = confidence interval; Q_{wi} = homogeneity within groups; Q_w = homogeneity overall groups; Q_{BET} = homogeneity between groups; OT = treatments which produce outliers.

* $p < .05$*. ** $p < .01$.

5.3.1 Review Status

The moderator *review status* split SRL treatments into two homogeneous groups. Treatments from peer-reviewed studies showed a mean effect on academic achievement of Δ_{Glass} = .82, $p < .01$, whereas the effect of treatments from non-peer-reviewed studies was Δ_{Glass} = .23, $p < .05$. As hypothesized, a significant test for homogeneity between groups resulted in relevant differences between categories. However, the omnibus test for homogeneity across all categories of the moderator turned out to be just significant. Figure 4 presents a box plot of the results (Greenhouse & Iyengar, 1994).

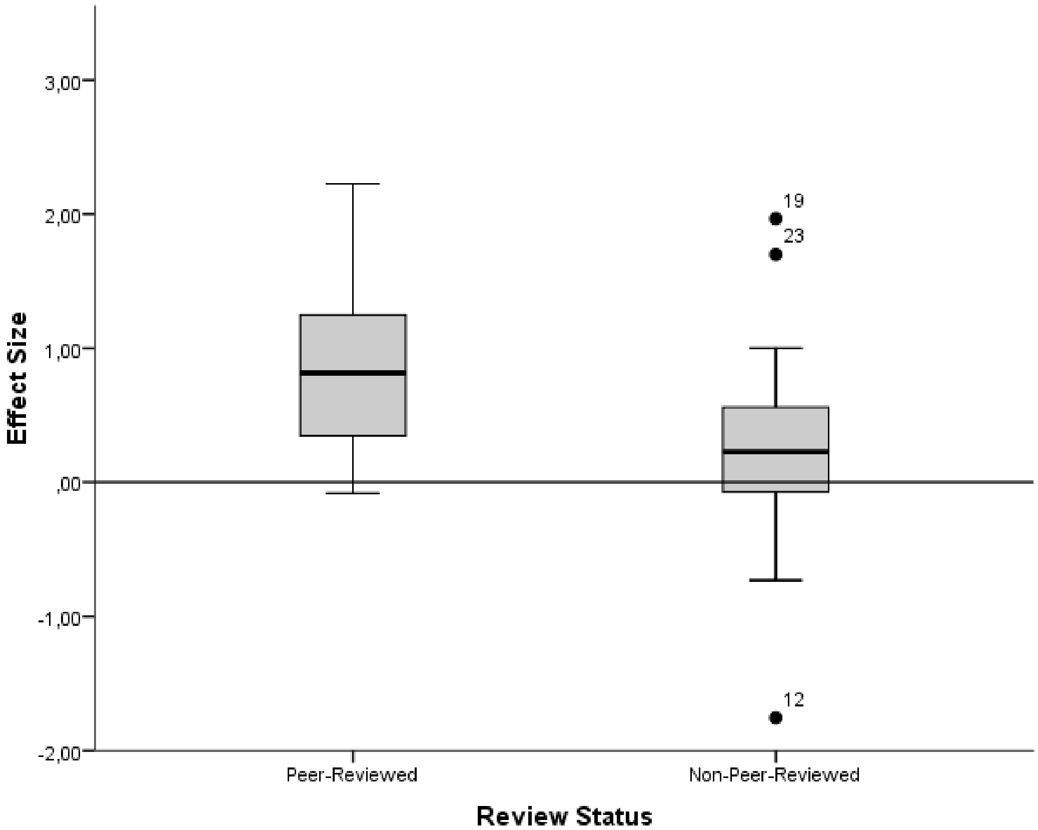


Figure 4. Box plot of the moderator review status. The mean effect size is unweighted.

5.3.2 Research Design

A perfect model fit was reached for the moderator research design as indicated by homogeneous groups, as well as overall homogeneity. Experimental interventions showed an effect on academic achievement of Δ_{Glass} = .19, $p > .05$, and quasi-experimental interventions of Δ_{Glass} = .74, $p < .01$. Both effects differed significantly as indicated by heterogeneity between groups. A box plot of these results is presented in Figure 5.

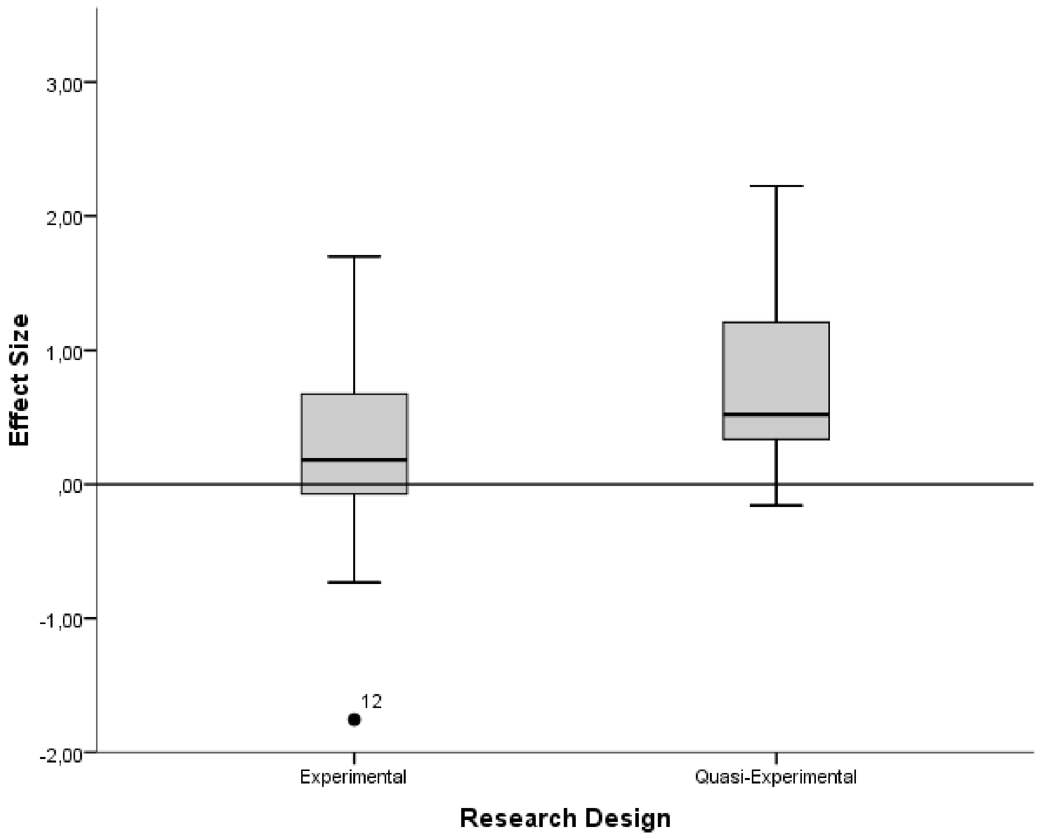


Figure 5. Box plot of the moderator research design. The mean effect size is unweighted.

5.3.3 Instance of Delivery of Intervention

In a first moderator analysis, the category *delivery by a human tutor* was taken as a whole, not distinguishing between researchers and teachers. However, due to a heterogeneous group of treatments delivered by humans, as well as heterogeneity across all categories, no model fit was achieved. In order to dissolve heterogeneity, as well as to be able to compare our results to the findings of Hattie et al. (1996), Dignath, Buettner, et al. (2008) and Dignath and Buettner (2008), in a second analysis, we further distinguished between treatments delivered by researchers and treatments delivered by teachers. Two treatments that had been included in the previous analysis did not report information in detail and had to be excluded. This second analysis revealed homogeneity across all groups. SRL interventions delivered by teachers showed an effect on academic achievement of $\Delta_{Glass} = .85$, $p < .01$, and interventions delivered by researchers of $\Delta_{Glass} = .55$, $p < .01$. Treatments delivered by computers, paper, or humans and paper did not have significant effects. However, the model fit was diminished by a heterogeneous group of interventions delivered by researchers. Hence, the effect of this group should not be taken as an estimate of the population parameter. Figure 6 presents a box plot of the second analysis.

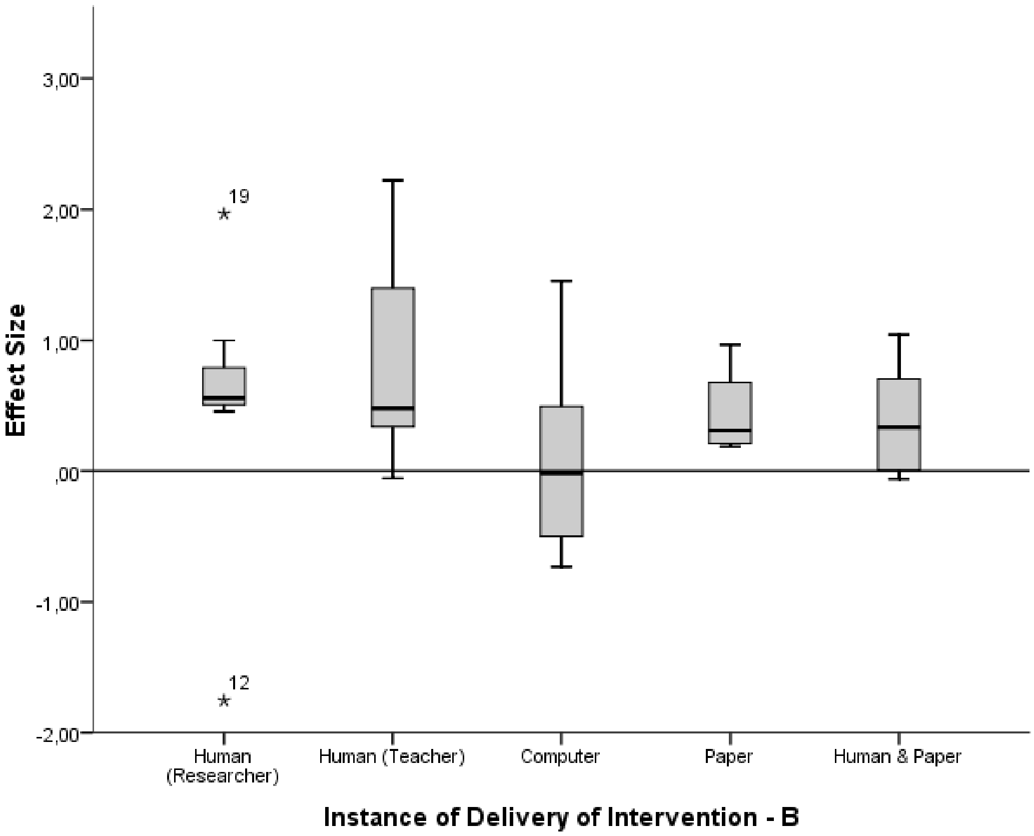


Figure 6. Box plot of the moderator instance of delivery of intervention - B. The mean effect size is unweighted.

5.3.4 Age of Participants

The moderator *age* also resulted in a perfect model fit. Homogeneity was reached within each group, as well as across all groups. The greatest effect of $\Delta_{Glass} = .81$, $p < .01$ was reached by SRL treatments that focused on learners between the ages of 9 to 13. With an effect size of $\Delta_{Glass} = .33$, $p < .05$, older learners between the ages of 19 to 37 benefited less, and adolescent learners between the age of 14 to 18 did not profit at all. As indicated by heterogeneity between groups, all groups differed significantly. A box plot is presented in Figure 7.

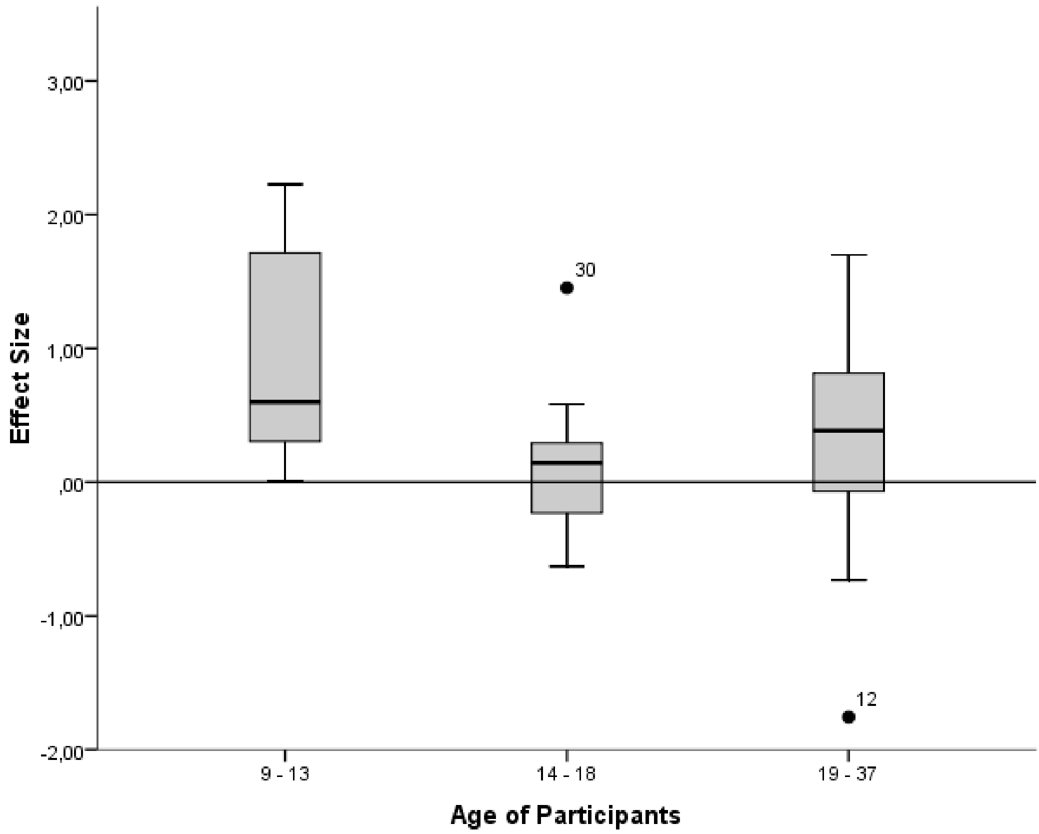


Figure 7. Box plot of the moderator age of participants. The mean effect size is unweighted.

5.4 Combined Effects

We chose an exploratory approach for revealing the reasons for heterogeneity within the groups of the moderators presented in the previous section. In the following sections, three models, which resolve the heterogeneity, are presented.

Table 9

Combined Effects of SRL Layer & Age, Type of Support & Age, and Domain of Learning & Age

Moderators / Categories	*k*T (*n*P)	ET	Mean Δ	S^2 (*SE*)	95% CI	Q_{Wi}	Q_W	Q_{BET}	OT
SRL Layer & Age									
Metacognitive; All Age Groups	4 (523)		.55*	.076 (.276)	.00, 1.09	$\chi^2(3) = 1.74$, $p > .05$			
Metacognitive, Cognitive; 9 - 13	7 (478)		1.30**	.050 (.224)	.86, 1.74	$\chi^2(6) = 11.18$, $p > .05$			
Metacognitive, Cognitive; 14 - 18	8 (667)	4	.14	.040 (.201)	-.25, .54	$\chi^2(7) = 10.82$, $p > .05$	$\chi^2(37) = 47.28$, $p > .05$	$\chi^2(5) = 220.06$, $p < .01$	30
Metacognitive, Cognitive; 19 - 37	7 (283)		.12	.053 (.230)	-.33, .57	$\chi^2(6) = 12.40$, $p > .05$			12
Metacognitive, Motivational; All Age Groups	4 (214)		.30	.086 (.294)	-.27, .88	$\chi^2(3) = 3.79$, $p > .05$			

Moderators / Categories	*k*T (*n*P)	ET	Mean Δ	S^2 *(SE)*	95% CI	Q_{Wi}	Q_W	Q_{BET}	OT
Metacognitive,	13		.43**	.025	.12, .74	$\chi^2(12) = 7.36$,			
Cognitive,	(1401)			(.158)		$p > .05$			
Motivational;									
All Age Groups									
Type of Support & Age									
Strategy Instruction;	8		1.21**	.045	.79, 1.6]	$\chi^2(7) = 12.23$,			
9 - 13	(529)			(.213)		$p > .05$			
Strategy Instruction;	8		.30	.045	-.12, .72	$\chi^2(7) = 13.05$,			12
19 - 37	(482)	14,		(.212)		$p > .05$			
Process Support;	6	17,	.08	.057	-.39, .55	$\chi^2(5) = 9.69$,			30
14 - 18	(436)	18,		(.239)		$p > .05$	$\chi^2(34) = 46.01$,	$\chi^2(4) = 240.88$,	
Process Support;	6	31,	.45*	.065	-.05, .95	$\chi^2(5) = 9.02$,	$p > .05$	$p < .01$	23,
19 - 37	(228)	37		(.255)		$p > .05$			35
Strategy Instruction &	11		.18	.029	-.15, .51	$\chi^2(10) = 2.02$,			
Process Support;	(1716)			(.169)		$p > .05$			
All Age Groups									
Domain of Learning & Age									
Mathematics;	7		1.09**	.050	.65, 1.52	$\chi^2(6) = 11.18$,			38
9 - 13	(604)			(.224)		$p > .05$			

Moderators / Categories	*k*T (*n*P)	ET	Mean Δ	S^2 *(SE)*	95% CI	Q_{Wi}	Q_W	Q_{BET}	OT
Mathematics; 14 - 18	9 (629)	5, 12,	.07	.037 (.193)	-.31, .45	$\chi^2(8) = 9.92$, p > .05			30
Language; 9 - 13	4 (821)	14, 18,	.72**	.076(.276)	.18, 1.26	$\chi^2(3) = 6.43$, p > .05	$\chi^2(33) = 40.60$, p > .05	$\chi^2(4) = 233.87$, p < .01	19
Science; All Age Groups	9 (697)	31, 44	.49**	.038 (.194)	.11, .87	$\chi^2(8) = 6.51$, p > .05			35
Other; All Age Groups	9 (854)		.28	.039 (.198)	-.11, .66	$\chi^2(8) = 6.56$, p > .05			23

Note: *k*T = number of treatments; *n*P = number of participants; ET = treatments excluded; Mean Δ = weighted mean effect size; CI = confidence interval; Q_{wi} = homogeneity within groups; Q_w = homogeneity overall groups; Q_{BET} = homogeneity between groups; OT = treatments which produce outliers.

* $p < .05$*. ** $p < .01$.

5.4.1 SRL Layer and Age

As presented in Table 8, no model fit was reached for the moderator *SRL layer*. The large group of treatments aimed at fostering the metacognitive and cognitive layers constitute a heterogeneous category. Homogeneity across all groups was also not achieved. In order to dissolve heterogeneity within the group of metacognitive and cognitive treatments, the category was split into the three groups of *age*. One treatment did not report adequate information for it to be assigned to one of the subcategories, and was therefore excluded from this analysis. Homogeneous groups were left untouched. Results are presented in Table 9. The formerly heterogeneous group of metacognitive and cognitive treatments by *age* was resolved into three homogenous subgroups. Treatments that supported the metacognitive and cognitive layers and focused on learners between the ages of 9 to 13 years showed a very high effect on academic achievement of $\Delta_{Glass} = 1.30$, $p < .01$. Both other groups did not show significant effects of achievement scores. This variety in effectiveness, which is illustrated in Figure 8, explains heterogeneity within the entire group of metacognitive and cognitive treatments. The effect of metacognitive; metacognitive and motivational; as well as metacognitive, cognitive, and motivational treatments remained the same, as they were not split into subgroups. A perfect model fit was reached as indicated by homogeneity within all groups and homogeneity across all groups.

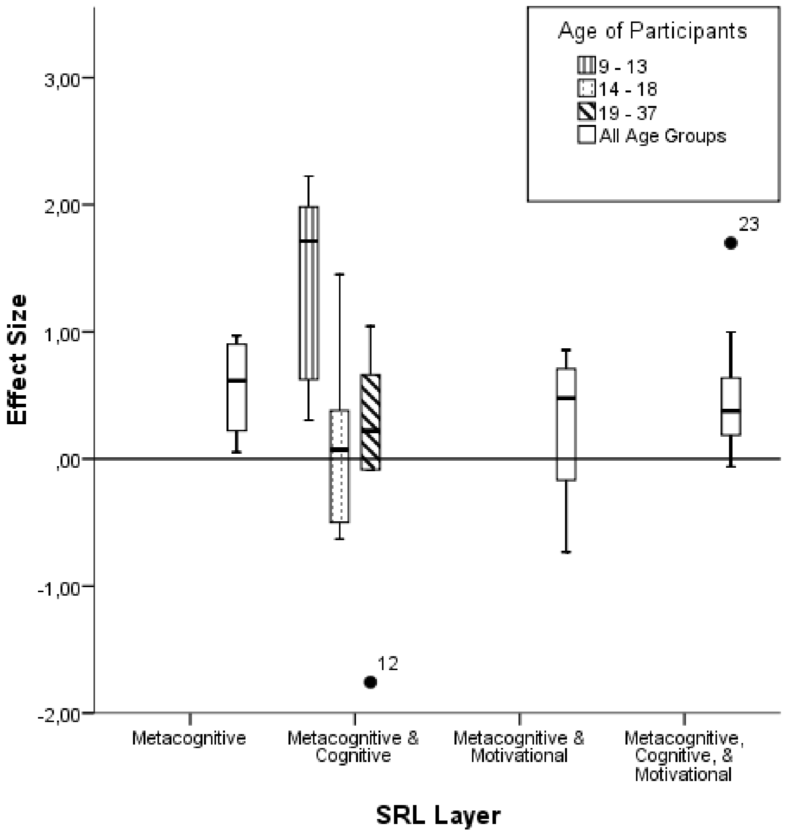


Figure 8. Box plot of the combined moderators SRL layer and age of participants. The mean effect size is unweighted.

5.4.2 Type of Support and Age

Due to heterogeneity across all groups, as well as heterogeneity within the category *strategy instruction*, no model fit was reached for the moderator *type of support*. In order to dissolve the heterogeneous group of strategy instructions, it was combined with the moderator *age*. However, only one strategy instruction was conducted on 14 to 18 year-old learners. This subcategory was therefore excluded from the analysis. This was also the case for another treatment that did not report detailed information for the ages of participants. Hence, in a first step, treatments that used strategy instruction were subdivided into the age groups of 9 to 13 and 19 to 37. The category *process support*, as well as *strategy instruction and process support* were left untouched, as they constituted homogeneous groups. However, in this new model, the previously homogeneous group of treatments that provided process support turned out to be heterogeneous. Consequently, we performed the same procedure again by also combining this category with the moderator *age*. Since only two treatments utilized process support for learners between the ages of 9 to 13, this subcategory was also excluded from the analysis. Again, one treatment did not report information on the age of participants and therefore could not be included within this further segmentation. As presented in Table 9, a perfect model fit was reached. Treatments that conducted strategy instruction on young learners between the ages of 9 to 13 showed a very high effect of $\Delta_{Glass} = 1.21$, $p < .01$ on academic achievement. Strategy instruction on adult learners, however, did not have a significant effect, which explains the former heterogeneity of the entire group. Subdividing *process support* into two subgroups revealed a medium effect (Cohen, 1992) of $\Delta_{Glass} = .45$, $p < .05$ for adult learners and a very small effect of $\Delta_{Glass} = .08$, $p > .05$ for adolescent learners. However, due to the subdivision of groups, the formerly significant effect of treatments using process support disappeared. A box plot is presented in Figure 9.

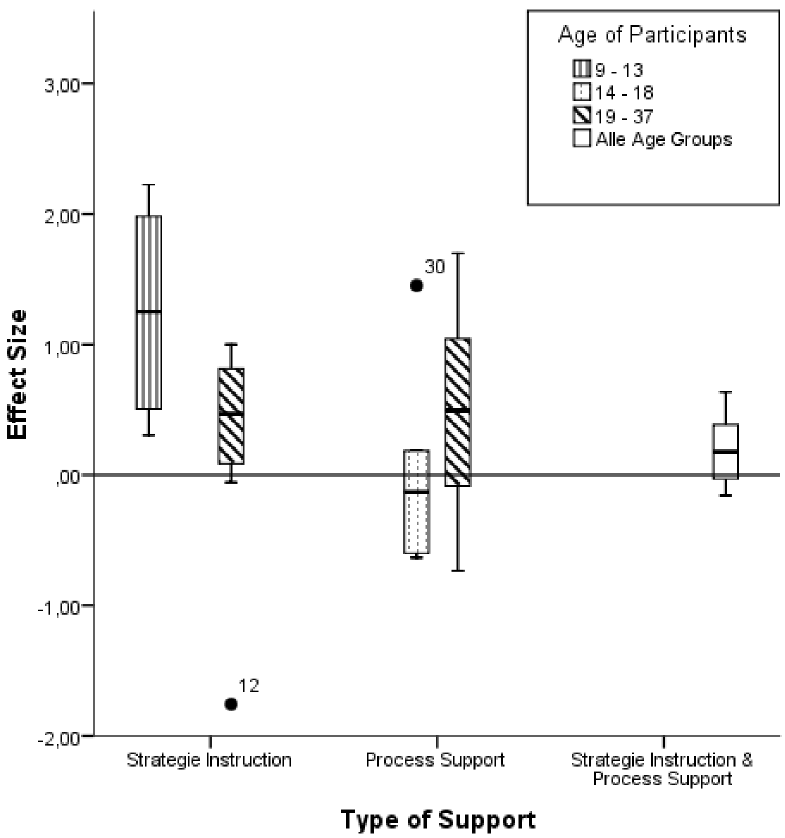


Figure 9. Box plot of the combined moderators type of support and age of participants. The mean effect size is unweighted.

5.4.3 Domain of Learning and Age

Applying *domain of learning* as a moderator in a categorical model revealed overall heterogeneity, as well as two heterogeneous groups. Treatments that took place within a mathematical or a language context did not constitute a homogeneous sample and therefore were combined with the moderator *age*. We did not include categories consisting of fewer than four treatments in an analysis of combined effects. Only one treatment in a mathematical environment was conducted for participants between the ages of 19 to 37, and was therefore excluded. This was also the case for one treatment focusing on 14 to 18 year-old learners, and two treatments focusing on 19 to 37 year-old learners in a language context. Also, two treatments had to be excluded because of a lack of reported information. Table 9 presents the results of the analysis. A perfect model fit was reached. Treatments conducted on 9 to 13 year-old learners in a mathematical context showed a very high effect on academic achievement of $\Delta_{Glass} = 1.09$, $p < .01$. In contrast, treatments aimed at fostering adolescent learners' SRL in mathematical environments were not effective, which explains heterogeneity within the entire group of treatments conducted in a mathematical context. Also, treatments conducted on 9 to 13 year-old learners in a language context significantly affected academic achievement. This was also the case for SRL treatments in a science context across all age groups. Figure 10 presents a box plot.

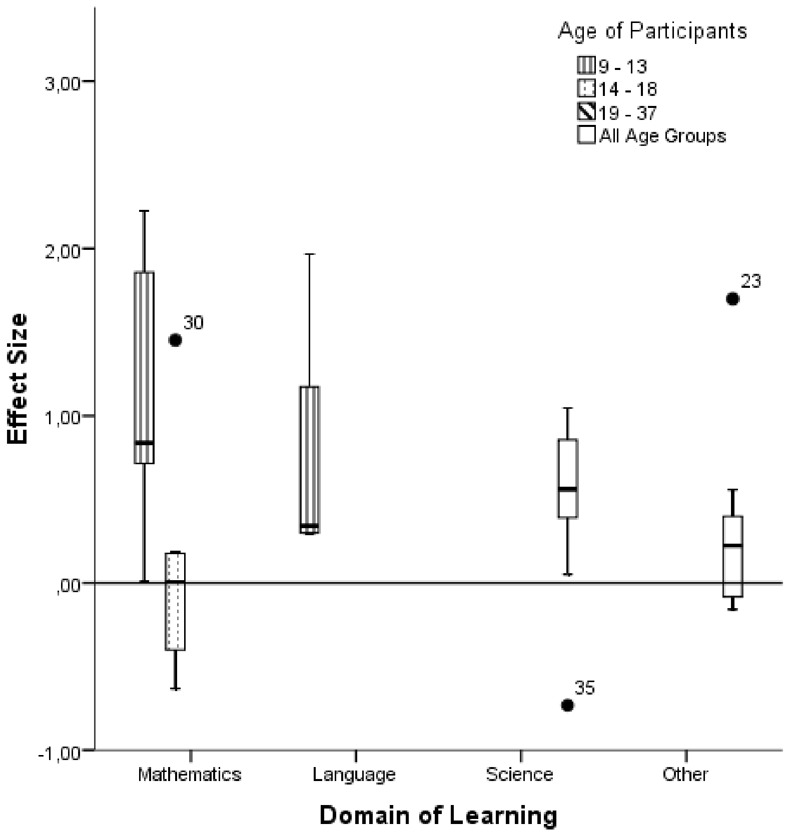


Figure 10. Box plot of the combined moderators domain of learning and age of participants. The mean effect size is unweighted.

6. Discussion

6.1 Effectiveness of SRL Interventions

In order to answer our first set of research questions regarding whether scientists have managed to create SRL treatments that affect performance, to quantify this effect, and at the same time, to evaluate the relevance of SRL for learning, we included 44 independent treatments from 38 papers in this meta-analysis. Based on 4047 learners, we found a significant weighted mean effect of Δ_{Glass} = .45, $p < .01$ for SRL treatments on academic achievement. Computation of the fail-safe *n* indicated that even if there existed 19 (50) additional SRL treatments with no effect at all, their inclusion in our analysis would still reveal an effect size of .31 (.21). Hence, the significance of the effect is beyond doubt.

However, as indicated by the distribution of the study effect sizes (Figure 3), as well as by the significance of the test for homogeneity, effect sizes within our sample did vary. Considering our broad approach toward the independent variable by defining few integration criteria, this heterogeneity is not a surprise. In contrast, it was our goal to account for a great deal of variability in treatments, and to dissolve differences within the sample by applying our proposed moderators. An analysis of the four outliers (treatments 12, 19, 27, & 28) in order to find explanations for the extreme effect sizes of those treatments did not reveal any pattern. Summing up this first analysis, as hypothesized, SRL interventions do have a positive impact on performance. Since synthesizing the effects of SRL interventions balances the effects of different conceptualizations of SRL treatments, the true effect of SRL on academic achievement is isolated. Following this line of argumentation, our results support the relevance of SRL for learning. Regardless of how learners are supported, the application of SRL provides a gain in performance. However, as a consequence of the heterogeneity within our sample of treatments, the effect of $\bar{\Delta}_{\text{Glass}} = .45$, $p < .01$ of SRL interventions on academic achievement is not to be taken as an estimate of the population parameter, but should serve rather as a descriptive result.

Hattie et al. (1996) reported an effect of $\bar{g}_{\text{Hedges}} = .57$ of study skills interventions on performance, which, in comparison to our result, is somewhat larger. This difference in effect might be due to a different scope of treatments. Focusing on study skills interventions that did not take place within a regular teaching context, Hattie et al. applied different integration criteria, as well as another set of search terms. Furthermore, Hattie et al. based their analyses on studies that had been published between 1982 and 1992, and therefore generated a different sample of studies. Ignoring dependencies between study effect sizes might also have enlarged their overall effect, as interventions with large effect sizes might have been very influential.

Comparing our result to the findings of Dignath, Buettner, et al. (2008), who reported an effect size of $\bar{g}_{\text{Hedges}} = .62$, reveals the same pattern. Again, the difference in effect sizes might be due to a different scope of treatments. When taking a look at the integration criteria of Dignath, Buettner, et al., a striking match between characteristics of selected treatments by Dignath, Buettner, et al. and the most effective categories of our moderators can be observed. As we found very high effects for peer-reviewed and quasi-experimental studies, as well as for treatments that applied strategy instruction, took place over 3 to 6 weeks, and focused on young learners between the ages of 9 to 13, a larger effect when including interventions for this group does not seem surprising. Hence, it seems like Dignath, Buettner, et al. chose a sample of SRL interventions that produces large effects on performance.

6.2 Toward a Framework of Fostering SRL

The relevance of evaluated and effective SRL treatments has been pointed out in this article. It has also been mentioned that, at the moment, a striking lack of theoretical guidance for creating those SRL interventions can be observed. In order to close this gap and, hence, to provide a path toward the development of a framework for the creation of effective SRL interventions, the second objective of this meta-analysis was put forth. We aimed to identify features of SRL interventions that have proven effective, as well as to point out properties of interventions that have escaped attention in past research. To address these matters, as a first step, we conducted several hypothesis-driven one-way moderator analyses. In a second step, we followed an exploratory approach aimed at resolving heterogeneous categories of moderators by combining effects. Pure interactions could not be calculated because of a restricted number and variety of studies within this rather young area of research.

A good model fit was reached for the moderator *review status*, and therefore allowed for the generalizing of results. In line with our hypothesis and the results of Hattie et al. (1996), peer-reviewed studies (Δ_{Glass} =.82, p < .01) did report larger effects than non-peer-reviewed studies (Δ_{Glass} =.23, p < .01). This difference in effectiveness underlines the adequacy of our strategy to put great effort into obtaining grey literature. Only focusing on peer-reviewed studies apparently would have overestimated the effect of SRL interventions on performance. Hence, the so called "file drawer problem" (Rosenthal, 1979) or publication bias (Rothstein, Sutton & Borenstein, 2005), once more, has been documented.

For the moderator *research design*, a perfect model fit was reached, which allowed for the generalizing of results. At a first glance, the superiority of studies that did not randomly assign learners to conditions seemed somewhat counterintuitive; most certainly, we must question whether this finding is an artifact of non-equivalent groups. However, since most quasi-experimental studies are conducted in real classrooms, other influences can be presumed to be present. Besides the Hawthorne-Effect (Roethlisberger & Dickson, 1939), which proposes that participants who receive an intervention will be more motivated than participants who follow their daily routine, it is very likely that teachers are responsible for this difference in effectiveness. Once teachers realize that a treatment is not as effective as expected, an increase in effort might occur in order to make up for poor treatments, and to provide their students with the required competences. This explanation is also underlined by the superiority of treatments delivered by teachers, compared to treatments carried out by researchers.

In a first analysis of the moderator *instance of delivery of intervention*, treatments delivered by humans were taken as a whole. In line with our hypothesis, treatments delivered by humans had the greatest effect on academic achievement (Δ_{Glass} =.76, $p < .01$). However, the very small group of four treatments that exclusively applied paper as the instance of delivery (Δ_{Glass} =.42, $p > .05$), was more effective than treatments delivered by computers (Δ_{Glass} =.11, $p > .05$). The validity of a result based on such a small group must be questioned, especially when considering the medium effect of treatments delivered by humans and paper (Δ_{Glass} =.37, $p < .05$). Since no model fit was reached, results can be interpreted only descriptively. In order to dissolve heterogeneity within the group of treatments that were delivered by humans, in a second analysis, we distinguished between treatments delivered by researchers and by teachers. Whereas treatments delivered by teachers were found to constitute a homogeneous group with an effect of Δ_{Glass} = .85, $p < .01$, treatments delivered by researchers were still assessed to be a heterogeneous subsample with a mean effect of Δ_{Glass} =.55, $p < .01$. Regarding the great variability of interventions conducted by researchers who apply treatments of a variety of shapes, heterogeneity is not surprising. In contrast, interventions conducted by teachers usually appear in combination with certain study features from the areas of *research design*, *type of support*, or *duration of intervention*. Hence, there is a good chance that homogeneity is due to confounding variables (Matt & Cook, 1994).

With respect to the effectiveness of treatments, we found an effect that was contrary to the findings of Hattie et al. (1996), Dignath, Buettner, et al. (2008) and Dignath and Buettner (2008), who reported that treatments delivered by researchers were more effective. First of all, as teachers can be expected to increase their efforts to make up for poor treatments in order to support their students, a superiority of teacher-delivered instructions does not seem unsubstantiated. However, to find reasons for these contradictory results is somewhat speculative. By excluding interventions that took place in regular teaching contexts (Hattie et al.) or by focusing on peer-reviewed studies that implemented strategy instruction within primary schools (Dignath, Buettner, et al.), the two previous meta-analyses followed a more narrow approach with regard to the independent variable. It is possible that the variety within our sample of treatments might have distorted the effect. Verifying this assumption by reducing our sample to the samples used in the previous meta-analyses, however, is not possible because there is no overlap with Hattie et al., and an overlap of only two studies with

Dignath, Buettner, et al. and Dignath and Buettner. Nonetheless, the heterogeneity within the group of treatments delivered by researchers is indicative of great variability, and allows for the assumption that the samples of Hattie et al., Dignath, Buettner, et al. and Dignath and Buettner are located somewhere in the higher region of effect sizes. With respect to future research, it has to be stated that computer-based interventions revealed astonishingly small effect sizes. More effort should be put into creating ways to make use of technology-enhanced learning. Furthermore, in order to investigate SRL more systematically, the number of studies in the category of paper-based interventions should be increased.

Without doubt, one of the major findings of this meta-analysis is the great influence of the *age of participants* on the effect of SRL treatments. Regardless of the nature of the treatment they are exposed to, young learners between the ages of 9 to 13 benefit the most from SRL interventions (Δ_{Glass} = .81, $p < .01$). As hypothesized, adult learners between the ages of 19 to 37 were found to profit moderately from SRL treatments (Δ_{Glass} = .33, $p < .05$). A nonsignificant small effect of interventions performed on adolescents between the ages of 14 to 18 once more underlines the resistance to treatments that is so prevalent for this age group. As a perfect model fit is reached by this moderator, these results may be generalized. Comparing our results to the findings of Hattie et al. (1996), a slightly different pattern can be observed, as they reported interventions for upper secondary students to be most effective in terms of performance; the secondary students were followed by primary, preprimary, lower secondary, university students, and adults.

Due to the large heterogeneous group of metacognitive and cognitive treatments, no model fit was reached for the moderator *SRL layer of intervention*. In contrast to our hypothesis and the results of Dignath, Buettner, et al. (2008), the group of metacognitive treatments showed the highest effect on performance (Δ_{Glass} = .55, $p < .05$). However, the cell size was very small. As we paid great attention to the correct handling of dependencies by only integrating the most complex treatment of a study, in this case, cell size does not represent occurrence in current research, but is rather an artifact of our approach. Hence, the application of another procedure would have resulted in a different distribution of cell sizes. At any rate, as metacognitive treatments constitute a homogeneous group, the true effect size is indicated. This is also the case for metacognitive and motivational treatments (Δ_{Glass} =.30, $p > .05$). A significant medium effect of Δ_{Glass} = .43, $p < .01$ was found for the group of metacognitive, cognitive, and motivational treatments. In order to investigate reasons for the great variability within metacognitive and cognitive treatments, as well as to reveal possible explanations for results that differed from Dignath, Buettner, et al. and Hattie et al. (1996), *SRL layer* was combined with other moderators. A perfect model fit allowing for generalizing results could be found when grouping metacognitive and cognitive treatments by *age*. All three subgroups turned out to be homogeneous. Whereas metacognitive and cognitive treatments focusing on adolescent and adult learners did not significantly affect academic achievement, a large effect size of $\Delta_{Glass} = 1.30$, $p < .01$ was found for young learners. Hence, our results underline the findings of Hattie et al., who suggested conducting treatments within the teaching of content, and further specify this recommendation to be valid for young learners.

For the moderator *type of support* no model fit was reached. In contrast to our hypothesis, interventions combining strategy instruction and process support had the smallest effect ($\Delta_{Glass} = .18$, $p > .05$), whereas treatments that used process support had a medium effect ($\Delta_{Glass} = .33$, $p < .05$), and treatments providing strategy instructions had a large effect ($\Delta_{Glass} = .73$, $p < .01$) on performance. These results seem somewhat counterintuitive, and might be due to confounding variables. When combining *type of support* with *age*, a perfect model fit, allowing for generalization was reached. Strategy instruction on young learners did have a large effect on performance ($\Delta_{Glass} = 1.21$, $p < .01$), whereas adults benefited less from this type of support. The pattern for process support seems to point in the opposite direction. A medium effect was found for process support on adult learners ($\Delta_{Glass} = .45$, $p < .05$), and a small effect for adolescent learners ($\Delta_{Glass} = .08$, $p > .05$). Referring to mediation and production deficiencies, these results imply that young learners do suffer from mediation deficiencies, and hence benefit most from strategy instruction, whereas process support helps older learners to apply strategies they are already equipped with. Considering the relevance of adaptively meeting the needs of learners, it is striking that in our sample of SRL interventions, only one treatment adaptively accounted for individual requirements of learners by generating prompts on the basis of a metacognitive test (Schwonke, 2005).

Applying *domain of learning* as a moderator did not reveal a model fit. Comparing the descriptive means of our sample to the results of Dignath and Buettner (2008), however, revealed the same pattern. Treatments within a mathematical context ($\Delta_{Glass} = .54$, $p < .01$) were more effective than treatments within science ($\Delta_{Glass} = .49$, $p < .01$), language ($\Delta_{Glass} = .38$, $p < .05$), and other ($\Delta_{Glass} = .28$, , $p > .05$) contexts. In a further analysis, combining heterogeneous groups with *age* revealed a perfect model fit and therefore allowed for generalization. Fostering SRL within mathematical ($\Delta_{Glass} = 1.09$, $p < .01$) and language ($\Delta_{Glass} = .72$, $p < .01$) contexts was very effective for young learners. These results provide a partial match with the findings of Dignath and Buettner who found a larger effect of treatments on primary school students than on secondary school students within a mathematical context. However, since our sample did not allow for investigating the effect of treatments on adolescents within a language context, Dignath and Buettner's finding of the superiority of treatments on secondary school students in comparison to treatments on primary school students within a language context could not be replicated.

For *SRL level of intervention*, *duration of intervention*, and *measure of academic achievement*, no model fit could be reached. Hence, the results should not be generalized but indicate the main characteristics of our sample. Interventions that focused on the micro level had a large effect of $\Delta_{Glass} = .56$, $p < .01$ on academic achievement, but were a heterogeneous subsample. Interventions that focused on the mid level and on the micro and mid level constituted homogeneous groups, but had smaller effects on performance. In order to investigate differences in effectiveness between interventions that are focused on both levels, more research on the mid level is needed. However, since all interventions that conduct treatments on the mid level measure performance within school or university courses, effects are diluted by learners' abilities on the micro level. In contrast, micro-level interventions also use tests requiring minimum transfer and therefore eliminate the influence of the mid level. Even though mid-level interventions seem to be determined to produce smaller effects on academic achievement than micro-level interventions, their importance should not be diminished.

To adequately regard *duration of an intervention*, the moderator was broken up into hours of treatment and length of intervention. All treatments of more than 1 hour constituted homogeneous groups and resulted in large effect sizes. Very short treatments of less than 1 hour in sum did not have an effect on performance. However, the group was heterogeneous, which implies a great variation in effectiveness. With regard to the length of an intervention, it seems like very short interventions of 1 day do not provide enough time for learners to actually benefit from the treatment ($\Delta_{Glass} = .24$, $p > .05$). Although interventions that are carried out over 3 to 6 weeks were highly effective ($\Delta_{Glass} = .78$, $p < .01$), this effect was diminished, possibly by a decrease in learners' motivation, when interventions were conducted over 2 to 7 months ($\Delta_{Glass} = .45$, $p < .01$). In sum, we could replicate neither Dignath and Buettner's (2008) finding of a linear trend between hours of intervention and effect, nor Hattie's et al. (1996) negative curvilinear trend between length of intervention and effect. However, due to different objectives of short and long interventions in terms of altering states or traits, and varying gaps between treatment and assessment, these contradictory results do not seem surprising.

Following SRL theory, we aimed to categorize measures of *academic achievement* in terms of the degree of freedom learners were confronted with during assessment. However, since general levels of complexity could not be applied due to the different nature of tasks, categories were iteratively developed based on the sample. Considering the environment of the performance measure allows for a comparison between our results and the findings of Dignath, Buettner, et al. (2008), who reported large effects for mathematics performance, followed by other performances and reading/writing performance. As we found medium effects for grades and undefined measures ($\Delta_{Glass} = .44$, $p < .01$), as well as measures of problem solving within mathematics ($\Delta_{Glass} = .37$, $p > .05$) and writing quality ($\Delta_{Glass} = .40$, $p > .05$), our results do not support these findings. Instead, according to our hypothesis that SRL plays a more important role for complex tasks than for tasks that confront learners with a small degree of freedom, comprehension measures ($\Delta_{Glass} = .92$, $p < .01$) were affected more by SRL interventions than knowledge measures ($\Delta_{Glass} = .18$, $p > .05$).

6.3 Implications for Practice

It is striking that for one-way moderator analyses, a model fit could only be reached for two moderators, which are concerned with the specific design of SRL treatments. Besides *age of participants* and *instance of delivery*, the formal and methodological moderators *review status* and *research design* are variables that break up studies into homogeneous groups. So, what do we learn from this meta-analysis? At a first glance, it seems like researchers are free to follow their noses when creating SRL interventions as long as they focus on young learners between the ages of 9 to 13. Also, very large effects can be reached if treatments are carried out by teachers. However, this meta-analysis revealed several descriptive results, which indicate large effects for treatments that deliver support on the metacognitive layer, focus on the micro level, equip learners with strategies, are conducted over 3 to 6 weeks, are embedded in a mathematical environment, and assess comprehension. It is the task of those conducting primary research on SRL to systematically conduct more SRL interventions in order to be able to judge the relevance of these results (Cooper & Hedges, 1994). In particular, more research is needed on categories of SRL interventions that currently lack effectiveness. For example, considering the potential of technology-enhanced learning and the importance of supporting strategy application during learning, computer-based learning environments and process support produce surprisingly small effect sizes. Also, regarding our analysis of the occurrence of study features in current SRL research, interventions on the mid level of SRL, treatments to foster writing quality, and paper-based instructions have lacked attention. However, small cell sizes of interventions on the metacognitive as well as the metacognitive and motivational layer can be considered artefacts of our approach. Nevertheless, regarding the restricted number of studies within this rather young area of research, the performing of more and various SRL interventions seems appropriate.

Referring to the analyses of interactions between moderators, a model fit could be reached for *SRL layer of intervention*, *type of support*, and *domain of learning* in combination with *age of participants*. Hence, on the level of two-way moderator analyses, the design of a treatment gains importance when considering the age of participants. Very large effects of SRL treatments can be reached when providing young learners with support on the metacognitive and cognitive layers, when conducting strategy instructions, and when selecting a mathematical learning environment. Accordingly, as age obviously accounts for a great deal of variance, suggestions about the design of SRL interventions should always consider the age of the group of interest.

7. Conclusion

Summing up 17 years of SRL research, in this meta-analysis, a broad approach was followed with regard to the independent variable. By only excluding interventions for participants with learning disabilities, learning difficulties or special needs, as well as treatments that exclusively focused on the cognitive and/or motivational layers, a great variety of SRL treatments were accounted for. With regard to the dependent variable, however, measures were narrowed down to academic performance. Paying great attention to avoid mechanisms of dependency, only one effect size was calculated per independent treatment. The drawback of this procedure was that information provided by primary research was not totally utilized. Hence, some study features were underrepresented in our sample, as was the case for metacognitive treatments, for example. However, in order to avoid applying the same piece of information more than once within the same analysis, we accept this loss of information.

Investigating our first set of research questions, we found that scientists have managed to create effective SRL treatments. As the integration of many SRL interventions eliminates the influence of specific treatment designs and therefore isolates the influence of SRL on academic achievement, this analysis also bolsters the importance of SRL for learning. In order to provide a path toward the development of a framework for the creation of effective SRL interventions, moderator analyses were performed. When only considering one moderator, two indicators for the specific design of SRL treatments could be found. Treatments conducted with young learners and interventions delivered by teachers were highly effective. In addition, peer review status and research design were influential variables. Furthermore, interventions on the mid level of SRL, treatments to foster writing quality, and paper-based instructions were identified as study features that have lacked attention in current SRL research. When analyzing combined effects, treatments that focused on the metacognitive and cognitive layer, strategy instructions, and treatments conducted within mathematical learning environments turned out to be most effective for young learners. Hence, as age of participants accounted for a great deal of variance, a framework for the creation of SRL interventions should always consider the group of interest. Another achievement of this meta-analysis was the continuation of the development of variables that could be used to classify SRL interventions in order to systematize much-needed future primary research.

Acknowledgments

The authors would like to thank Karim Hirsch, who bravely worked his way through the studies. The research was part of the doctoral dissertation of Bastian F. Benz. An earlier version of this article was presented at the 2008 Congress of Psychology in Berlin, Germany.

References

References marked with an asterisk indicate studies included in the meta-analysis.

Alexander, P. A. (1997). Mapping the multidimensional nature of domain learning: The interplay of cognitive, motivational and strategic forces. In M. L. Maehr & P. R. Pintrich (Eds.), Advances in Motivation and Achievement (Vol. 10, pp. 213-250). Greenwich, CT: JAI Press.

Anderson, J. R. (1993). *Rules of the mind.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

*Azevedo, R., & Cromley, J. G. (2004). Does training on self-regulated learning facilitate students' learning with hypermedia? *Journal of Educational Psychology, 96(3)*, 523-535. doi:10.1037/0022-0663.96.3.523

*Azevedo, R., Cromley, J. G., & Seibert, D. (2004). Does adaptive scaffolding facilitate students' ability to regulate their learning with hypermedia? *Contemporary Educational Psychology, 29*(3), 344-370. doi:10.1016/j.cedpsych.2003.09.002

*Azevedo, R., Cromley, J. G., Thomas, L., Seibert, D., & Tron, M. (2003). *Online process scaffolding and students' self-regulated learning with hypermedia*. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, IL. Retrieved from ERIC database (ED478070)

Bangert-Drowns, R. L. (1986). A review of developments in meta-analytic method. *Psychological Bulletin, 99(3)*, 388-399. doi: 10.1037/0033-2909.99.3.388

Beaton, A. E., Mullis, I. V. S., Martin, M. O., Gonzalez, E. J., Kelly, D. L., & Smith, T. A. (1996). *Mathematics achievement in the middle school years: IES's third international mathematics and science study*. Chestnut Hill, MA: Boston College.

*Beck, E., Guldimann, T., & Zutavern, M. (1994). Eigenstaendiges Lernen verstehen und foerdern. [The comprehension and promotion of self-directed learning]. In K. Reusser & M. Reusser-Weyeneth (Eds.), *Verstehen: Psychologischer Prozess und didaktische Aufgabe (S. 207-225)*. Switzerland: Huber.

Becker, B.J. (1988). Synthesizing standardized mean-change measures. *British Journal of Mathematical and Statistical Psychology, 41(2)*, 257-278.

Benz, B. F., Polushkina, S., Schmitz, B., & Bruder, R. (2007). Developing Learning Software for the Self-Regulated Learning of Mathematics. In M. B. Nunes. & M. McPherson (Ed.). *IADIS International Conference e-Learning (pp. 200-204)*. Portugal: IADIS Press.

*Berthold, K. (2006). *Lernen aus Loesungsbeispielen: Verstaendnis foerdern mit multiplen Repraesentationen, einer Integrationshilfe und Selbsterklaerungsprompts* [Learning from worked-out examples: Multiple representations, an integration help, and self-explanation prompts all foster understanding]. Unpublished doctoral dissertation, Universitaet Freiburg i. Br., Germany.

Boekaerts, M. (1999). Self-regulated learning: Where we are today. *International Journal of Educational Research*, 31(6), 445-457. doi:10.1016/S0883-0355(99)00014-2

*Boone, G. L. (1999). Cognitive self-instruction to foster self-regulation in regular education students. (study skills, metacognition, academic achievement). *Dissertation Abstracts International, 60*, 5-A. (UMI No. Aeh9929020)

*Bruder, S. (2006). *Die Foerderung von Selbstregulation bei Kindern unter Einbeziehung ihrer Eltern* [Promotion of children's self-regulation integrating their parents]. Germany: Logos-Verlag.

Campbell, D. T., & Stanley, J. C. (1966). *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.

*Campillo, M. (2006). Acquisition and transfer of a writing revision strategy: A self-regulatory analysis. *Dissertation Abstracts International*, 67, 3-A. (UMI No. Aai3213136)

Cohen, J. (1968). Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin, 70(4)*, 213-220. doi: 10.1037/h0026256

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112(1)*, 155-159. doi: 10.1037/0033-2909.112.1.155

Colorado, J.T. (2006). The relationship of self-regulated learning and academic performance in an online course environment (Doctoral dissertation, University of Kansas, 2006). *Dissertation Abstracts International, 67*, 527.

Cooper, H. (1982). Scientific guidelines for conducting integrative research reviews. *Review of Educational Research, 52(2)*, 291-302. doi: 10.2307/1170314

Cooper, H. (1985). Literature searching strategies of integrative research reviewers. *American Psychologist, 40(11)*, 1267-1269. doi: 10.1037/0003-066X.40.11.1267

Cooper, H., & Hedges, L. V. (1994). Potentials and limitations of research synthesis. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis (pp. 521-529).* New York: Russell Sage Foundation.

*Cuevas, H. M. (2005). Transforming learning into a constructive cognitive and metacognitive activity: Use of a guided learner-generated instructional strategy within computer-based training. *Dissertation Abstracts International*, 66, 2-B. (UMI No. Aai3163597)

Deci, E., & Ryan, R. (Eds.). (2002). *Handbook of self-determination research*. Rochester, NY: University of Rochester Press.

Demetriou, A. (2000).Organization and development of self-understanding and self-regulation: Toward a general theory. In M. Boekaerts, P. R. Pintrich & M. Zeidner (Eds.), *Handbook of self-regulation (pp. 209-251).* San Diego, CA: Academic Press.

Dignath, C., & Buettner, G. (2008). Components of fostering self-regulated learning among students. A meta-analysis on intervention studies at primary and secondary school level. *Metacognition and Learning, 3(3)*, 231-264. doi: 10.1007/s11409-008-9029-x

Dignath, C., Buettner, G., & Langfeldt, H.-P. (2008). How can primary school students learn self-regulated learning strategies most effectively? A meta-analysis on self-regulation training programmes. *Educational Research Review, 3(2)*, 101-129. doi: 10.1016/j.edurev.2008.02.003

*Duke, B. L. (2004). The influence of using cognitive strategy instruction through writing rubrics on high school students' writing self-efficacy, achievement goal orientation, perceptions of classroom goal structures, self-regulation, and writing achievement. *Dissertation Abstracts International, 64*, 10-A. (UMI No. Aai3109071)

Durlak J. A. (2000). How to evaluate a meta-analysis. In D. Drotar (Ed.), *Handbook of research in pediatric and clinical child psychology (pp. 395-407)*. New York: Kluwer Academic/Plenum Publishers.

Durlak J. A., & Lipsey, M. W. (1991). A practitioner's guide to meta-analysis. *American Journal of Community Psychology, 19(3)*, 291-332. doi: 10.1007/BF00938026

*Eide, B. J. (1999). Integrating learning strategy instruction in an intermediate accounting course: Changes in accounting majors' learning strategy use and course performance. *Dissertation Abstracts International, 59*, 7-A. (Umi No. Aam9839195)

Endnote (Version 6) [Computer software]. New York: Thomson Reuters.

Flavell, J. H. (1970). Developmental studies of mediated memory. In H. W. Reese & L. P. Lipsitt (Eds.), *Advances in child development and behaviour: Vol. 5 (pp. 181-211).* New York: Academic Press.

*Fuchs, L. S., Fuchs, D., Prentice, K., Burch, M., Hamlett, C. L., Owen, R., & Schroeter, K. (2003). Enhancing third-grade student' mathematical problem solving with self-regulated learning strategies. *Journal of Educational Psychology, 95(2)*, 306-315. doi: 10.1037/0022-0663.95.2.306

*Gargallo-López, B. (2001). Learning strategies. A training programme. *European Journal of Psychology of Education, 16(1)*, 49-65. doi: 10.1007/BF03172994

Greenhouse, J. B., & Iyengar, S. (1994). Sensitivity analysis and diagnostics, In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis (pp. 383-398)*. New York: Russell Sage Foundation.

*Glaser, C. (2005). *Foerderung der Schreibkompetenz bei Grundschuelern: Effekte einer integrierten Vermittlung kognitiver Schreibstrategien und selbstregulatorischer Ferti*gkeiten [Promotion of writing skills in elementary school students: Effects of the integrated teaching of cognitive writing strategies and self-regulation skills]. Unpublished doctoral dissertation, Universitaet Potsdam, Germany.

Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher, 5(10)*, 3-8. doi: 10.3102/0013189X005010003

Glass, G. V., McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research.* Beverly Hills, CA: Sage.

Gleser, L. J., & Olkin, I. (1994). Stochastically dependent effect sizes. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis (pp. 339-355)*. New York: Russell Sage Foundation.

*Guan, J. S. K. (1995). The acquisition of learning strategies to improve academic performance. *Dissertation Abstracts International*, 56, 2-A. (UMI No. Aam9520369)

*Guertler, T. (2003). *Trainingsprogramm zur Foerderung selbstregulativer Kompetenz in Kombination mit Problemloesestrategien - PROSEKKO* [Training program for the improvement of problem solving and self-regulating competencies - PROSEKKO]. Germany: Peter Lang.

Hattie, J., Biggs, J., & Purdie, N. (1996). Effects of learning skills interventions on student learning: A meta-analysis. *Review of Educational Research, 66(2)*, 99-136. doi: 10.2307/1170605

Hedges, L. V. (1981). Distribution theory for Glass´s estimator of effect size and related estimators. *Journal of Educational Statistics, 6(2)*, 107-128. doi: 10.3102/10769986006002107

Hedges, L. V. (1982). Estimation of effect size from a series of independent experiments, *Psychological Bulletin, 92(2)*, 490-499. doi: 10.1037/0033-2909.92.2.490

Hedges, L. V. (1990). Directions for future methodology. In K. W. Wachter & M. L. Straf (Eds.), *The future of meta-analysis (pp. 11-26)*. New York: Russel Sage Foundation.

Hedges, L. V. (1994). Fixed effects models. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis (pp. 285-299)*. New York: Russell Sage Foundation.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. New York: Academic Press.

Hedges, L. V., & Vevea, J. L. (1998). Fixed- and random-effects models in meta-analysis. *Psychological Methods, 3(4)*, 486-504. doi: 10.1037/1082-989X.3.4.486

*Heo, H. (1999). The effect of self-regulated learning strategies on learner achievement and perceptions on personal learning responsibility. *Dissertation Abstracts International*, 59, 10-A. (UMI No. Aam9911444)

Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis (2nd ed.)*. Newbury Park, CA: Sage.

Kauffman, D. F. (2004). Self-regulated learning in web-based environments: Instructional tools designed to facilitate cognitive strategy use, metacognitive processing, and motivational beliefs. *Journal of Educational Computing Research, 30(1-2)*, 139-161. doi: 10.2190/AX2D-Y9VM-V7PX-0TAD

*Kauffman, D. F. (2002). Self-regulated learning in web-based environments: Instructional tools designed to facilitate cognitive strategy use, metacognitive processing, and motivational beliefs. *Dissertation Abstracts International*, 62, 8-A. (UMI No. Aai3022640)

*Keith, N. (2005). *Self-regulatory processes in error management training*. Unpublished doctoral dissertation, Justus-Liebig-Universitaet Giessen, Germany.

*Kramarski, B., & Gutman, M. (2006). How can self-regulated learning be supported in mathematical e-learning environments? *Journal of Computer Assisted Learning, 22(1)*, 24-33. doi: 10.1111/j.1365-2729.2006.00157.x

*Kramarski, B., & Hirsch, C. (2003). Using computer algebra systems in mathematical classrooms. *Journal of Computer Assisted Learning, 19(1)*, 35-45. doi: 10.1046/j.0266-4909.2003.00004.x

*Kramarski, B., & Mizrachi, N. (2006a). Online discussion and self-regulated learning: Effects of instructional methods on mathematical literacy. *Journal of Educational Research, 99(4)*, 218-230. doi: 10.3200/JOER.99.4.218-231

*Kramarski, B., & Mizrachi, N. (2006b). Online interactions in a mathematical classroom. *Educational Media International, 43(1)*, 43-50. doi: 10.1080/09523980500490778

*Kramarski, B., & Zeichner, O. (2001). Using technology to enhance mathematical reasoning: effects of feedback and self-regulation learning. *Educational Media International, 38(2-3)*, 77-82. doi: 10.1080/09523980110041458

*Lan, W. Y. (1996). The effects of self-monitoring on students' course performance, use of learning strategies, attitude, self-judgment ability, and knowledge representation. *Journal of Experimental Education, 64(2)*, 101-115.

*Leutner, D., Barthel, A., & Schreiber, B. (2001). Studierende koennen lernen, sich selbst zum Lernen zu motivieren: Ein Trainingsexperiment [Students can learn to motivate themselves for learning - A training experiment]. *Zeitschrift für Paedagogische Psychologie, 15(3-4)*, 155-167.

*Leutner, D., & Leopold, C. (2003). Selbstreguliertes Lernen als Selbstregulation von Lernstrategien - Ein Trainingsexperiment mit Berufstaetigen zum Lernen aus Sachtexten [Self-regulated learning as self-regulation of learning strategies - A training experiment on learning from texts]. *Unterrichtswissenschaft, 31(1)*, 38-56.

Light, R. J., Singer, J. D., & Willet, J. B. (1994). The visual presentation and interpretation of meta-analyses. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis (pp. 439-453)*. New York: Russell Sage Foundation.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

Mansfield, R. S., & Busse, T. V. (1977). Meta-analysis of research: A rejoinder to Glass. *Educational Research, 6(9)*, 3. doi: 10.3102/0013189X006009003

*Masui, C., & De Corte, E. (2005). Learning to reflect and to attribute constructively as basic components of self-regulated learning. *British Journal of Educational Psychology, 75(3)*, 351-372. doi: 10.1348/000709905X25030

Matt, G. E., & Cook, T. D. (1994). Threats to validity of research syntheses. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis (pp. 503-520)*. New York: Russell Sage Foundation.

*McGovern, A. R. (2005). The influence of mental simulation in an expressive writing context on academic performance and the moderating influence of self-efficacy, optimism, and gender. *Dissertation Abstracts International*, 66, 1-B. (UMI No. Aai3162697)

*Mosley, J. D. (2006). How does the implementation of self-regulated learning strategies affect the academic achievement of eighth-grade algebra students. *Dissertations Abstracts International, 67*, 1-A. (UMI No. Aai3206366)

Organisation for Economic Co-operation and Development. (2001). *Knowledge and skills for life. First results from PISA 2000*. France: Author.

Orwin, R. G. (1994). Evaluating coding decisions. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis (pp. 139-162)*. New York: Russell Sage Foundation.

Overton, R. C. (1998). A comparison of fixed-effects and mixed (random-effects) models for meta-analysis tests of moderator variable effects. *Psychological Methods, 3(4)*, 354-379. doi: 10.1037/1082-989X.3.3.354

*Parcel, W. J. (2005). Teaching self-management of writing strategies to promote transfer. *Dissertation Abstracts International, 61*, 12-A. (UMI No. Aai3169435)

*Perels, F., Guertler, T., & Schmitz, B. (2005). Training of self-regulatory and problem-solving competence. *Learning and Instruction, 15(2)*, 123-139. doi: 10.1016/j.learninstruc.2005.04.010

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In P. R. Pintrich, M. Boekaerts & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 452-502). San Diego, CA: Academic Press.

Pintrich, P. R., & DeGroot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, 82(1), 33-40. doi: 10.1037/0022-0663.82.1.33

Raudenbush, S. W. (1994). Random effects models. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis (pp. 301-321)*. New York: Russell Sage Foundation.

Reed, J. G., & Baxter, P. M. (1994). Using reference databases, In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis (pp. 57-70)*. New York: Russell Sage Foundation.

Reese, H. W. (1962). Verbal mediation as a function of age level. *Psychological Bulletin*, 59(6), 502-509. doi: 10.1037/h0040739

Roethlisberger, F., & Dickson, W. (1939). *Management and the worker*. Cambridge, MA: Harvard University Press.

Rosenthal, R. (1979). The 'file drawer problem' and tolerance for null results. *Psychological Bulletin, 86(3),* 638–641. doi: 10.1037/h0040739

Rosenthal, R. (1991). *Meta-analytic procedures for social research*. Newbury Park, CA: Sage.

Rosenthal, R. (1994). Parametric measures of effect size. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis (pp. 231-244)*. New York: Russell Sage Foundation.

Rothstein, H. R., Sutton A. J., & Borenstein, M. (2005). *Publication Bias in Meta-Analysis*. West Sussex: Wiley.

Schmitz, B., & Wiese, B.S. (2006). New perspectives for the evaluation of training sessions in self-regulated learning: Time-series analyses of diary data. *Contemporary Educational Psychology, 31(1)*, 64-96. doi:10.1016/j.cedpsych.205.02.002

*Schwartz, L. S. (1996). The effects of a goal-setting instructional intervention on achievement, goal-setting habits, and self-efficacy for self-regulated learning. *Dissertation Abstracts International, 57*, 3A. (UMI No. Aam9623120)

Shadish, W. R., & Haddock, C. K. (1994). Combining estimates of effect size. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis (pp. 261-281)*. New York: Russell Sage Foundation.

*Schwonke, R. (2005). *A computer-supported writing-to-learn approach. The composition of individual and collaborative learning journals as a learning method in higher education and its support by a computer-based writing environment.* Unpublished doctoral dissertation, Universitaet Freiburg i. Br., Germany.

Smith, M. L. (1980). Publication bias and meta-analysis. *Evaluation and Education, 4(1)*, 22-24.

*Souvignier, E., & Mokhlesgerami, J. (2006). Using self-regulation as a framework for implementing strategy instruction to foster reading comprehension. *Learning and Instruction, 16(1)*, 57-71. doi: 10.1016/j.learninstruc.2005.12.006

Stock, W. A. (2001). Systematic coding for research synthesis. In H. Cooper & L. V. Hedges (Eds.), The *handbook of research synthesis (pp. 521-529)*. New York: Russell Sage Foundation.

*Trenk-Hinterberger, I. (2006).*Lesekompetenz im Schulunterricht. Evaluation und Optimierung eines strategieorientierten Programms in der fuenften Jahrgangsstufe* [Reading competence in school teaching: Evaluation and optimization of a strategy-oriented training program for fifth graders]. Germany: Kovac.

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. doi: 10.1007/s11409-006-6893-0

*Walser, T. M. (2001). Teaching self-management of writing strategies to promote transfer. *Dissertation Abstracts International, 61*, 12-A. (UMI No. Aai9999118)

White, H. D. (1994). Scientific communication and literature retrieval, In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis (pp. 41-55)*. New York: Russell Sage Foundation.

Winne, P., & Hadwin, A. (2008). The weave of motivation and self-regulated learning. In D. Schunk & B. J. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and applications* (pp. 297–314). New York: Taylor & Francis.

Woodworth, G. (1994). Managing meta-analytic databases. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis (pp. 177-189)*. New York: Russell Sage Foundation.

*Xiao, L. (2006). How goal orientations, perceived competence, and strategy training affect college students' use of self-regulated learning strategies and achievement in learning foreign languages. *Dissertation Abstracts International, 67*, 4-A. (UMI No. Aai3216561)

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich & M. Zeidner (Eds.), *Handbook of self-regulation (pp. 13-39).* San Diego, CA: Academic Press.

Zimmerman, B. J., & Kitsantas, A. (1997). Developmental phases in self-regulation: Shifting from process goals to outcome goals. *Journal of Educational Psychology, 89(1)*, 29-36. doi: 10.1037/0022-0663.89.1.29

Zimmerman, B. J., & Schunk, D. J. (Eds.). (2001). *Self-regulated learning and academic achievement: Theoretical perspectives*. New York: Lawrence Erlbaum Associates.

Study 2: Benz, B. F., Scholl, P., Boehnstedt, D., Schmitz, B., Rensing, C., & Steinmetz, R. (2010). *Improving the Quality of E-Learning. Scaffolding Self-Regulated Learning on the World Wide Web.* Manuscript submitted for publication.

Abstract

With the goal of enhancing the quality of the learning process and the learning outcome of web-based learning, the authors optimized the World Wide Web (WWW) as a learning environment by integrating scaffolds, which offer functions that can be used to complete a learning task on the WWW and simultaneously provide metacognitive assistance, into the Firefox web browser. To evaluate the effectiveness of the scaffolds in this study, undergraduate students (N = 64) were randomly assigned to four conditions to learn about *Classical Antiquity* on Wikipedia for 45 minutes. Experimental groups were either (a) free to apply the scaffolds of their own accord or (b) received additional invasive prompts directing them to engage in metacognitive processes. Two control groups did not receive scaffolding. The quality of the learning process was assessed by collecting self-reports and automatically generated log files. The quality of the learning outcome was assessed by gain in factual knowledge. For experimental conditions, additionally, video analyses of the learning period were conducted, and the quality of the created goal-resource structure was assessed. Learners who received scaffolding during web-based learning were more involved in self-regulated learning and experienced more positive emotions, but could not gather more factual knowledge. Learners who could apply the scaffolds of their own accord deployed more self-monitoring processes and created a goal-resource structure of a higher quality, but engaged in fewer reflection processes than those who received additional prompts. All in all, the study indicates that optimizing the WWW as a learning environment by providing learners with scaffolds, which embody functionality and metacognitive support, is a powerful concept for enhancing the quality of web-based learning.

Keywords: self-regulated learning, metacognition, scaffolding, computer assisted instruction, hypermedia, achievement

1. Introduction

Nowadays the World Wide Web (WWW) is used as a resource for learning in various settings (United Nations [UN], 2008). However, as it is a nonlinear and unstructured environment (Jonassen, 1996) that provides an enormous degree of freedom, the quality of web-based learning very much relies on the skills and strategies of the learners. It was the goal of our approach to support learners to overcome the obstacles they are confronted with during learning on the WWW and thereby to enhance the quality of the learning process and the learning outcome of web-based learning.

On the basis of research on self-regulated learning (SRL), it was assumed that the deployment of SRL processes enhances the quality of learning on the WWW. From a social cognitive theoretical perspective, SRL is defined as learners' self-generated thoughts, feelings, and actions that are systematically oriented toward the attainment of their learning goals (Zimmerman & Schunk, 2001). Based on current models of SRL (Alexander, 1997; Boekaerts, 1999; Pintrich, 2000; Schmitz & Wiese, 2006; Winne & Hadwin, 2008; Zimmerman, 2000), we derived six metacognitive processes for which deployment is assumed to enhance the quality of learning on the WWW.

With the goal of supporting the deployment of those metacognitive processes during web-based learning, we pursued an indirect approach of providing assistance (Friedrich & Mandl, 1992). We optimized the WWW as a learning environment by creating an extension to the Mozilla Firefox web browser called *E-Learning knoWledge Management System* (ELWMS). Following a new concept of support, ELWMS provides scaffolds (Palincsar, 1998), which, on the one hand offer functions that can be used to conduct a web-based learning task and on the other hand simultaneously induce metacognitive processes. In the standard version of ELWMS, learners are able to decide if, how, and when to apply those scaffolds during learning on the WWW. However, considering research that suggests that learners might not be able to make use of scaffolds of their own accord (Aleven, Stahl, Schworm, Fischer, & Wallace, 2003; Oliver & Hannafin, 2000), we created a second version of ELWMS that provides additional invasive and directive prompts.

It was the aim of the current study to evaluate the impact of the indirect scaffolding as it is realized in our ELWMS software on the quality of the learning process and the learning outcome of web-based learning. We compared two groups that worked with the two versions of our ELWMS software to two control groups that did not receive any scaffolding, a weak control group that just worked with Firefox and a strong control group that was additionally allowed to used pen and paper. We further investigated whether learners who received additional prompts would reach a higher quality of web-based learning than those who had the freedom to apply the scaffolds of their own accord.

In this study, we followed a multi-method approach to evaluate the quality of web-based learning. Current research suggests that *offline* measures, which are assessed with a temporal distance to the learning process, reveal different results than *online* measures, which are collected during the ongoing learning process (Veenman, 2007). Taking these findings into account, in addition to the assessment of retrospective self-reports on the learning process, we automatically generated log files and conducted analyses of screen recordings. This multi-method approach allowed for analyzing the validity of our learning measures.

1.1 The World Wide Web as a Learning Environment

In the past 2 decades the internet has gained great importance in modern life. Nowadays, in developed countries, 50-82% of individuals between the age of 15 and 74 use the internet on a daily basis and an additional 14-41% use the internet at least once a week (UN, 2008). The World Wide Web is a hypermedia system that is accessible through the internet. It provides an immense and permanently growing amount of information about all kinds of topics represented as text, graphics, animation, audio, and video in a nonlinear fashion (Jacobson & Archodidou, 2000; Jonassen, 1996). In recent years, a paradigm change from web 1.0 to web 2.0 technologies has taken place. Nowadays, internet users are not in the position of passive consumers of information anymore, but are in the role of active authors who can easily create and publish content on the WWW. The web-based encyclopedia, Wikipedia, which is collaboratively created by users all over the world, is an example of this new web 2.0 technology.

Individuals use the World Wide Web as a resource (Rakes, 1996) for obtaining information and for utilizing education or learning activities in vocational, educational, and private settings (UN, 2008). However, now more than ever with the introduction of web 2.0 technologies, information on the WWW is provided by all sorts of entities and is commonly not presented in a manner that benefits learning. Accordingly, when employing the World Wide Web as a learning environment, is the responsibility of the users to master the freedom they are confronted with and to profit from the environment. They have to decide what to learn, how much to learn, how to learn, and how much time to spend learning. Further, they have to navigate within the environment, find relevant resources, determine whether they understand the material, judge the trustworthiness of sources (epistemology), and decide when to abandon or modify plans and strategies and when to increase effort. Also, they have to decide when to stop looking for information and how to organize their findings, as well as learn and elaborate upon relevant information (Williams, 1996). In sum, if learners are not able to cope with the obstacles they are confronted with during learning on the WWW, their learning process as well as their learning outcome will be of poor quality. The majority of studies have shown that this is often the case (Dillon & Gabbard, 1998; Shapiro & Niederhauser, 2004).

1.2 Self-Regulated Learning and the World Wide Web

Based on SRL research, engaging in SRL processes enables learners to cope with the obstacles they are confronted with during learning in environments with high degrees of freedom (Zimmerman & Schunk, 2001). Accordingly, the deployment of SRL processes during web-based learning is considered to be an indicator for a learning process of high quality, which in turn is assumed to entail a learning outcome of high quality. In the following, we describe current SRL models in order to derive the aspects of SRL that were relevant to our approach.

Boekaerts (1999) has suggested three systems that have to be regulated for successful learning: a motivational/emotional, a cognitive, and a metacognitive system. In the motivational/emotional system, which may be described metaphorically as the engine of learning, the learner manages motivational and emotional states to get started, to stay on task, and to overcome obstacles and manage negative emotions. The cognitive system provides the basic armamentarium for learning, such as strategies for calculating a math task or for reading a text. By means of the metacognitive system, the learner takes a metaperspective in order to plan, monitor, regulate, and evaluate his learning process. Zimmerman (2000) has taken a process view of SRL by allocating processes, which are presumed to enhance the quality of learning, within three cyclical phases of action. The forethought phase precedes efforts to act and sets the stage for it, the performance or volitional control phase involves processes that occur during action, and the self-reflection phase involves processes that occur after performance. This approach has been adopted by Schmitz and Wiese (2006), who have focused on states and have distinguished between a preaction, an action, and a postaction phase. Winne and Hadwin (2008) and Pintrich (2000) have taken a process view of SRL as well. Further, Alexander (1997) has proposed three levels of SRL that confront learners with different demands: a micro, a mid, and a macro level. On the micro level learners are concerned with elementary tasks like web-based learning. On the mid level, learners organize their daily study routine, whereas on the macro level learners are involved in life and career management.

In our approach, we focused on micro level learning and combined the metacognitive system with a process view. More specifically, we assumed that the quality of an elementary task, which is conducted on the WWW, is enhanced by the deployment of the six metacognitive processes of goal setting & planning, self-monitoring & process-regulation, and reflection & modification. Those processes are located in the three cyclical phases of learning (see Figure 1). In the preaction phase, before the actual learning has begun, it is considered essential to define relevant goals in order to lead learning in a beneficial direction. The involvement in planning processes then enables the attainment of previously set learning goals. In the action phase, during the actual learning, carrying out self-monitoring activities allows for the detection of inefficient and ineffective processes of learning. By engaging in process-regulation, those disadvantageous processes may be altered during the ongoing learning process, and beneficial processes can be reestablished. In the postaction phase, after the actual learning, the reflection on the learning process and the learning outcome allows for elaborating content and provides a basis for the modification of learning strategies for the next learning episode. It is assumed that the deployment of those six metacognitive processes is an indicator for a learning process of high quality, which in turn entails a learning outcome of high quality. However, current research provides evidence that outcomes of different complexities, like factual knowledge, structure, or understanding, might be affected by the learning process in different ways (Azevedo & Cromley, 2004; Azevedo, Cromley, & Seibert, 2004; Bannert, 2006). It is further assumed that the deployment of those six metacognitive processes enables learners to cyclically adapt their learning on the WWW and thereby to become experts in the long run.

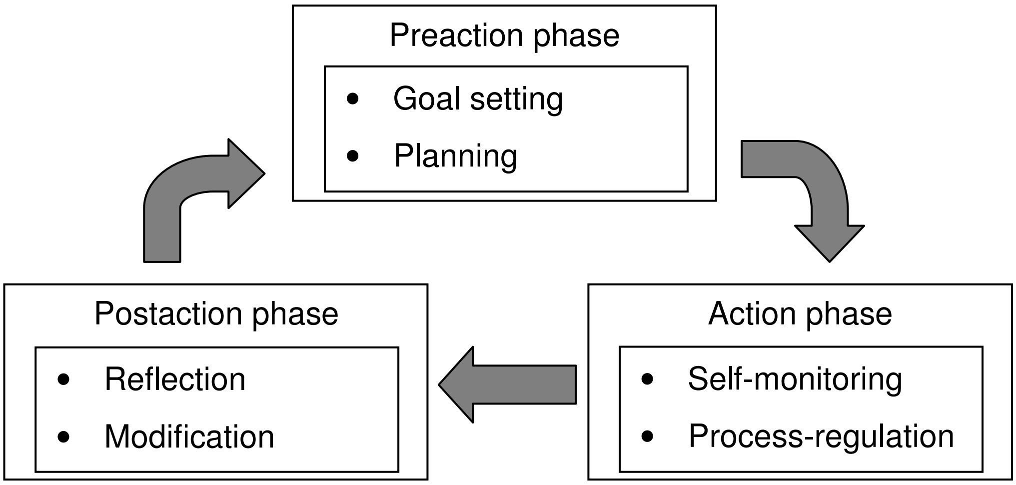


Figure 1. Six metacognitive processes located in the three phases of action.

1.3 Scaffolding Self-Regulated Learning on the World Wide Web

It was the goal of our approach to help learners to overcome obstacles that they are confronted with during learning on the WWW and thereby to enhance the quality of the learning process and the learning outcome of web-based learning. To provide individuals with adequate assistance, we applied the concept of *scaffolding*, which in everyday life refers to a support, such as a temporary framework that supports workers during the construction of a building. Scaffolding is grounded in the developmental theories of Vygotsky (1978), who suggested a *zone of proximal development* (ZPD), which is the area between what a child can accomplish without help and what the same child can accomplish with assistance. The assistance that is provided in the ZPD is called *scaffolding*. The concept has been introduced to the learning sciences by Bruner (Wood, Bruner, & Ross, 1976) and, in the past 2 decades, along with the advancement of technology and the growing possibilities of providing assistance through computers, has found its way into hypermedia learning (Puntambekar & Hubscher, 2005). In the following, based on current research on scaffolding in hypermedia learning, we derive our approach of scaffolding self-regulated learning on the WWW.

In designing scaffolds to help learners to overcome the obstacles of web-based learning, we had to decide *what* to support, *how* and *when* to support it, and to *whom* the support should be administered (Pea, 2004). In the previous section (1.2), it was already mentioned that in our approach we focused on the processes of goal setting & planning, self-monitoring & process-regulation, and reflection & modification during learning on the

WWW. Accordingly, this was in line with current research that aimed to support metacognitive processes to foster micro level hypermedia learning (e.g., Brush & Saye, 2001). However, in contrast to those studies, which mainly supported single metacognitive processes like self-monitoring or reflection, we pursued a holistic concept of assistance by supporting metacognitive processes in all three cyclical phases of learning. Therefore, we not only aimed to enhance the quality of the learning process and the learning outcome, but also provided support for learners to cyclically adapt their web-based learning and thereby to become experts in the long run.

In general, two complementary approaches of providing learners with metacognitive assistance to enhance the quality of hypermedia learning could be pursued (Benz & Schmitz, 2009). On the one hand, researchers have provided strategy instructions to equip learners with a repertoire of relevant strategies (e.g., Azevedo & Cromley, 2004). This direct training approach (Friedrich & Mandl, 1992) is applied if researchers focus on individuals who suffer from a mediation deficiency and who are hence not in the possession of relevant strategies (Reese, 1962). On the other hand, researchers have integrated process support of various shapes in learning environments to induce the deployment of beneficial SRL processes during hypermedia learning (e.g., Azevedo, Cromley et al., 2004; Brush & Saye, 2001; Greene & Land, 2000; Land & Zembal-Saul, 2003). This indirect approach (Friedrich & Mandl, 1992) is applied if researchers focus on individuals who suffer from a production deficiency and who are then accordingly already equipped with SRL strategies, but do not manage to apply them (Flavell, 1970).

We followed an indirect approach of providing assistance in order to help learners to overcome their production deficiencies (Flavell, 1970), and thereby to help them master the obstacles they are confronted with during web-based learning. In contrast to studies that have delivered assistance through instances outside the computer, like a human tutor (e.g., Azevedo, Cromley et al., 2004) or a sheet of paper (e.g., Greene & Land, 2000), we pursued the approach of integrating computer-based scaffolds within a hypermedia environment (e.g., Land & Zembal-Saul, 2003). On the left hand side of the Firefox web browser, we embedded a sidebar that supported the deployment of the six metacognitive processes derived in the previous section (1.2). This means that we optimized the window through which the WWW is seen, and thereby optimized the WWW as a learning environment itself. The creation of a tool (Jonassen & Reeves, 1996) that may be flexibly applied in the WWW, an open-ended learning environment of great relevance for modern life, exceeds current research, which has mainly focused on closed hypermedia environments.

We followed a new concept of designing indirect metacognitive support. Previous research on hypermedia learning has mainly focused on simply adding metacognitive support to the learning environment, like a window for writing down one's planning or reflection (e.g., Land & Zembal-Saul, 2003). Such an additive approach does not put the optimization of the tools, which are applied to complete a learning task, into focus, but leaves them unaltered. However, it aims at optimizing the way learners apply those tools during task implementation by providing additional metacognitive support. One of the problems of such a two-dimensional approach is that learners, focused on completing their learning task, oftentimes do not perceive additional metacognitive support as instrumental. Instead, they tend to be resistant against alterations to their accustomed learning processes and experience deeper metacognitive processing as an extra burden. In turn, we followed a one-dimensional approach by developing a tool that combines functionality and metacognitive assistance. More specifically, our tool offers several functions that may be employed to complete a learning task on the WWW and that simultaneously induce the deployment of the six metacognitive processes. In other words, our scaffolds serve as functions that can be used to complete a learning task on the WWW, and in applying the scaffolds, learners are bound to engage in metacognitive processes. Hence, we created a learning environment that metacognitive support is immanent to.

In the standard version of our extension, the scaffolds are offered in a nonembedded way (Clarebout & Elen, 2006), leaving the decision of if, how, and when to apply them during web-based learning to the learners. However, following SRL research, the six metacognitive processes of goal setting & planning, self-monitoring & process-regulation, and reflection & modification are considered most beneficial when carried out during a specific phase of learning (see Figure 1). Based on this assumption, our scaffolds can be considered most effective when applied at a certain point during learning on the WWW. However, research suggests that learners may not be able to use nonembedded scaffolds of their own accord (Aleven et al., 2003; Oliver & Hannafin, 2000). We did not know whether these results were also true for our scaffolds that embody functionality and metacognitive assistance. Studies on prompting (e.g., Bannert, 2006; Horz, Winter, & Fries, 2009; Kramarski, & Zeichner, 2001; Schwonke, Hauser, Nuckles, & Renkl, 2006) have aimed to investigate the effects of scaffolds that invasively disrupt learners at various points during the process of learning, such as in the preaction, action, or postaction phase (Schmitz & Wiese, 2006), in order to direct learners to deploy certain processes. Based on this research, we created a second version of our Firefox extension, supplementing functions of our first version by two invasive and directive metacognitive prompts. Thereby we aimed to support learners to make use of the functions in a specific way in order to increase the probability that the metacognitive processes would be carried out as intended.

In sum, it was the goal of our approach to support learners to overcome the obstacles they are confronted with during learning on the WWW and thereby to enhance the quality of the learning process and the learning outcome of web-based learning. Following an indirect approach, we optimized the WWW as a learning environment by integrating scaffolds in the Firefox web browser. Supporting the six metacognitive processes in the three cyclical phases of learning, the scaffolds are based on a holistic concept of assistance. Further, we provided scaffolds that offer functions that can be used to complete a micro level task and at the same time provide metacognitive assistance. In the standard version of our Firefox extension the scaffolds may be applied by learners of their own accord. In an extended version learners are additionally scaffolded by two invasive and directive metacognitive prompts.

1.4 The ELWMS Sidebar

We realized our approach of optimizing the WWW as a learning environment by creating an extension to the Firefox web browser called ELWMS (see Figure 2). The extension is implemented as a sidebar, embedded on the left hand side of the browser. Based on our new concept of indirect metacognitive support, ELWMS provides scaffolds, which offer functions that can be used to complete a learning task on the WWW and provide assistance of the six metacognitive processes. Three main functions are provided: the management of goals, the handling of resources, and the illustration of the created goal-resource structure. In the standard version of ELWMS, we did not integrate any invasive and directive prompts, but rather provided learners with the freedom to apply all functions of their own accord. To simplify matters, in the following, the functions are described within a prototypical scenario (see Table 1).

Before actually starting to learn (preaction phase), the learner is scaffolded to define goals and to plan the process of implementation. ELWMS provides a goal-setting function, which allows for specifying a name, a description, and a tag, as well as the current state of goal completion. For example, if a learner is looking for information on ancient Rome and more specifically on the members of the first triumvirate, he might create a goal with the name "First Triumvirate," add the description "What are the members of the first triumvirate?" organize the goal with the tag "Person," and specify the goal progress as "not

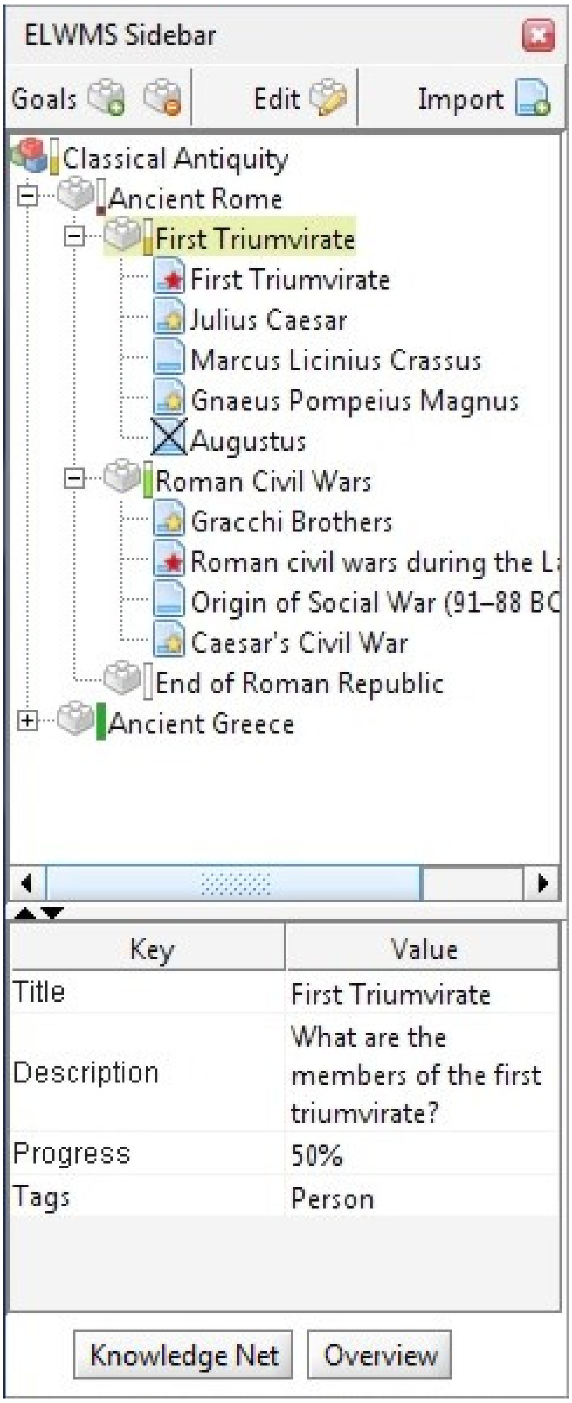


Figure 2. Screen shot of the ELWMS sidebar.

started." If he is further interested in the Roman civil wars and the end of the Roman Republic, he might create several levels of subgoals. ELWMS offers the opportunity to arrange goals in a hierarchical structure in order to organize the upcoming learning phase into sequences.

During the actual learning (action phase), the learner is scaffolded to monitor his learning, and if necessary to engage in process-regulation. ELWMS provides an import function, which allows the user to gather snippets of information by highlighting words, phrases, or paragraphs from web pages. For each resource, a name and a tag may be defined and its relevance may be judged. Snippets are automatically saved within the description of a resource and may be adapted by the user. It is further possible to bookmark whole web pages. Upon the import, resources are assigned to associated goals and a goal-resource structure is created. This structure is illustrated in the sidebar, but may also be viewed in detail when entering the knowledge net or the overview. The knowledge net is similar to a mind map, which presents defined goals and imported resources in a netlike overview. Its advantage lies in the illustration of tags, which are used across goal paths. The overview pictures goals and resources in a hierarchical structure, but displays the full content of persisted resources. It therefore provides a good basis for (re)viewing persisted content. Both knowledge net and overview are nonembedded and may be accessed on demand by clicking on a button. With an increasing number of relevant resources, the progress toward the completion of specific goals may be adapted. In the case of the need for a change in strategy during the action of learning, new goals may be defined and existing goals and resources may be edited, restructured, or deleted. Also the web page that a resource was originally retained from may be reopened.

Table 1

Scaffolds Provided in the Standard Version of ELWMS: Functions and Supported SRL Processes

Phase	Preaction	Action		Postaction
Metacognitive processes	Goal setting & planning	Self-monitoring	Process-regulation	Reflection & modification
Function	Defining goals	Assigning resources to goals	Defining new goals	Viewing goals
	Structuring goals	Defining relevance of resources	Redefining goals	Viewing resources
		Defining progress toward goal completion	Deleting goals	Opening resources
		Viewing knowledge net	Restructuring goals	Viewing goal-resource structure
		Viewing overview	Adapting resources	Viewing knowledge net
			Deleting resources	Viewing overview
			Restructuring resources	
			Opening resources	

Note. ELWMS = E-Learning knoWledge Management System; SRL = self-regulated learning.

Toward the end of learning (postaction phase), reflection and modification processes are scaffolded. Previously defined goals and persisted resources may be viewed in order to reflect on the search process and to review the results. Further, each resource allows the learner to go back to the web page that it was originally retained from. Of course the sidebar, the knowledge net, and the overview, which have already been described, are three illustrations of the created goal-resource structure, and also serve for elaboration purposes.

1.5 Research Questions

Our research is based on our model of adaptive learner support for the enhancement of learning quality (see Figure 3). The model sketches the relation between stable and varying learner characteristics, situational parameters, and the effectiveness of certain realizations of support. In the case of a fit between the preconditions and the design of the scaffolds, the quality of the learning process is enhanced. This may be indicated by the application of SRL strategies, or more specifically by the deployment of metacognitive, cognitive, or motivational/emotional processes. The quality of the learning process in turn has an impact on the achieved learning outcome, which may be affected at different levels of complexity.

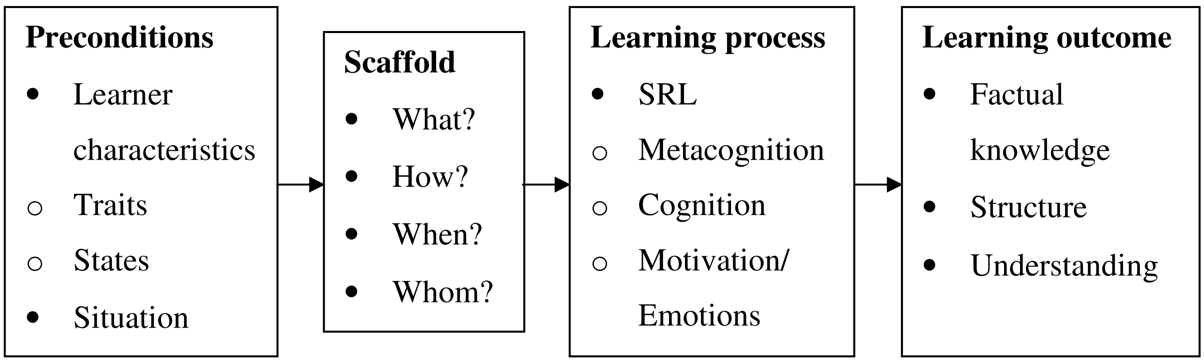


Figure 3. Model of adaptive learner support for the enhancement of learning quality.

The current study served to investigate whether indirect support, as it is realized by our ELWMS software, enhances the quality of learning on the WWW. More specifically, we examined whether optimizing the WWW as a learning environment by integrating scaffolds that embody functionality and at the same time support the six metacognitive processes of goal setting & planning, self-monitoring & process-regulation, and reflection & modification during the three cyclical phases of learning into the Firefox web browser, enhances the quality of the learning process and the learning outcome of web-based learning. We also investigated whether it would be sufficient to provide learners with those scaffolds, leaving the decision of if, how, and when to apply the scaffolds to the learners, or whether learners would achieve higher scores on dependent variables if they received additional invasive and directive prompting to apply the scaffolds.

Our first set of research questions was concerned with the effect of indirect scaffolding on the quality of the learning process during web-based learning. We hypothesized that our scaffolds, which embody functionality and support of the six metacognitive processes in the three phases of learning, would enhance the deployment of SRL processes during learning on the WWW. Accordingly, we presumed that learners who worked with ELWMS would display a learning process of higher quality than learners who did not receive any scaffolding and either just worked with Firefox or additionally were allowed to use pen and paper. We expected to find this pattern for *offline* self-reports, which assessed metacognitive, cognitive, and motivational/emotional processes that had been carried out during task implementation, and for automatically generated objective *online* log files. Based on research on prompting (Aleven et al., 2003; Horz et al., 2009; Oliver & Hannafin, 2000), we further hypothesized that learners who worked with ELWMS and received additional invasive and directive prompts to deploy the processes of goal setting, planning, and reflection at specific points during learning would display a learning process of higher quality than learners who could decide freely when to apply the scaffolds. We expected to find this pattern for *offline* self-reports and for quantitative and qualitative metacognitive scales attained from *online* video analyses. This multi-method approach of collecting data on the learning process allowed for analyses of validity.

Our second set of research questions was concerned with the effect of indirect scaffolding on the quality of the learning outcome during web-based learning. We hypothesized that integrating scaffolds into the Firefox web browser that embody functionality and provide support of the six metacognitive processes in the three phases of learning would enhance the gain of factual knowledge during web-based learning measured by an achievement test. Again, we expected that learners who worked with ELWMS would achieve a learning outcome of higher quality than learners who did not receive any scaffolding and either just worked with Firefox or additionally were allowed to use pen and paper. Referring to research on prompting (Aleven et al., 2003; Horz et al., 2009; Oliver & Hannafin, 2000), we also hypothesized that learners who worked with ELWMS and received additional invasive and directive prompts to deploy the processes of goal setting, planning, and reflection at specific points during learning, would outperform learners who were provided only with the indirect scaffolding of the standard version of ELWMS. For this latter contrast, besides the acquisition of factual knowledge, we evaluated the quality of the created goal-resource structure.

2. Method

2.1 Participants

Participants were $N = 64$ Bachelor of Science Psychology students who were provided with constructive feedback on their learning process and received credit for their participation. Their mean age was 23.1 years. Fifty-three of them were female, 11 were male. All of them were freshmen; however, 23 had previously been enrolled in another career with 1 having graduated. Nineteen had conducted a vocational apprenticeship prior to their studies. The pretest confirmed that all participants had little knowledge of the period of Classical Antiquity.

2.2 Groups

In order to investigate our research questions, we randomly assigned participants to one of four conditions. Experimental group 1 (EG-Tool, $n = 15$) worked with the standard version of our ELWMS software. In this condition learners were provided with the freedom to apply the scaffolds that embodied functionality and provided assistance of the six metacognitive processes if, how, and when they preferred to. Experimental group 2 (EG-Prompt, $n = 15$) worked with an extended version of ELWMS that provided the same scaffolds, but due to two embedded instructional prompts, was more invasive and more directive. At the beginning of the learning period in the preaction phase, participants were prompted to set search goals, which were related to their knowledge gaps in the achievement pretest, and to engage in planning by creating a goal hierarchy prior to their actual search. Five minutes prior to the end of the learning period in the postaction phase, learners were prompted to reflect on their search results and to prepare for the achievement post test.

We created two control conditions to be able to compare the scaffolding conditions to real-world WWW learning approaches. Control group 1 (CG-Firefox, $n = 16$) was only allowed to use the Firefox web browser. The most useful function they were provided with was the bookmarking function, which allowed for saving and organizing links to web pages. Control group 2 (CG-Pen&Paper, n = 18), in addition to the functions of the Firefox web browser, was equipped with pen and paper. Accordingly, as participants in this group had the freedom to apply their established study strategies, control group 2 constituted a very strong condition.

2.3 Procedure

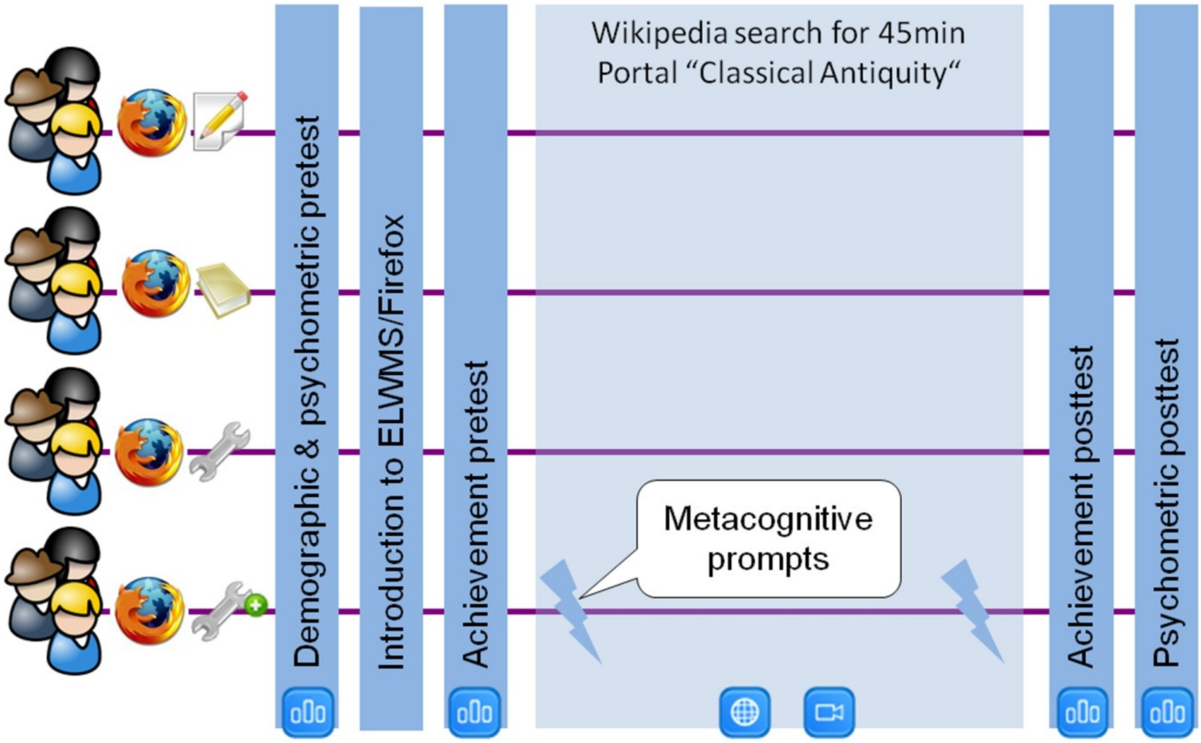


Figure 4. Overview of the design of the study.

The study was conducted in a laboratory that was equipped with 30 up-to-date computers with internet access. To avoid close contact between learners and to keep disturbances at a low level, the 64 participants were divided into 6 groups. To simplify matters, if possible, each group was exposed to the same condition. Prior to each trial, the required software was installed on the computers. Facilitating the navigation through the experiment, we added an *evaluation-menu* to the Firefox menu bar, which allowed for moving from one element of the study to the next. One run (see Figure 4) lasted about 110 minutes.

When entering the lab, participants were randomly assigned to a computer. After a short welcome, the term *hypermedia* was explained and participants were informed that it would be their task to learn in such an environment later in this study. A brief overview of the session was provided and instructions were given concerning the completion of the web-based questionnaires and the handling of the computers. We further asked them not to discuss this experiment with their fellow students, but to wait until the debriefing.

Participants were then given 10 minutes to work on a demographic pretest and a psychometric pretest, which assessed metacognitive skills, computer literacy, and state measures for motivation and self-efficacy. We then conducted a short 5-minute introduction to either Firefox or ELWMS depending on the condition learners had been assigned to. For control groups we emphasized the navigation in Wikipedia, as well as the creation and usage of bookmarks. For experimental groups, besides the navigation in Wikipedia, the handling of goals and resources and the options to display the created goal-resource structure were demonstrated. When carrying out the introduction, we made great efforts to avoid suggesting a learning strategy to participants, and rather presented functions ambiguously.

In the following 20 minutes, participants were administered a 30-item multiple-choice questionnaire on Classical Antiquity. Prior to the test, they were told that they would get the same achievement test as a posttest after having had the chance to look for information in a hypermedia environment. We also pointed out that on the pretest some of the questions might be experienced as challenging, but after the learning period would be much easier.

Having filled out the achievement pretest, participants were instructed to use the next 45 minutes to search for information only by navigating through Wikipedia in order to prepare for an identical achievement posttest. Providing them with a little support, we clarified that it would be the best strategy to establish an overview over the period of Classical Antiquity, focusing on central events, developments, and persons in Ancient Rome and Ancient Greece. We further pointed out that it would not be possible to use the established material (goal-resource structure; bookmarks; paper) when working on the achievement posttest. EG-Tool worked with the standard version of ELWMS, EG-Prompt with a more invasive and more directive version that provided two instructional prompts. CG-Firefox was only allowed to work with Firefox, whereas CG-Pen&Paper was also equipped with pen and paper. As the web-based achievement pretest had already been submitted, participants could not review the questions or their answers, but had to rely on their memories. During the period of learning, log files were collected and a screen recording was conducted. Participants were notified about the remaining time after 25 minutes, and again when 40 minutes had past. After the period of learning, ELWMS and Firefox were automatically closed and notes made on paper were collected.

Participants were allowed to work on the achievement posttest at their preferred pace. Finally, a psychometric posttest was administered assessing self-reports on the learning process, specifically deployed SRL strategies, experienced emotions, as well as state measures for motivation and self-efficacy.

2.4 Learning Environment

At the beginning of the period of learning, the Wikipedia portal *Classical Antiquity* was opened automatically in the browsers of all participants. For this study, learners were limited to navigating within the German Wikipedia, which provided multiple informational sources like text, photographs, as well as static and animated diagrams. During learning, participants were allowed to navigate freely within Wikipedia and to use all incorporated functions, such as the search function and hyperlinks. From the portal, all relevant information for the achievement test could be obtained by either searching for a proper term or by following 1 or 2 hyperlinks.

As Wikipedia is not a standardized learning environment developed for experimental purposes, two disadvantages arose from its usage. First, since Wikipedia is a web 2.0 technology, it is a dynamic environment based on user-generated content. To assure that pages relevant to this study were not essentially changed from the first to the last trial, we conducted the whole study within 1 week. To be on the safe side, we also checked the history of changes of relevant pages. Second, the number of words per Wikipedia page differs. Accordingly, the position of relevant information on a page and the ratio of test-relevant information to non-test-relevant information varied. This imbalance could be compensated for by participants when applying the Firefox search function, which allowed for finding specific terms on a web page. Utilizing Wikipedia, a learning environment with a great relevance for modern life, increased the external validity of our study.

2.5 Measures

The measures applied in this study differed with regard to their natures. On the one hand, we acquired *offline* measures by administering a demographic pretest, as well as psychometric and achievement pre and posttests. All questionnaires were created using a web-based survey application called LimeSurvey (Version 1.5.3) and accordingly could be accessed through a URL. We added an *evaluation-menu* to the Firefox menu bar, which was linked to the questionnaires. In the experimental condition, the ELWMS sidebar was hidden while participants were working on the questionnaires, appeared automatically when the learning period began, and was hidden again when posttests were accessed. On the other hand, *online* measures were obtained by collecting log files and by recording computer screens.

2.5.1 Demographic Pretest

We assessed the common demographic variables age, gender, native language, career, semester, career and/or apprenticeship before current career, and overall high school GPA.

2.5.2 Psychometric Pretest

With the psychometric pretest, we assessed metacognitive skills, computer literacy, and state measures for motivation and self-efficacy. Metacognitive skills were measured by the *planning*, *self-monitoring,* and *regulating* scales from the German version of the Motivated Strategies for Learning Questionnaire (MSLQ; Pintrich, Smith, Garcia, & McKeachie, 1993) and by the *goal directed attention-a* and *goal maintenance* scales from the German version of the Volitional Components Questionnaire II (VCQII; Kuhl & Fuhrmann, 1998). Computer literacy was assessed by a self-developed 14-item questionnaire that contains the scales *general computer skills*, *web search skills*, *experience with Firefox,* and *experience with web 2.0 technologies.* State motivation and state self-efficacy each were evaluated by 2 items that were also developed by the authors. For pretest measures, no differences could be determined between groups.

2.5.3 Achievement Pre- and Posttest

The achievement test, which was used as pre and posttest, contained 30 multiple-choice questions on the period of Classical Antiquity. All participants received the questions in the same order. The first question was a general one on Classical Antiquity, whereas questions 2 to 16 referred to Ancient Rome, and questions 17 to 30 to Ancient Greece. Each block comprised one ranking question, which asked learners to place the phases of Ancient Rome and Ancient Greece in the correct order. The remaining 28 questions were classical multiple-choice items that offered four alternatives. One option always represented the correct answer and another option was designed to closely resemble the correct answer. Two further options constituted reasonable alternatives, which were related to each other, but not to the correct answer.

All questions were very carefully developed from existing Wikipedia pages, making sure that the answers could be found by the learners. Further, we did not want Wikipedia to offer support for any of the multiple-choice answers that we considered to be incorrect. Due to the close relation between the questions and the available Wikipedia information, while participants perceived questions as being quite difficult on the pretest, this impression changed after having had the chance to learn on Wikipedia.

For the correct answer for each of the 28 multiple-choice questions we assigned 1 point; all other options were not rewarded. For the ranking questions, participants received proportional points for each item that was ranked in the correct spot, or 1 point if they had put the phases into the correct order. Accordingly, on the pre and post achievement tests, a maximum of 30 points and a minimum of 0 points could be achieved.

2.5.4 Psychometric Posttest

The psychometric posttest contained a self-developed questionnaire for obtaining self-reports on SRL processes that had been deployed during learning on the WWW. In addition, it evaluated emotions that had been experienced, and state measures on motivation and self-efficacy were collected once again. Further, the appraisal of SRL support during web-based learning and more specifically the perceived utility of ELWMS were evaluated.

Based on scales of the MSLQ (Pintrich et al., 1993) and VCQ II (Kuhl & Fuhrmann, 1998) we created a 48-item self-report questionnaire to assess SRL processes that had been deployed during learning on the WWW. In order to evaluate learning in a hypermedia environment, items were rephrased in a more specific manner. On a 4-point scale learners were supposed to indicate whether they had carried out a process and how they had experienced their learning. Items were presented in three packages that specifically referred to the three phases of learning. The questionnaire was designed in such a way that two items assessed several subprocesses of the scales *goal setting*, *planning* (preaction), *self-monitoring*, *process-regulation*, *cognition*, *motivation*, *subjective experience* (action), *reflection,* and *modification* (postaction). Accordingly, an overall scale for the implementation of metacognitive strategies that included the six metacognitive scales, and an overall scale for self-regulated learning that included all scales, could be created.

Emotions that had been experienced during the period of learning were assessed with the Positive Affect Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988).

The evaluations of state motivation and state self-efficacy were identical to the pretest. Further, to evaluate the acceptance of existing and possible future functions of our ELWMS software, we administered a questionnaire asking for the appraisal of specific support during web-based learning, like goal setting or reflection prompts. In addition, participants had the chance to propose useful functions. We also wanted to know how participants usually learned on the WWW and if they would like to use ELWMS in their everyday lives.

2.5.5 Log Files

In addition to our ELWMS sidebar, we created a Firefox add-on that automatically collected log files that were relevant for our research questions (see Table 2). As participants in the control conditions worked with a regular version of Firefox and learners in the experimental conditions used ELWMS, there were log files that could be obtained for all groups, as well as other log files that could be obtained for only experimental or control groups. To get a better impression of the actions performed by participants in the experimental conditions, log files that were collected for these groups were arranged on time lines that represented the 45-minute learning period of individual learners. An example of such a time line is presented in Appendix A.

Table 2

Automatically Collected Log Files for All Conditions, for Experimental Groups, and for Control Groups

Log files		
All conditions	Experimental conditions	Control conditions
# web pages browsed	# goals defined	# bookmarks created
# web pages uniquely browsed	# resources persisted	
# Wikipedia searches performed	# goals redefined	
# Wikipedia images opened	# resources adapted	
# tabs opened	# goals/resources restructured	
	# goals/resources deleted	
	# resources opened	
	# knowledge net viewed	
	# overview viewed	

Note. # = number of.

2.5.6 Screen Recordings

We used Camtasia Studio (Version 3) to record computer screens during the 45-minute period of learning. In order to analyze the activities of participants who had worked with ELWMS in more detail, quantitative and qualitative video analyses were performed, whereas the individual timelines that had been created on the automatically generated log files served as the basis for the analyses. We developed a system of categories that was used for coding three basic types of ELWMS activities: actions related to goals, actions related to resources, and actions related to navigation. The system was extended during video analyses when nonanticipated activities occurred. The final categories are presented in Appendix B.

Besides the time of occurrence, we supplemented categories by information that was related to the goal, the resource, or the web page that the action was performed on. Goal- and resource-related categories were specified by the name, the description, and the goal tree of the instance. To resource related categories, the URL of a web page that a resource had been obtained from was also added. Goals and resources were given an ID, which allowed for pursuing how they were being adapted throughout the period of learning. If an instance was not deleted but was part of the final goal-resource structure, its final state was indicated. Categories that described the navigation on web pages were supplemented by the corresponding URL.

In order to be able to perform an additional qualitative analysis, goals, resources, and web pages were also assigned a relevance with regard to the achievement test. To keep work at a manageable level, we did not rate the main goal of each goal-resource structure or large resources like entire web pages and information of more then 10 rows. The rating 1 was assigned if a goal had a clear relation to at least one question on the achievement test. Accordingly, those goals represented a question in a way that allowed learners to find the information necessary for answering the question correctly. In turn, a resource that contained such information was also rated 1. Goals were given a 2 if they made sense but did not specifically relate to a question on the achievement test. Correspondingly, resources were assigned a 2 if they helped to narrow down the alternatives for an answer, but did not suffice to determine the correct option. Goals and resources that were rated 3 did not relate to the achievement test and, hence, were not of any help in answering a question. We used two categories to rate entire web pages that learners had been navigating through. Pages were assigned a 2 if their title was promising and a 3 if a visit did not make sense. For instances that had been rated 1 or 2, we also identified the corresponding questions on the achievement test. Accordingly, it was possible to determine whether a resource and the goal it had been assigned to referred to the same question. This relation was described as the *fit* between a goal and a resource.

Coding was completed by four student assistants who had received training and were equipped with detailed material about coding rules. One video was always coded by two raters. Following coding, student assistants met with a researcher to compare their coding. If time of an action, coded category, or assigned relevance and question did not match, discrepancies were discussed in detail and settled in consensus. After the coding, we subdivided categories into the three phases of learning based on time of occurrence. Accordingly, it was possible to determine the number of actions, as well as the number of relevant actions in a specific learning phase. In the next step, on the basis of quantitative as well as qualitative measures of video analyses, we created quantitative and qualitative scales for goal setting, planning, self-monitoring, process-regulation, and reflection and overall scales for metacognition (Appendix C). Hence, for the learning processes of participants from the experimental conditions, two sets of metacognitive scales were established.

Further, an achievement measure was established by determining a value for the quality of the final goal-resource structure that a learner had created. The value was generated by adding the number of goals and resources that were part of the final structure and had received a relevance rating of 1 or 2. However, the higher quality of very relevant goals and resources was accounted for by weighting instances with a relevance of 1 by the factor 2.

2.6 Validating *Offline* and *Online* Measures

To investigate the validity of measures, we correlated *offline* self-reports on the learning process and quantitative and qualitative *online* measures acquired from video analyses. As videos were analyzed only for experimental conditions, results were based on 30 participants who had worked with ELWMS. In this paper we focus on the presentation of the correlations of the aggregated scales.

For matters of clarity, in Table 3 correlations are only presented if they showed at least marginal ($\alpha = .10$) significance. Following the multitrait-multimethod approach (Campbell & Fiske, 1959), when examining correlations between similar and dissimilar measures, convergent validity, indicated by the correlations in the grey cells, is supposed to be high, whereas discriminant validity, represented by the correlations in the white cells, is supposed to be low. In our data we found correlations for process-regulation and for reflection measured by self-reports and by quantitative categories of video analyses. However, process-regulation acquired through quantitative video analyses also correlated with other self-report scales. We further found a positive correlation for reflection, measured by self-reports and qualitative categories of video analyses, and a negative correlation for planning. However, again, process-regulation attained from video analyses and reflection based on self-reports significantly correlated with other variables.

In sum, we found indicators for both convergent and discriminant validity when correlating self-report measures on the learning process with quantitative and qualitative variables from video analyses. Accordingly, *online* as well as *offline* measures raised partly similar and partly different aspects of the process of learning on the WWW. This finding, which in the literature is a well described pattern (Perry & Winne, 2006; Veenman, Prins, & Verheij, 2003; Winne & Jamieson-Noel, 2002), underlined the importance and appropriateness of our multi-method approach.

Table 3

Correlations of Self-Report Scales and Quantitative and Qualitative Scales Attained from Video Analyses

Variables from video analyses	Self-reports						
	SRL	Meta-cognition	Goal setting	Planning	Self-monitoring	Process-regulation	Reflection
Quantitative							
Metacognition							
Goal setting							
Planning							
Self-monitoring							
Process-regulation	.38*	.34*	.25$^{\#}$	.28$^{\#}$	.31*	.24$^{\#}$	
Reflection							.34*
Qualitative							
Metacognition							
Goal setting							
Planning				-.30*			-.25$^{\#}$
Self-monitoring							.32*
Process-regulation				.35*	-.29$^{\#}$		
Reflection							.31*

Note. SRL = self-regulated learning.

$^{\#}p < .10$, one tailed. $^{*}p < .05$, one-tailed.

3. Results

3.1 Quality of the Learning Process

Our first set of research questions was concerned with the effect of indirect scaffolding on the quality of the learning process during web-based learning. More specifically, we investigated whether the integration of our scaffolds, which served as functions and simultaneously provided metacognitive assistance, into the Firefox web browser would have beneficial effects on the learning process. And further, whether learners, who received invasive and directive prompts to apply the scaffolds, would perform a more sophisticated learning process than learners who had the freedom to use the scaffolds of their own accord. We conducted two sets of ANOVAs with the four conditions as levels of the independent variable, and *offline* self-report measures as well as automatically generated log files as dependent variables. The differences between EG-Tool and EG-Prompt were additionally investigated by contrasting quantitative and qualitative *online* scales attained from video analyses.

3.1.1 Self-Reports

Results from the first set of ANOVAs, which applied the four conditions EG-Tool, EG-Prompt, CG-Firefox, and CG-Pen&Paper as levels of the independent variable and self-report scales as dependent variables, are presented in Table 4. We found highly significant differences between groups for the overall scale of *self-regulated learning*. Experimental groups reported deploying significantly more SRL processes then CG-Firefox or CG-Pen&Paper. The same pattern was found for the overall scale of *metacognition* and for the scale of *process-regulation,* as well as the positive items of the PANAS (Watson et al., 1988). Groups also differed significantly in their reports on *planning*, *reflection,* and *motivation*. Experimental conditions were significantly more involved in *planning* and *reflection* than CG-Firefox, and further were more motivated than CG-Pen & Paper. With regard to g*oal setting* and *subjective experience,* groups differed marginally. Based on self-reports, for *self-monitoring*, *cognition*, the negative items of the PANAS (Watson et al., 1988), and mean *state motivation* as well as mean *state self-efficacy,* no significant differences between groups could be obtained. Further, we did not find differences between experimental groups on self-report scales.

Table 4

Differences Between Groups in the Quality of the Learning Process Based on Offline Self-Reports

	Conditions					Planned comparisons		
	EG-Tl (*n* = 15)	EG-Ppt (*n* = 15)	CG-Ffx (*n* = 16)	CG-PnP (*n* = 18)		EG-Tool & EG-Ppt vs. CG-Ffx[a]	EG-Tool & EG-Ppt vs. CG-PnP[a]	EG-Tool vs. EG-Ppt[a]
Self-report scales	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*F*			
SRL	2.51 (0.28)	2.50 (0.37)	2.14 (0.34)	2.19 (0.30)	$F(3, 63) = 5.95$, $p = .001$, $\eta^2 = .23$	$t(60) = 3.66$, $p < .001$, $\eta^2 = .18$	$t(60) = 3.29$, $p = .001$, $\eta^2 = .15$	$t(60) = .09$, $p = .463$, $\eta^2 = .00$
Metacognition	2.53 (0.36)	2.53 (0.38)	2.11 (0.43)	2.27 (0.34)	$F(3, 63) = 4.76$, $p = .005$, $\eta^2 = .19$	$t(60) = 3.62$, $p < .001$, $\eta^2 = .18$	$t(60) = 2.32$, $p = .012$, $\eta^2 = .08$	$t(60) = .03$, $p = .487$, $\eta^2 = .00$
Goal setting	2.83 (0.51)	2.63 (0.70)	2.31 (0.73)	2.29 (0.69)	$F(3, 63) = 2.48$, $p = .070$, $\eta^2 = 0.11$	$t(60) = 2.05$, $p = .023$, $\eta^2 = .07$	$t(60) = 2.23$, $p = .015$, $\eta^2 = .08$	$t(60) = .83$, $p = .206$, $\eta^2 = .01$
Planning	2.13 (0.65)	1.87 (0.41)	1.45 (0.45)	1.86 (0.67)	$F(3, 63) = 3.93$, $p = .013$, $\eta^2 = .16$	$t(60) = 3.16$, $p = .001$, $\eta^2 = .14$	$t(60) = 0.83$, $p = .204$, $\eta^2 = .01$	$t(60) = 1.31$, $p = .099$, $\eta^2 = .03$

Self-report scales	Conditions: EG-Tl (*n* = 15) *M* (*SD*)	EG-Ppt (*n* = 15) *M* (*SD*)	CG-Ffx (*n* = 16) *M* (*SD*)	CG-PnP (*n* = 18) *M* (*SD*)	*F*	Planned comparisons: EG-Tool & EG-Ppt vs. CG-Ffx[a]	EG-Tool & EG-Ppt vs. CG-PnP[a]	EG-Tool vs. EG-Ppt[a]
Self-monitoring	2.23 (0.46)	2.25 (0.63)	2.05 (0.67)	2.18 (0.46)	$F(3, 63) = .42$, $p = .739$, $\eta^2 = .02$	$t(60) = 1.12$, $p = .134$, $\eta^2 = .02$	$t(60) = 0.37$, $p = .359$, $\eta^2 = .00$	$t(60) = -0.08$, $p = .468$, $\eta^2 = .00$
Process-regulation	2.77 (0.50)	2.86 (0.62)	2.22 (0.74)	2.43 (0.50)	$F(3, 63) = 3.87$, $p = .013$, $\eta^2 = .16$	$t(60) = 3.21$, $p = .001$, $\eta^2 = .15$	$t(60) = 2.17$, $p = .017$, $\eta^2 = .07$	$t(60) = -0.41$, $p = .343$, $\eta^2 = .00$
Reflection	2.43 (0.35)	2.63 (0.42)	2.15 (0.47)	2.33 (0.36)	$F(3, 63) = 3.92$, $p = .013$, $\eta^2 = .16$	$t(60) = 3.09$, $p = .002$, $\eta^2 = .14$	$t(60) = 1.67$, $p = .051$, $\eta^2 = .04$	$t(60) = -1.36$, $p = .089$, $\eta^2 = .03$
Cognition	2.55 (0.48)	2.58 (0.50)	2.20 (0.55)	2.40 (0.57)	$F(3, 63) = 1.69$, $p = .179$, $\eta^2 = .08$	$t(60) = 2.23$, $p = .015$, $\eta^2 = .08$	$t(60) = 1.05$, $p = .149$, $\eta^2 = .02$	$t(60) = -0.13$, $p = .449$, $\eta^2 = .00$
Motivation	2.33 (0.47)	2.51 (0.62)	2.22 (0.45)	1.90 (0.28)	$F(3, 63) = 5.17$, $p = .003$, $\eta^2 = .21$	$t(60) = 1.39$, $p = .086$, $\eta^2 = .03$	$t(60) = 3.79$, $p < .001$, $\eta^2 = .19$	$t(60) = -1.09$, $p = .141$, $\eta^2 = .02$

Self-report scales	Conditions: EG-Tl (n = 15) *M* (*SD*)	EG-Ppt (n = 15) *M* (*SD*)	CG-Ffx (n = 16) *M* (*SD*)	CG-PnP (n = 18) *M* (*SD*)	*F*	Planned comparisons: EG-Tool & EG-Ppt vs. CG-Ffx[a]	EG-Tool & EG-Ppt vs. CG-PnP[a]	EG-Tool vs. EG-Ppt[a]
Subjective experience	2.11 (0.38)	2.07 (0.49)	1.94 (0.39)	1.81 (0.23)	$F(3, 63) = 2.22$, $p = .095$, $\eta^2 = .10$	$t(33.3) = 1.21$, $p = .119$, $\eta^2 = .04$	$t(42.4) = 2.94$, $p = .003$, $\eta^2 = .17$	$t(26.3) = 0.28$, $p = .391$, $\eta^2 = .00$
PANAS positive	2.83 (0.71)	2.91 (0.71)	2.48 (0.75)	2.26 (0.68)	$F(3, 63) = 3.06$, $p = .035$, $\eta^2 = .13$	$t(60) = 1.78$, $p = .040$, $\eta^2 = .05$	$t(60) = 2.91$, $p = .003$, $\eta^2 = .12$	$t(60) = -0.31$, $p = .380$, $\eta^2 = .00$
PANAS negative	1.56 (0.49)	1.68 (0.70)	1.36 (0.38)	1.42 (0.56)	$F(3, 63) = 1.08$, $p = .366$, $\eta^2 = .05$	$t(60) = 1.55$, $p = .064$, $\eta^2 = .04$	$t(60) = 1.22$, $p = .114$, $\eta^2 = .02$	$t(60) = -0.57$, $p = .285$, $\eta^2 = .01$
Mean state motivation	2.85 (0.49)	2.92 (0.54)	2.80 (0.62)	2.58 (0.79)	$F(3, 63) = 0.88$, $p = .454$, $\eta^2 = .04$	$t(60) = 0.44$, $p = .330$, $\eta^2 = .00$	$t(60) = 1.60$, $p = .058$, $\eta^2 = .04$	$t(60) = -0.29$, $p = .387$, $\eta^2 = .00$
Mean state self-efficacy	2.70 (0.44)	2.60 (0.51)	2.81 (0.70)	2.68 (0.59)	$F(3, 63) = 0.37$, $p = .778$, $\eta^2 = .02$	$t(60) = -0.92$, $p = .182$, $\eta^2 = .01$	$t(60) = -0.18$, $p = .430$, $\eta^2 = .00$	$t(60) = 0.48$, $p = .317$, $\eta^2 = .00$

Note. EG-Tl = experimental group working with the extension; EG-Ppt = experimental group receiving additional prompts; CG-Ffx = control group working with Firefox; CG-Pnp = control group additionally working with pen and paper; SRL = self-regulated learning; PANAS positive = positive items of the PANAS (Watson et al., 1988); PANAS negative = negative items of the PANAS.

[a]one-tailed.

3.1.2 Log Files

In a second set of ANOVAs we applied the four conditions EG-Tool, EG-Prompt, CG-Firefox, and CG-Pen&Paper as levels of the independent variable and automatically generated log files as the dependent variable. For log files that could be attained for all groups, a significant difference could be found for *number of browsed web pages,* $F(3, 63) = 3.65$, $p = .017$, $\eta^2 = .15$. EG-Tool ($M = 22.80$, $SD = 9.57$) and EG-Prompt ($M = 16.87$, $SD = 7.17$) had surfed significantly fewer web pages than CG-Firefox ($M = 28.75$, $SD = 12.78$), $t(60) = -2.82$, $p = 004$ (1-tailed), $\eta^2 = .12$, but not than CG-Pen&Paper ($M = 20.94$, $SD = 10.34$), $t(60) = -0.37$, $p = .359$ (1-tailed), $\eta^2 = .00$. The difference in surfed web pages between EG-Prompt and EG-Tool was only marginally significant, $t(60) = 1.59$, $p = .059$ (1-tailed), $\eta^2 = .04$. Further, we could not find differences between groups for *number of web pages uniquely browsed*, $F(3, 63) = 2.11$, $p = .109$, $\eta^2 = .10$; *number of Wikipedia searches performed*, $F(3, 63) = 0.52$, $p = .673$, $\eta^2 = .03$; *number of Wikipedia images opened*, $F(3, 63) = 1.71$, $p = .175$, $\eta^2 = .08$; or *number of tabs opened*, $F(3, 63) = 1.93$, $p = .134$, $\eta^2 = .09$.

For both control groups, the *number of bookmarks* created was evaluated. Calculating a *t* test we found a significant difference, $t(15.1) = 2.87$, $p = .006$ (1-tailed), $\eta^2 = .35$, with CG-Firefox ($M = 2.69$, $SD = 3.67$) on average having created more bookmarks than CG-Paper ($M = 0.06$, $SD = 0.236$). We did not analyze differences between experimental groups based on log files, as this question was covered in more detail by video analyses.

3.1.3 Quantitative and Qualitative Video Analyses

In order to further investigate differences in the quality of the learning process between learners who could freely decide if, how, and when to apply our scaffolds and learners who additionally received invasive and directive prompts, we compared EG-Tool and of EG-Prompt based on quantitative and qualitative scales attained from video analyses (see Table 5). Conducting *t* tests, we identified EG-Tool as significantly more involved in *self-monitoring* than EG-Prompt. In turn, EG-Prompt carried out significantly more *reflection* processes. For the overall scale of *metacognition* as well as the scales of *goal setting*, *planning,* and *process-regulation,* no significant differences between experimental conditions could be determined. For quantitative and qualitative scales from video analyses, differences between groups followed the same pattern.

Table 5

Differences Between Experimental Groups in the Quality of the Learning Process Based on Quantitative and Qualitative Scales Attained from Video Analyses

	Conditions		
	EG-Tl	EG-Ppt	
Scales from	(n = 15)	(n = 15)	
video analyses	*M* (*SD*)	*M* (*SD*)	t^a
Quantitative			
Metacognition	11.64 (3.62)	10.15 (4.05)	$t(28) = 1.07, p = .148, \eta^2 = .04$
Goal setting	4.47 (2.56)	5.00 (3.36)	$t(28) = -0.49, p = .315, \eta^2 = .01$
Planning	-1.33 (1.68)	-1.33 (2.02)	$t(28) = 0.00, p = .500, \eta^2 = .00$
Self-monitoring	36.20 (12.39)	27.20 (9.77)	$t(28) = 2.21, p = .018, \eta^2 = .15$
Process-regulation	22.33 (10.20)	18.40 (8.86)	$t(28) = 1.13, p = .135\ \eta^2 = .04$
Reflection	-3.47 (3.78)	1.47 (2.07)	$t(28) = -4.44, p < .001, \eta^2 = .41$
Qualitative			
Metacognition	4.37 (1.44)	4.05 (1.35)	$t(28) = 0.63, p = .268, \eta^2 = .01$
Goal setting	0.40 (0.63)	0.60 (1.06)	$t(28) = -0.63, p = .267, \eta^2 = .01$
Planning	-1.00 (1.07)	-1.00 (0.85)	$t(28) = 0.00, p = .500, \eta^2 = .00$
Self-monitoring	16.40 (5.26)	13.13 (4.78)	$t(28) = 1.78, p = .043, \eta^2 = .10$
Process-regulation	8.53 (4.19)	7.87 (3.23)	$t(28) = 0.49, p = .315, \eta^2 = .01$
Reflection	-2.47 (1.68)	-0.33 (0.49)	$t(16.3) = -4.71, p < .001, \eta^2 = .58$

Note. EG-Tl = experimental group working with the extension; EG-Ppt = experimental group receiving additional prompts.

[a]one-tailed.

3.2 Quality of the Learning Outcome

Our second set of research questions was concerned with the effect of indirect scaffolding on the quality of the learning outcome during web-based learning. Again, we investigated whether the integration of scaffolds, which served as functions and simultaneously provided metacognitive assistance, would have beneficial effects on the learning outcome. We also examined whether learners who received invasive and directive prompts to apply the scaffolds would perform better than learners who were left with the freedom to decide if, how, and when to use the scaffolds. We conducted an ANCOVA to analyze differences between groups on the achievement posttest. The difference between experimental conditions was additionally investigated by contrasting the quality of the created goal-resource structures.

3.2.1 Achievement Test

Differences between groups on the achievement posttest were analyzed by applying a 1-factor ANCOVA with the four conditions EG-Tool, EG-Prompt, CG-Firefox, and CG-Pen&Paper as levels of the independent variable and achievement on the pretest (EG-Tool: $M = 11.93$, $SD = 4.06$; EG-Prompt: $M = 12.22$, $SD = 2.99$; CG-Firefox: $M = 13.00$, $SD = 4.14$; CG-Pen&Paper: $M = 12.11$, $SD = 2.80$) as the covariate. EG-Tool attained a mean of 16.72 ($SD = 4.14$) correct answers on the achievement posttest, whereas EG-Prompt, CG-Firefox, and CG-Pen&Paper reached medium scores of 15.39 ($SD = 4.59$), 17.29 ($SD = 4.59$), and 17.53 ($SD = 4.25$). The covariate, achievement on the pretest, was significantly related to achievement on the posttest, $F(1,59) = 27.64$, $p < .001$, partial $\eta^2 = .32$. However, there was no significant effect of achievement on the posttest after controlling for achievement on the pretest, $F(3,59) = 1.10$, $p = .355$, partial $\eta^2 = .05$.

3.2.2 Created Goal-Resource Structure

To investigate whether participants who had received additional invasive and directive prompts had created a more sophisticated basis for preparing for the achievement posttest than participants who had the freedom to apply the scaffolds of their own accord, we contrasted the mean quality of the final goal-resource structures of the experimental groups (see Table 6). EG-Tool on average had established structures of higher quality than EG-Prompt. In further analyses we also found that the final structures of EG-Tool contained significantly more resources with a relevance of 1 than the final structures of EG-Prompt. This pattern could not be established for goals with a relevance of 1.

Table 6

Differences Between Experimental Groups on the Quality of the Final Goal-Resource Structure Based on Online Measures Attained from Video Analyses

	Conditions		
	EG-Tl	EG-Ppt	
Variables from	(n = 15)	(n = 15)	
video analyses	*M* (*SD*)	*M* (*SD*)	t^a
Goal-resource structure	39.40 (15.79)	24.33 (17.83)	$t(28) = 2.45$, $p = .011$, $\eta^2 = .18$
Goals relevance 1	5.20 (3.19)	3.53 (3.34)	$t(28) = 1.40$, $p = .087$, $\eta^2 = .07$
Resources relevance 1	10.60 (4.91)	5.53 (3.89)	$t(28) = 3.13$, $p = .002$, $\eta^2 = .26$

Note. EG-Tl = experimental group working with the extension; EG-Ppt = experimental group receiving additional prompts.
[a]one-tailed.

4. Discussion

4.1 Quality of the Learning Process

In our first set of research questions, we investigated the effect of indirect scaffolding on the quality of the learning process. We examined whether scaffolds, which serve as functions for web-based learning and simultaneously provide metacognitive assistance in the three cyclical phases of learning, have beneficial effects on the quality of the learning process compared to two control groups of different strengths. Further, we evaluated whether providing learners with additional invasive and directive prompting to engage in goal setting, planning, and reflection would be more effective than letting the learners decide if, how, and when to apply our scaffolds. To answer this first set of research questions we assessed self-reports on SRL processes, which had been deployed during task implementation. We also automatically generated log files during the 45-minute period of learning. As participants worked with the original Firefox or with ELWMS, there were log files that could be collected for all groups and log files that could only be collected for experimental or control conditions.

We further analyzed screen recordings of experimental groups to create quantitative and qualitative measures for the quality of the learning process.

4.1.1 Experimental Conditions Versus CG-Firefox

In a first contrast, we compared experimental groups against CG-Firefox, which was allowed only to set bookmarks and accordingly was very restricted in their application of learning strategies. Based on self-reports, learners in the two scaffolding conditions deployed significantly more SRL processes during learning on the WWW. Further, significantly more metacognitive processes were carried out. More specifically, learners in the experimental groups were significantly more involved in planning in the preaction phase, in process-regulation in the action phase and in reflection in the postaction phase. Additionally, they experienced more positive emotions. Based on automatically generated log files, learners who received scaffolding visited a significantly smaller number of web pages. Our results indicate that providing learners with indirect scaffolds that embody functionality and metacognitive assistance helps them to learn more directly, to overcome obstacles by adapting their actions, and to elaborate on their learning. Accordingly, they need to visit fewer web pages to find relevant information and experience more positive emotions.

4.1.2 Experimental Conditions Versus CG-Pen&Paper

In a second contrast, we compared experimental conditions to CG-Pen&Paper, which was allowed to use pen and paper in addition to Firefox, and therefore had more freedom to apply preferred strategies than CG-Firefox. Based on self-reports, experimental groups were significantly more involved in SRL during learning on the WWW. They deployed more metacognitive processes overall and specifically carried out more process-regulation in the action phase. Further, their motivation was significantly higher and they experienced more positive emotions. On the basis of automatically generated log files, no differences between experimental groups and CG-Pen&Paper could be perceived.

Integrating these results and the findings from contrasting experimental conditions to CG-Firefox, it seems like the scaffolds as they are realized in our ELWMS software, in general, foster SRL and metacognition during learning on the WWW. Also, learners who receive indirect scaffolding conduct more process-regulation during action and experience more positive emotions. Accordingly, optimizing the WWW as a learning environment by integrating scaffolds, which offer functions that can be used to complete a web-based learning task and at the same time induce the six metacognitive processes in the three cyclical phases of learning, enhances the quality of the learning process of web-based learning. Our results are consistent with results of other studies that also have found beneficial effects of processes support (Azevedo, Guthrie, & Seibert, 2004; Brush & Saye, 2001; Greene & Land, 2000).

However, we found significant differences between the experimental groups and CG-Firefox for planning and reflection, but not between the experimental conditions and CG-Pen&Paper. Also, unlike for the contrast between the experimental groups and CG-Firefox, for the experimental conditions and CG-Pen&Paper, we did not find a significant difference in the number of web pages visited. As CG-Firefox created significantly more bookmarks than CG-Pen&Paper, our results provided evidence that both groups applied different strategies to conduct learning on the WWW. Whereas CG-Firefox had to rely on bookmarks, CG-Pen&Paper chose to make use of pen and paper. It seems like the application of pen and paper allowed learners to carry out more metacognitive processes and to perform a more directed learning. Accordingly, being equipped with adequate tools helps learners to overcome a production deficiency (Flavell, 1970), whereas not being equipped with adequate tools inhibits the deployment of strategies that learners are in possession of. These findings point out the importance of providing learners with adequate tools during learning and also confirm our approach of creating scaffolds, which embody functionality and metacognitive support.

We further found a significant contrast for motivation between the experimental groups and CG-Pen&Paper, but not between the experimental conditions and CG-Firefox. Hence, our study provides evidence that the application of pen and paper during web-based learning benefits the deployment of metacognitive processes, but at the same time overwhelms learners' motivational/emotional systems. The latter is not the case for tools, which are integrated within computer-based environments as indicated by a nonsignificant contrast between the experimental conditions and EG-Firefox. In sum, in our study we found beneficial effects of indirect scaffolding on the quality of the learning process, and at the same time are encouraged to pursue our approach of designing computer-based scaffolds that optimize the WWW as a learning environment and simultaneously embody functionality and metacognitive support in the three phases of learning.

4.1.3 EG-Tool Versus EG-Prompt

By contrasting experimental groups, we could not find significant differences based on self-reports. This was also the case for the qualitative and quantitative scales of *goal setting*, *planning*, and *process-regulation* attained from video analyses. However, for the qualitative and quantitative scales of *self-monitoring* and *reflection*, video analyses revealed differences between groups. These contrasting results between *offline* and *online* measures are not a new phenomenon in current research on learning (Perry & Winne, 2006; Veenman et al., 2003; Winne & Jamieson-Noel, 2002).

As we could not find differences between groups on self-report scales and on three scales attained from video analyses, our results indicate that scaffolds as they are realized in our ELWMS software, do not necessarily have to be supplemented by invasive and directive prompts. In contrast to previous research, which indicated that learners are not able to make use of scaffolds of their own accord (Aleven et al., 2003; Horz et al., 2009; Oliver & Hannafin, 2000), learners seem to be able to manage the application of the scaffolds that are provided by our ELWMS software without help. We assume that this result is mainly due to the functionality of our scaffolds. As they may be applied for completing a learning task on the WWW, learners do not perceive them as a burden, but as a constructive tool whose application is beneficial. Accordingly, we again found evidence for the combination of functionality and metacognitive support as powerful concepts for enhancing the quality of the learning process of web-based learning.

Further, learners who worked with ELWMS and received prompts for goal setting, planning, and reflection carried out fewer self-monitoring processes in general and fewer self-monitoring processes with relevance to the achievement test than those who had the freedom to apply the scaffolds of their own accord. Accordingly, our study provides evidence that too much additional guidance may inhibit the deployment of SRL processes that are not prompted to the same extent, and hence cause a production deficiency (Flavell, 1970). However, considering that EG-Prompt, which was prompted to reflect on their learning in the postaction phase, carried out more reflection processes in general and more reflection-processes that were relevant to the achievement test, the picture is not clear. More research will have to be conducted in order to investigate how prompts affect processes that are meant to be fostered and how they affect processes that are not in focus. Also, our study does not clarify when it is best to provide learners with the freedom to apply scaffolds that embody functionality and provide metacognitive support of their own accord and when they profit from additional invasive and directive prompts.

4.2 Quality of the Learning Outcome

In our second set of research questions we investigated the effect of indirect scaffolding on the quality of the learning outcome. We examined whether scaffolds, which embody functionality and provide metacognitive assistance in the three cyclical phases of learning, foster the quality of the learning outcome compared to two control groups of different strengths. We also evaluated whether providing learners with additional invasive and directive prompting to engage in goal setting, planning, and reflection would be more effective than letting the learners decide if, when, and how to apply the scaffolds. To answer this second set of research questions, we assessed factual knowledge on the period of Classical Antiquity with a 30-item multiple-choice achievement test before and after the 45-minute period of learning. Additionally, based on video analyses, we created a value for the quality of the final goal-resource structure for the learners of both experimental groups.

4.2.1 Experimental Conditions Versus CG-Tool / CG-Pen&Paper

Experimental as well as control groups profited from the 45-minute period of learning by gaining factual knowledge on the period of Classical Antiquity. We could not find differences between the experimental conditions and either one of the control groups on the achievement posttest when controlling for the achievement pretest. It seems like indirect scaffolding, as it has been realized in our ELWMS software, does not provide an additional advantage for the acquisition of factual knowledge. Considering that the two scaffolding conditions deployed more SRL and more metacognitive processes and experienced more positive emotions during learning on the WWW, these results appear counterintuitive.

Based on SRL research (Azevedo, Guthrie et al., 2004; Zimmerman & Schunk, 2001), a learning process of high quality should have beneficial affects on the learning outcome. However, we have to acknowledge that learners working with ELWMS had received only a 5-minute introduction to the tool and no time for practicing. This group had to adapt very quickly to a new learning environment. Besides having to handle new software, they were required to diverge from their familiar approach of learning on the WWW and to apply a new set of SRL strategies. Keeping this in mind, it can be considered remarkable that those learners achieved the same quality of learning outcome as the learners who could apply well-known tools and strategies. At any rate, we have to deal with the question of how our scaffolds can be optimized in order to foster the quality of the learning outcome. However, besides investigating changes in the independent variable to examine effects on the dependent variable, it might be beneficial to utilize further dependent variables. There is evidence that indirect scaffolds may not have an effect on the acquisition of factual knowledge, but on achievement measures of higher complexity, like structure or understanding (Azevedo & Cromley, 2004; Azevedo, Cromley et al., 2004; Bannert, 2006).

4.2.2 EG-Tool Versus EG-Prompt

We could not find a difference between experimental groups on the achievement posttest when controlling for the achievement pretest. However, when looking at the quality of the goal-resource structures that learners had created, we found the group that could apply the scaffolds of their own accord to be superior to the group that was additionally administered invasive and directive prompts. Further analysis provided evidence that this superiority was rather due to the finding of a higher quantity of relevant resources than to the creation of a larger number of relevant goals.

Integrating these results and the findings from research question one, which proved EG-Tool to be more involved in self-monitoring processes in the action phase and EG-Prompt in reflection processes in the postaction phase, this study provides evidence that self-monitoring processes were the key for EG-Tool to achieve a goal-resource structure of higher quality. By comparing their actions to their goals during the action phase, learners who could apply the scaffolds of their own accord were able to find a larger quantity of relevant resources. However, those learners did not deploy enough reflection processes in order to memorize their high-quality findings and therefore could not reach higher scores on the achievement posttest. In turn, learners who were additionally prompted to engage in goal setting, planning, and reflection did not carry out sufficient self-monitoring processes to find the same number of resources of high relevance than EG-Tool. Accordingly, EG-Prompt had missed the opportunity to create the basis to profit from the reflection processes, which they deployed significantly more than EG-Tool. These results indicate that learners tend to follow prompts, but do not deploy processes that are not fostered to the same extent. They are consistent with the findings of other researchers who found learners to use prompts superficially, failing to engage in deeper processing (Greene & Land, 2000).

4.3 Limitations

In this study we have shown that the quality of learning on the WWW can be enhanced by providing learners with indirect support as it is realized in our EWLMS software.

However, there are some open questions with respect to our measures. When correlating self-reports on the learning process, which were attained through adapted MSLQ (Pintrich et al., 1993) and VCQ II (Kuhl & Fuhrmann, 1998) scales, and with quantitative and qualitative scales from video analyses, we found indicators for convergent and for discriminant validity (Campbell & Fiske, 1959). Accordingly, when contrasting EG-Tool and EG-Prompt, no differences based on self-reports could be identified; however, differences could be found on *online* measures. More specifically, our results suggest that the participants of EG-Tool and EG-Prompt experienced their learning processes in the same way, even though objective data indicate that they performed different actions during their learning. Such diverging results between *offline* and *online* measures are a well-known phenomenon in current research on learning (Perry & Winne, 2006; Veenman et al., 2003; Winne & Jamieson-Noel, 2002); however, which factors are responsible for the moderate correspondence remains an open issue.

When applying self-report measures as indicators of the quality of the learning process, one always infers on the basis of subjective impressions of the learners. Hence, the great advantage of asking learners for self-reports is that we gain insight into how learners experienced their learning. For researchers, being aware of the subjective impression of learners is not only valuable per se, but also helps to provide an understanding of alterations in other variables, like motivational and emotional states. However, as we do not know how and to what extent learners' impressions relate to the processes that were actually carried out, we have to be aware that self-reports also have their limitations. In our study, we found differences based on self-reports between the groups that worked with ELWMS and the groups that did not work with ELWMS, but we did not find differences when contrasting the two groups that worked with ELWMS. Perhaps as our experimental conditions were very similar we induced learning processes that were too similar for learners to experience their learning in different ways. Hence, the value of self-reports may decline with an increasing similarity of conditions. In addition, we also have to consider that the experiences of the learners were biased by memory effects due to retrospective assessment and that our questionnaires were not adequately sensitive for assessing slight variations in the learning process.

In turn, when using log files and screen recordings to infer the quality of the learning process, overt actions are applied as indicators for covered processes. Hence, the great advantage of such measures is that we do not have to rely on subjective impressions, but that we can refer to objective indicators. However, as we also do not know how and to what extent overt actions of learners relate to the processes that are actually carried out, log files and screen recordings have their limitations as well. Considering the shortcomings of each method, more than one of them should be applied when evaluating the quality of the learning process. It is a good advice to pursue a multi-method approach, applying each method to its advantage. In this regard, we should not interpret moderate correspondence between measures as a drawback, but we should be aware of the fact that each method is able to provide us with a specific piece of the "actual learning process" puzzle.

4.4 Future Perspectives

Future research will have to deal with two major topics: (a) the design of support for web-based learning, and (b) the assessment of the quality of web-based learning as a precondition for evaluating the effectiveness of the support. With regard to the questions of *what* to support, *how* and *when* to support it, and to *whom* the support should be administered (Pea, 2004), this study provided evidence for the effectiveness of our concept of designing scaffolds. Optimizing the WWW as a learning environment by integrating scaffolds, which offer functions that can be used to complete a task and at the same time support the six metacognitive processes, fostered the deployment of SRL processes, and more specifically of the metacognitive processes of planning, process-regulation, and reflection. However, future research will have to deal with the question of why scaffolds that are based on this powerful concept do not reveal significant gains in factual knowledge.

Considering that following a goal-oriented approach is a precondition for deploying other processes of self-regulated learning and to achieve a performance gain, the fact that we found a difference of only marginal significance between groups for goal setting might explain missing effects on factual knowledge. Accordingly, future research will have to focus on optimizing the design of goal-setting support. However, the small effects in goal setting might not be due only to the design of the goal setting scaffold, but also to failures in learners' prerequisite processes. In our study, to set relevant goals, learners had to get involved in successful self-diagnoses to identify their knowledge gaps on the achievement pretest, memorize and recall those gaps, and transform them into relevant goals. To help learners to pursue a goal-oriented approach during learning on the WWW, we suggest supporting these processes as well.

In addition, the orientation on previously set goals during web-based learning can be considered a precondition for achieving a performance gain. Accordingly, the fact that we did not find significant differences between groups for the process of self-monitoring based on self-reports might also explain missing effects on factual knowledge. This assumption is supported by the finding that based on video analyses a significant difference between experimental groups in the process of self-monitoring was accompanied by a significant difference in the quality of the goal-resource structure that was created. Hence, our study indicates that goal orientation during action plays a key role for finding relevant pages and relevant resources. It will be the task of future research to optimize self-monitoring support during web-based learning to help learners to create a better goal-resource structure. As a consequence, reflection processes may be performed on a high-quality basis and result in higher gains in factual knowledge.

This study also suggests that learners do not necessarily need invasive and directive prompts in addition to scaffolds, which embody functionality and metacognitive support. As learners seem to follow invasive and directive prompts but seem to suppress processes that are not supported to the same extent, it remains unclear if, when, and to what extent it is beneficial to supplement our scaffolds by additional prompts. Future research will have to clarify advantages and disadvantages of invasive and directive prompting. It will also have to deal with the questions whether the prompting of more crucial metacognitive processes may entail a higher gain in factual knowledge, or whether too much additional guidance may rather have negative effects. With respect to the learning outcome, it will also be interesting to investigate differential effects of support on achievement measures of different complexity.

Referring to the assessment of the quality of web-based learning, it will be necessary to reflect on the methods and instruments that are applied in current research. It has already been stated that questionnaires like the MSLQ (Pintrich et al., 1993) and the VCQ II (Kuhl & Fuhrmann, 1998), which assess general learning processes, have only limited value for evaluating web-based learning. In our study, we administered a self-developed questionnaire based on adapted MSLQ and VCQ II scales to be able to assess processes that were relevant for our scenario. It will be the task of future research to construct specific questionnaires for evaluating processes that occur during learning on the WWW.

With regard to *online* measures of learning, in our study we evaluated overt actions of experimental groups by combining log file analyses with very complex video analyses. We could not contrast experimental and control groups based on log files and categories of video analyses because working with and without ELWMS resulted in the implementation of different actions during web-based learning. It will be the challenge of future research to improve the assessment of log files in order to replace time-consuming video analyses by automatically generated data. In this regard it will be important to harmonize the collection of log files with the design of the study.

However, to fully exploit the potential of assessment methods, it will not be sufficient to focus on each method independently. Rather, we need to develop corresponding instruments that are aligned to assess identical processes. For example, when aiming to evaluate the process of goal setting, we recommend creating a methodology to automatically collect relevant actions and at the same time to design a questionnaire that assesses subjective impressions of the process in focus. Such a synchronized multi-method approach should provide better insight into the quality of the learning process and reveal improved correlations between measures.

Acknowledgments

This research was carried out within the Interdisciplinary Research Training Group *Feedback Based Quality Management in eLearning* funded by the German Research Foundation (DFG). The authors would like to thank Jianhua Lambrecht, Anna Schmidt, Karim Hirsch, and Thomas Loesch for their assistance. The research was part of the doctoral dissertation of Bastian F. Benz. An earlier version of this article was presented at the 2008 biannual meeting of the Special Interest Group Metacognition (SIG16) of the European Association for Research on Learning and Instruction (EARLI) in Ioannina, Greece, and at the 2008 Congress of Psychology in Berlin, Germany.

References

Aleven, V., Stahl, E., Schworm, S., Fischer, F., & Wallace, R. M. (2003). Help seeking and help design in interactive learning environments. *Review of Educational Research, 73(2)*, 277-320. doi:10.3102/00346543073003277

Alexander, P. A. (1997). Mapping the multidimensional nature of domain learning: The interplay of cognitive, motivational and strategic forces. In M. L. Maehr & P. R. Pintrich (Eds.), *Advances in Motivation and Achievement* (Vol. 10, pp. 213-250). Greenwich, CT: JAI Press.

Azevedo, R., & Cromley, J. G. (2004). Does training on self-regulated learning facilitate students' learning with hypermedia? *Journal of Educational Psychology, 96(3)*, 523-535. doi:10.1037/0022-0663.96.3.523

Azevedo, R., Cromley, J. G., & Seibert, D. (2004). Does adaptive scaffolding facilitate students' ability to regulate their learning with hypermedia? *Contemporary Educational Psychology, 29*(3), 344-370. doi:10.1016/j.cedpsych.2003.09.002

Azevedo, R., Guthrie, J. T., & Seibert, D. (2004). The role of self-regulated learning in fostering students‘ conceptual understanding of complex systems with hypermedia. *Journal of Educational Computing Research, 30(1)*, 87-111. doi:10.1016/j.cedpsych.2003.09.002

Bannert, M. (2006). Effects of reflection prompts when learning with hypermedia. *Journal of Educational Computing Research, 35(4)*, 359-378. doi:10.2190/94V6-R58H-3367-G388

Benz, B. F., & Schmitz, B. (2009). *Self-regulated learning and academic success: Do self-regulated learning interventions enhance performance? A meta-analysis.* Manuscript submitted for publication.

Boekaerts, M. (1999). Self-regulated learning: Where we are today. *International Journal of Educational Research*, 31(6), 445-457. doi:10.1016/S0883-0355(99)00014-2

Brush, T., & Saye, J. (2001). The use of embedded scaffolds with hypermedia-supported student-centered learning. *Journal of Educational Multimedia and Hypermedia, 10(4)*, 333-356.

Campbell, D. T; & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56(2)*, 81-105. doi:10.1037/h0046016

Camtasia (Version 3) [Computer software]. Okemos, MI: TechSmith Corporation.

Clarebout, G., & Elen, J. (2006). Tool use in computer-based learning environments: Towards a research framework. *Computers in Human Behavior, 22(3)*, 389–411. doi:10.1037/h0046016

Dillon, A., & Gabbard, R. (1998). Hypermedia as educational technology: A review of the quantitative research literature on learner comprehension, control, and style. *Review of Educational Research, 68(3)*, 322-349. doi:10.2307/1170600

Flavell, J. H. (1970). Developmental studies of mediated memory. In H. W. Reese & L. P. Lipsitt (Eds.), *Advances in child development and behaviour* (Vol. 5, pp. 181-211). New York: Academic Press.

Friedrich, H. F., & Mandl, H. (1992). Lern- und Denkstrategien – ein Problemaufriss [Learning and thinking strategies: A sketch of the problem]. In H. Mandl & H. F. Friedrich (Eds.), *Lern- und Denkstrategien. Analyse und Intervention* (pp. 3-54). Goettingen, Germany: Hogrefe.

Greene, B., & Land, S. (2000). A qualitative analysis of scaffolding use in a resource-based learning environment involving the World Wide Web. *Journal of Educational Computing Research, 23(2)*, 151-179. doi:10.2190/1GUB-8UE9-NW80-CQAD

Horz, H., Winter, C., Fries, S. (2009). Differential benefits of situated instructional prompts. *Computers in Human Behavior, 25(4)*, 818-828. doi:10.1016/j.chb.2008.07.001

Jacobson, M., & Archodidou, A. (2000). The design of hypermedia tools for learning: Fostering conceptual change and transfer of complex scientific knowledge. *Journal of the Learning Sciences, 9(2)*, 149-199. doi: 10.1207/s15327809jls0902_2

Jonassen, D. (1996). *Computers as mind tools for schools.* Columbus, OH: Merril.

Jonassen, D., & Reeves, T. (1996). Learning with technology: Using computers as cognitive tools. In D. Jonassen (Ed.), *Handbook of research for educational communications and technology* (pp. 694-719). NY: Macmillan.

Kramarski, B., & Zeichner, O. (2001). Using technology to enhance mathematical reasoning: effects of feedback and self-regulation learning. *Educational Media International, 38(2-3)*, 77-82. doi: 10.1080/09523980110041458

Kuhl, J., & Fuhrmann, A. (1998). Decomposing self-regulation and self-control: The volitional component inventory. In J. Heckhausen & C. S. Dweck (Eds.), *Motivation and self-regulation across the life span* (pp. 15–49). New York: Cambridge University Press.

Land, S. & Zembal-Saul, C. (2003). Scaffolding reflection and articulation of scientific explanations in a data-rich, project-based learning environment: An investigation of progress portfolio. *Educational Technology Research & Development, 51(4)*, 65-84. doi:10.1007/BF02504544

Oliver, K., & Hannafin, M.J. (2000). Student management of web-based hypermedia resources during open-ended problem solving. *Journal of Educational Research, 94(2)*, 75-93. doi: 10.1080/00220670009598746

Palincsar, A.S. (1998). Keeping the metaphor of scaffolding fresh—A response to C. Adison Stone's "The metaphor of scaffolding: Its utility for the field of learning disabilities". *Journal of Learning Disabilities, 31*(4), 370-373. doi: 10.1177/002221949803100406

Pea, R.D. (2004). The social and technological dimensions of scaffolding and related theoretical concepts for learning, education, and human activity. *Journal of the Learning Sciences, 13*(3), 423-451. doi:10.1207/s15327809jls1303_6

Perry, N. E., & Winne, P. H. (2006). Learning from learning kits: gStudy traces of students' self-regulated engagements using software. *Educational Psychology Review, 18(3)*, 211-228. doi:10.1007/s10648-006-9014-3

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In P. R. Pintrich, M. Boekaerts & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 452-502). San Diego, CA: Academic Press.

Pintrich, P. R., Smith, D. A., Garcia, T., & McKeachie, W. (1993). Reliability and predictive validity of the Motivated Strategies for Learning Questionnaire (MSLQ). *Educational and Psychological Measurement, 53(3)*, 801-813. doi:10.1177/0013164493053003024

Puntambekar, S., & Hubscher, R. (2005). Tools for scaffolding students in a complex learning environment: What have we gained and what have we missed? *Educational Psychologist, 40(1)*, 1-12. doi:10.1207/s15326985ep4001_1

Rakes, G. C. (1996). Using the internet as a tool in a resource-based learning environment. *Educational Technology, 36(5)*, 52-56.

Reese, H. W. (1962). Verbal mediation as a function of age level. *Psychological Bulletin*, *59(6)*, 502-509. doi:10.1037/h0040739

Schmitz, B., & Wiese, B.S. (2006). New perspectives for the evaluation of training sessions in self-regulated learning: Time-series analyses of diary data. *Contemporary Educational Psychology, 31(1)*, 64-96. doi:10.1016/j.cedpsych.205.02.002

Schmitz, C. (2009). Limesurvey (Version 1.53) [Computer software]. Hamburg, Germany.

Schwonke, R., Hauser, S., Nuckles, M., & Renkl, A. (2006). Enhancing computer-supported writing of learning protocols by adaptive prompts. *Computers in Human Behavior, 22(1),* 77-92. doi:10.1016/j.chb.2005.01.002

Shapiro, A., & Niederhauser, D. (2004). Learning from hypertext: Research issues and findings. In D. H. Jonassen (Ed.). *Handbook of Research for Education Communications and Technology* (2nd ed., pp. 605-620). Mahwah, NJ: Lawrence Erlbaum Associates.

United Nations. (2008). *The global information society: A statistical view*. Chile: United Nations.

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. doi:10.1007/s11409-007-9017-6

Veenman, M. V. J., Prins, F. J., & Verheij, J. (2003). Learning styles: Self-reports versus thinking-aloud measures. *British Journal of Educational Psychology, 73(3)*, 357-372. doi:10.1348/000709903322275885

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* Cambridge, MA: Harvard University Press.

Watson, D., Clark, L.A. & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology, 54(6)*, 1063-70. doi:10.1037/0022-3514.54.6.1063

Williams, M. (1996). Learner control and instructional technologies. In D. Jonassen (Ed.), *Handbook of research on educational communications and technology* (pp. 957-983. New York: Scholastic.

Winne, P., & Hadwin, A. (2008). The weave of motivation and self-regulated learning. In D. Schunk & B. J. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and applications* (pp. 297–314). New York: Taylor & Francis.

Winne, P. H., & Jamieson-Noel, D. (2002). Exploring students' calibration of self reports about study tactics and achievement. *Contemporary Educational Psychology, 27(4)*, 551-572. doi:10.1016/S0361-476X%2802%2900006-1

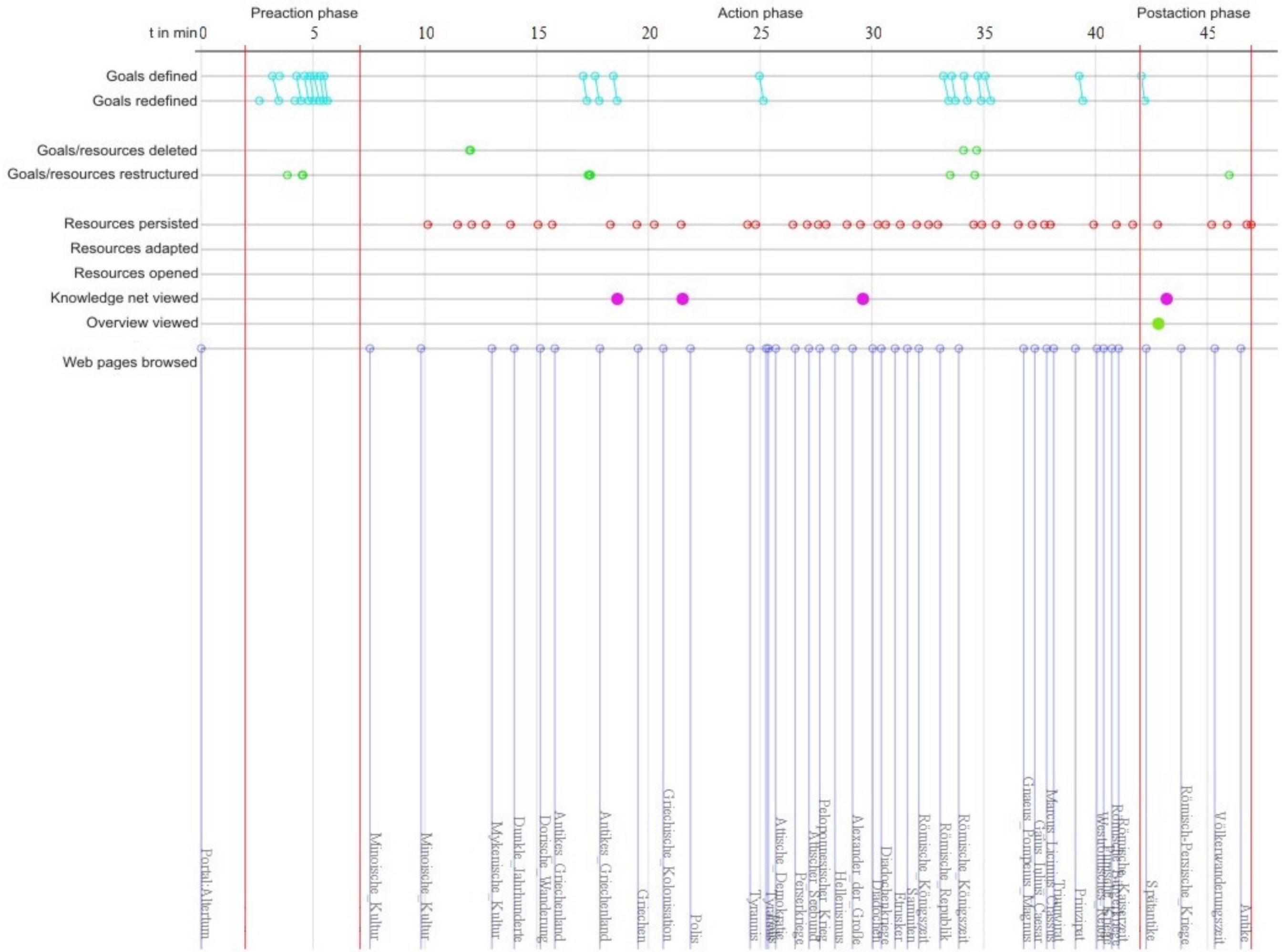
Wood, D. J., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry, 17(2)*, 89-100. doi:10.1111/j.1469-7610.1976.tb00381.x

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich & M. Zeidner (Eds.), *Handbook of self-regulation (pp. 13-39).* San Diego, CA: Academic Press.

Zimmerman, B. J., & Schunk, D. J. (Eds.). (2001). *Self-regulated learning and academic achievement: Theoretical perspectives*. New York: Lawrence Erlbaum Associates.

Appendix A

Timeline of Participant EZR3 Based on Automatically Generated Log Files During the 45-Minute Period of Learning



Appendix B

Final Categories of Video Analyses for Coding Basic Types of ELWMS Activities

Name of category	Relevance rating[a]
Goal categories	
Redefining name of main goal; from "XX/" to "YY/"	
(Re)defining description of main goal "XX/": (from "YY" to) "ZZ"	
Defining progress of main goal "XX/": from "YY%" to "ZZ%"	
Defining tag for main goal "XX/": "YY"	
Defining subgoal: "main goal/XX"	X
Defining new subgoal by redefining name of existing goal; from "main goal/XX" to "main goal/YY"	X
Redefining name of subgoal; from "main goal/XX" to "main goal/YY"	X
(Re)defining description of subgoal "main goal/XX": (from "YY" to) "ZZ"	X
Defining progress of subgoal "main goal/XX/": from "YY%" to "ZZ%"	X
Defining tag for subgoal "main goal/XX": "YY"	X
Deleting subgoal "main goal/XX"	X
Deleting subgoal "main goal/XX/YY" with parent: "main goal/ZZ"	X
Restructuring goals; moving subgoal "XX/ZZ": "YY/ZZ"	X
Restructuring goals; moving subgoal "XX" within goal "YY"	X
Resource categories	
Import of web page "XX" to goal "YY"	
Import of introductory paragraph "XX" (<= 10 rows) to goal "YY"	X
Import of introductory paragraph "XX" (> 10 rows) to goal "YY"	
Import of information "XX" (<= 10 rows) to goal "YY"	X
Import of information "XX" (> 10 rows) to goal "YY"	

Name of category	Relevance rating[a]
Import; Using description of subgoal "main goal/XX" as resource: (from "YY" to) "ZZ"	X
Redefinition; Using description of subgoal "main goal/XX" as resource: (from "YY" to) "ZZ"	X
Goal: “XX“; Redefining name of resource; from “YY” to “main goal/ZZ”	X
Goal: “XX“; (Re)defining description of resource "YY": (from "ZZ" to) "AA"	X
Goal: “XX“; Defining relevance of resource "YY": "ZZ"	X
Goal: “XX“; Defining tag for resource "YY": "ZZ"	X
Goal: “XX“; Redefining tag for resource "YY”: from "ZZ" to "AA"	X
Goal: “XX“; Deleting resource "YY"	X
Restructuring resource; moving resource "XX/YY ": "ZZ/YY"	X
Navigation categories	
Viewing knowledge net	
Viewing overview	
Following link; Entering new web page: “XX”	X
Following link; Entering previously visited web page: “XX”	X
Goal “XX“; Opening resource „YY”; Entering previously visited web page: “ZZ”	X
Searching for "XX"; Entering new web page: “YY”	X
Searching for "XX"; Entering previously visited web page: “YY”	X
Searching for "XX", Failure	
Opening link in new inactive tab: "XX"	X
Changing tab; Entering new web page: “XX”	X
Changing tab; Entering previously visited web page: “XX”	X

Name of category	Relevance rating[a]
Closing active tab; Entering new web page: “XX”	X
Closing active tab; Entering previously visited web page: “XX”	X
Closing inactive tab "XX"	X
Navigating "backwards" to previous web page: "XX"	X
Navigating "forward" to previous web page: "XX"	X

Note. ELWMS = E-Learning knoWledge Management System. XX, YY, ZZ, AA are variables for specific values within categories and are not related across categories.
[a]Categories that received a relevance rating are marked.

Appendix C

Scheme for Establishing the Quantitative and Qualitative Metacognitive Scales Based on Categories from Video Analyses

		Quantitative and qualitative metacognitive scales				
Name of category	Phase	Goal setting	Planning	Self-monitoring	Process-regulation	Reflection
Goal categories						
Defining subgoals	Pre	+ qt / r1				
	Act				+ qt / r1	
	Post					
Redefining main goal (name, description)	Pre	+ qt				
	Act				+ qt	
	Post					
Redefining subgoals (name, description)	Pre	+ qt / r1				
	Act				+ qt / r1	
	Post					
Defining progress of main goal	Pre		+ qt			
	Act			+ qt		
	Post					+ qt
Defining progress of subgoals	Pre		+ qt / r1			
	Act			+ qt / r1		
	Post					+ qt / r1
Defining tags for main goal	Pre		+ qt			
	Act			+ qt		
	Post					
Defining tags for subgoals	Pre		+ qt / r1			
	Act			+ qt / r1		
	Post					

Name of category	Phase	Quantitative and qualitative metacognitive scales				
		Goal setting	Planning	Self-monitoring	Process-regulation	Reflection
Deleting subgoals	Pre	+ qt / r3				
	Act				+ qt / r3	
	Post					
Restructuring subgoals	Pre		+ qt / r1			
	Act				+ qt / r1	
	Post					
Resource categories						
Importing resources (no relevance rating)	Pre		- qt			
	Act			+ qt		
	Post					- qt
Importing resources (with relevance rating)	Pre		- qt / r1			
	Act			+ qt / r1		
	Post					- qt / r1
Redefining resources (name, description, tag)	Pre					
	Act				+ qt / r1	
	Post					
Defining tag for resources	Pre					
	Act			+ qt / r1		
	Post					
Defining relevance of resources	Pre					
	Act			+ qt / r1		
	Post					
Deleting resources	Pre					
	Act				+ qt / r3	
	Post					

Name of category	Phase	Goal setting	Planning	Self-monitoring	Process-regulation	Reflection
		Quantitative and qualitative metacognitive scales				
Restructuring resources	Pre					
	Act				+ qt / r1	
	Post					
Viewing knowledge net and overview	Pre		+ qt			
	Act			+ qt		
	Post					+ qt
Navigation categories						
Entering new web page	Pre		- qt / r1			
	Act			+ qt / r1		
	Post					- qt / r1
Entering previously visited web page	Pre					
	Act				+ qt / r1	
	Post					
Opening resources	Pre					
	Act				+ qt / r1	
	Post					+ qt / r1

Note. Pre = preaction phase (< 5 minutes); Act = action phase (5 – 40 minutes); Post = postaction phase (> 40 minutes). qt = quantitative (all actions are considered); r1 = all actions with the relevance of 1 are considered; r3 = all actions with the relevance 3 are considered; + = value is added; - = value is subtracted. Quantitative scales were established by summing across all quantitative values. Qualitative scales were established by summing the number of all actions with the relevance 1 or 3.

Study 3: Benz, B. F., Scholl, P., Boehnstedt, D., Schmitz, B., Rensing, C., & Steinmetz, R. (2010). *Improving the Quality of Learning on the World Wide Web by Scaffolding Self-Regulated Learning.* Manuscript submitted for publication.

Abstract

With the goal of enhancing the quality of the process and the outcome of web-based learning, the authors optimized the World Wide Web (WWW) as a learning environment by integrating scaffolds, which offer functionality and metacognitive support, into Mozilla Firefox. Continuing their previous research, the authors investigated the effectiveness of their scaffolding approach by randomly assigning N = 108 undergraduate students to three conditions when learning for 45 min on Wikipedia about *Classical Antiquity*. Experimental groups (a) freely applied the scaffolds or (b) received additional intensive prompts to activate processes that were considered to enhance achievement. The control group worked with a no-scaffolding version of the software. The quality of the learning processes was evaluated by assessing deployed processes through self-reports as well as via quantitative and qualitative log data. The quality of the learning outcome was determined by assessing the established structure and gain in factual knowledge. Learners who received scaffolding deployed more metacognitive processes. This was also the case for learners who received prompting compared to learners who freely applied the scaffolds. All in all, the study provides further evidence that the authors' scaffolding approach is a powerful concept for enhancing the quality of web-based learning.

Keywords: self-regulated learning, metacognition, scaffolding, computer assisted instruction, hypermedia, achievement

1. Introduction

The World Wide Web (WWW), which nowadays is used as a resource for learning in various settings (United Nations [UN], 2008), is a nonlinear and unstructured environment (Jonassen, 1996) that provides an enormous degree of freedom. As a consequence, the quality of web-based learning very much relies on the skills and strategies of the learners. It was the goal of our approach to support learners to overcome the obstacles that they are confronted with during learning on the WWW, and thereby to enhance the quality of the learning process and the learning outcome.

The concept of self-regulated learning (SRL) provided a framework for achieving this aim. Referring to current models of SRL (Alexander, 1997; Boekaerts, 1999; Pintrich, 2000; Schmitz & Wiese, 2006; Winne & Hadwin, 2008; Zimmerman, 2000), we derived six metacognitive processes, which can be considered highly relevant for web-based learning. Aiming to enhance learners' involvement in these processes, we followed an indirect approach of assistance (Friedrich & Mandl, 1992) in our research. More specifically, our goal was to optimize the WWW as a learning environment by offering scaffolds (Palincsar, 1998) that serve as functions that can be used to conduct a web-based learning task and that simultaneously induce involvement in the metacognitive processes. As there is evidence that learners may not be able to apply scaffolds of their own accord (Aleven, Stahl, Schworm, Fischer, & Wallace, 2003; Oliver & Hannafin, 2000), we also followed the approach of providing individuals with additional prompting.

In our previous study (Benz, Scholl, Boehnstedt, Schmitz, Rensing, & Steinmetz, 2010), we had applied our indirect scaffolding approach by implementing the first generation of the Mozilla Firefox extension: *E-Learning knoWledge Management System* (ELWMS). We had created a standard version that provided learners with the freedom to apply the scaffolds of their own accord, as well as an extended version that provided learners with additional invasive and directive prompting. After a short introduction to the software, two experimental groups worked with the two versions of ELWMS, and two control groups with the standard version of the Firefox web browser when learning for 45 min on Wikipedia about *Classical Antiquity*. For all groups, the quality of the learning process was assessed by an *offline* self-report questionnaire and *online* basic log data (Veenman, 2007), whereas the quality of the learning outcome was evaluated by gain in factual knowledge on an achievement test. For the two experimental groups, we conducted detailed analyses of screen recordings, which allowed for evaluating the quality of the learning process by an analysis of quantitative and qualitative *online* data, and the quality of the learning outcome by an assessment of the established structure. As Firefox and ELWMS entailed overt actions of a different nature, we were not able to compare all groups on the basis of objective *online* data.When comparing experimental groups against the two control groups, we found positive effects of indirect scaffolding on the quality of the learning process, but not on the quality of the learning outcome. However, when contrasting the two experimental conditions, we found ambiguous effects of prompting.

With the current study, we aimed to gain further insight into the impact of our indirect scaffolding approach and into the effect of additional intensive prompting on the quality of web-based learning. Based on the lessons of our previous study (Benz et al., 2010), we applied the second generation of the ELWMS software, revised instruments, and an elaborated study design. To be able to compare all conditions on the basis of log data, we harmonized the design of the study with our method of collecting data by equipping the control group with a downgraded version of ELWMS that offered the same functions but did not provide metacognitive support. We also revised the extended version of ELWMS, providing learners with intensive prompting of processes that were considered to enhance achievement. Aiming to ensure that participants had understood how to operate ELWMS, each group had to complete a web-based training before learning on Wikipedia. The quality of the learning process was evaluated by applying a synchronized multi-method approach, which allowed for analyses of validity. As a pretest, we used *offline* self-reports to assess whether learners would carry out a specific process during web-based learning, and as a posttest, we used the same method to assess whether they had deployed the process during the learning task. In addition, we automatically generated *online* log data on overt actions that were indicators for the same process. In further analyses, we assigned a rating of relevance to logged actions, which allowed for examining the quality of the learning process by evaluating quantitative and qualitative *online* data. To evaluate the quality of the learning outcome, we assessed the established structure as well as the gain in factual knowledge with a revised instrument.

1.1 The World Wide Web as a Learning Environment

The WWW is a hypermedia system that is accessible through the internet. It provides an immense and continually growing amount of information about all kinds of topics represented as text, graphics, animation, audio, and video (Jacobsen & Archodidou, 2000;

Jonassen, 1996). In modern life, the WWW has become a major resource (Rakes, 1996) for obtaining information and for utilizing education or learning activities in vocational, educational, and private settings (UN, 2008). However, especially with the paradigm change to web 2.0 technologies like Wikipedia, information on the WWW is created by all sorts of entities and, hence, is commonly not organized in a manner that benefits learning. As a consequence, when employing the WWW as a learning environment, it is the responsibility of the users to master the freedom they are confronted with and thereby to profit from the environment. Besides having to decide what to learn, how much to learn, how to learn, and how much time to spend, individuals have to navigate through the WWW, find relevant resources, determine whether they understand the material, judge the trustworthiness of sources (epistemology), decide when to abandon or modify plans and strategies, and determine when to increase effort. Further, they have to decide when to stop looking for information, how to organize their findings, and they have to learn and to elaborate upon relevant information (Williams, 1996). If learners are not able to cope with the obstacles they are confronted with during learning on the WWW, their learning process as well as their learning outcome will be of poor quality. The majority of studies have shown that this is often the case (Dillon & Gabbard, 1998; Shapiro & Niederhauser, 2004).

1.2 Scaffolding Self-Regulated Learning on the World Wide Web

To help learners to overcome the obstacles that they are confronted with during learning on the WWW, we aimed to provide them with adequate assistance, referring to the concept of *scaffolding (*Puntambekar & Hubscher, 2005; Vygotsky, 1978; Wood, Bruner, & Ross, 1976). In designing scaffolds, we had to decide *what* to support, *how* and *when* to support it, and to *whom* the support should be administered (Pea, 2004). In the following, we present our approach of offering assistance, which had proven to be a promising concept in our previous research (Benz et al., 2010).

1.2.1 The SRL Approach

From a social cognitive theoretical perspective, SRL is defined as learners' self-generated thoughts, feelings, and actions that are systematically oriented toward the attainment of their learning goals (Zimmerman & Schunk, 2001). According to SRL research, engaging in SRL processes enables individuals to cope with learning environments that provide a large degree of freedom. Hence, engaging in SRL is considered to be an indicator for a learning process of high quality, which in turn entails a learning outcome of high quality. In the rest of this section, we derive our SRL approach.

Focusing on the metacognitive system (Boekaerts, 1999) and taking a process view of SRL (Pintrich, 2000; Schmitz and Wiese, 2006; Winne and Hadwin, 2008; Zimmerman, 2000), we allocated six metacognitive processes in the three cyclical phases of learning (see Figure 1). More specifically, we assumed that the deployment of goal setting and planning in the preaction phase, of self-monitoring and process-regulation in the action phase, and of reflection and modification in the postaction phase, would enhance the quality of the learning processes and the learning outcome. We thereby focused on an elementary learning task that is conducted on the WWW, and thus on the micro level of learning (Alexander, 1997).

In detail, when learning on the WWW, in the preaction phase before the actual learning is started, it is considered essential to define relevant goals in order to lead learning in a beneficial direction. Becoming involved in planning processes then enables the attainment of previously set learning goals. In the action phase, during the actual learning, carrying out self-monitoring activities allows for the detection of inefficient and ineffective processes of learning. By engaging in process-regulation, those disadvantageous processes can be altered during the ongoing learning process, and beneficial processes can be reestablished. In the postaction phase, after the actual learning, the reflection on the learning process and the learning outcome allows for elaborating content, and provides a basis for the modification of learning strategies for the next learning episode.

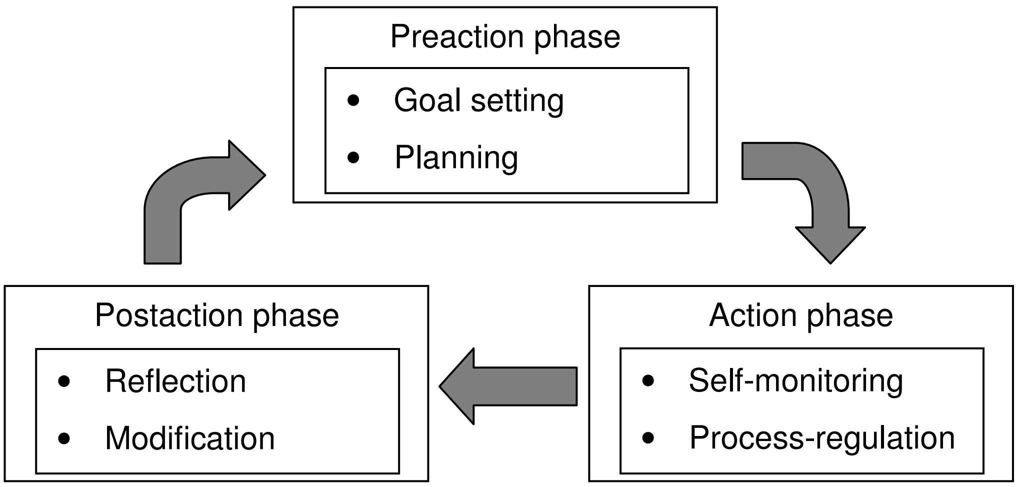


Figure 1. Six metacognitive processes located in the three phases of action.

In sum, we focused on enhancing learners' deployment of six metacognitive processes that are considered to be particularly relevant for web-based learning. In contrast to other research that has mainly aimed to support single metacognitive processes like self-monitoring or reflection in micro-level hypermedia learning (e.g., Brush & Saye, 2001), we pursued a holistic concept of support by fostering metacognitive processes in all three cyclical phases of learning. By doing so, additional goals were to enable individuals to cyclically adapt their learning on the WWW and to become experts in the long run.

1.2.2 The Indirect Approach of Assistance

To encourage individuals to deploy SRL processes during web-based learning, two approaches can potentially be pursued. On the one hand, strategy instructions may be implemented to equip learners who suffer from a mediation deficiency (Reese, 1962) with a repertoire of relevant strategies (e.g., Azevedo & Cromley, 2004). On the other hand, supplementing those direct training approaches, indirect approaches focus on learners who are already in possession of relevant strategies, but who are not managing to apply them. With the goal of helping individuals to overcome their production deficiencies (Flavell, 1970), indirect approaches of assistance aim to modify the learning environment in order to induce the deployment of relevant processes.

Following an indirect approach of offering assistance (Friedrich & Mandl, 1992), it was our goal to upgrade the web browser with scaffolds in order to induce learners' deployment of the six metacognitive processes during web-based learning. By doing so, we optimized the window through which the WWW was seen, and thereby optimized the WWW as a learning environment itself. Accordingly, in contrast to studies that had provided indirect assistance in hypermedia learning through instances outside the computer, like a human tutor (e.g., Azevedo, Cromley, & Seibert, 2004) or a sheet of paper (e.g., Greene & Land, 2000), we applied the computer as the instance of delivery (e.g., Land & Zembal-Saul, 2003). This approach allowed us to develop a tool (Jonassen & Reeves, 1996) that could be flexibly applied on the WWW. We thereby exceeded current research on hypermedia learning, which has mainly focused on closed environments.

1.2.3 The Integrated Scaffolding Approach

To provide individuals with indirect assistance, previous research on hypermedia learning has mainly focused on adding metacognitive support to the learning environment, like a window for writing notes about plans or reflections (e.g., Land & Zembal-Saul, 2003). Such an additive approach does not focus on the optimization of the tools that are applied to complete a learning task, but rather leaves them unaltered. However, it aims to optimize the way learners apply those tools during task implementation by providing additional metacognitive support. One of the problems of such an additive approach is that learners, focused on completing their learning task, oftentimes do not perceive additional metacognitive support as instrumental. Instead, they tend to be resistant against alterations to their accustomed learning processes and experience deeper metacognitive processing as an extra burden.

Aiming to avoid the disadvantageous effects of indirect support, we pursued an integrated scaffolding approach by combining functionality and metacognitive assistance. More specifically, our scaffolds constitute subtools that offer functions to complete a learning task on the WWW. Those subtools are designed to induce the deployment of the six metacognitive processes upon application. Accordingly, when applying our scaffolds to conduct a web-based learning task, individuals are bound to engage in metacognitive processes.

1.2.4 Supplementing the Scaffolding Approach by Prompts

As the six metacognitive processes are considered most beneficial when carried out during a specific phase of learning (see Figure 1), this is also the case for our integrated scaffolds. However, research that has focused on upgrading learning environments by offering additional aids has suggested that learners may not be able to apply scaffolds of their own accord (Aleven et al., 2003; Oliver & Hannafin, 2000). In turn, studies on prompting have reported beneficial effects of providing learners with more intensive guidance (e.g., Bannert, 2006; Horz, Winter, & Fries, 2009; Kramarski, & Zeichner, 2001; Schwonke, Hauser, Nuckles, & Renkl, 2006). Aiming to direct learners to apply our scaffolds as intended and thereby to deploy the metacognitive processes, we also pursued the approach of administering prompts in addition to our scaffolds.

1.3 The Standard Version of ELWMS

For our previous study (Benz et al., 2010), we had applied our indirect scaffolding approach by implementing the first generation of the Firefox extension ELWMS, which is embedded on the left-hand side of the browser. ELWMS provides three major functions that induce the deployment of the six metacognitive processes during the implementation of a web-based learning task: the management of goals, the handling of resources, and the illustration of the created goal-resource structure. In the standard version of ELWMS, the scaffolds are offered in a nonembedded way (Clarebout & Elen, 2006), which lets the learners decide if, how, and when to apply them during web-based learning.

Our previous study (Benz et al., 2010) revealed that participants, who had been working with the standard version of ELWMS had deployed significantly more SRL and metacognitive processes. Specifically, they were more involved in the metacognitive processes of planning, process-regulation, and reflection. However, for the process of goal setting we found only marginal differences, and for the process of self-monitoring, we were far from finding significant differences. In our approach, since goal setting is considered to be a precondition for the deployment of the consecutive metacognitive processes, and self-monitoring is considered to be essential for finding relevant pages and relevant resources, the fact that we had not found differences between groups on the achievement posttest was attributed to the low deployment of those processes.

With regard to these findings, we created a second generation of ELWMS for the current study (see Figure 2). Besides optimizing its usability and its appearance in order to enhance learners' awareness of goals that they are currently pursuing, ELWMS was upgraded by a goal-activation function. In addition, we created a function for viewing single goals and resources. In the previous version of ELWMS, this was possible only by opening the window for editing a goal or a resource. Providing a separate function for viewing single instances enabled us to specifically log this action. This makes sense as viewing an instance can be considered a self-monitoring process, whereas editing an instance is part of process-regulation. In the following paragraphs, we describe the functions of the second generation of ELWMS in a prototypical scenario (see Table 1).

Before actually starting to learn (preaction phase), the learner is scaffolded to define goals and to plan the process of implementation. ELWMS provides a goal management function, which allows the user to specify a name, a description, and a tag, as well as the current state of completion of a goal. For example, if a learner is looking for information on ancient Rome, and more specifically on the members of the first triumvirate, he might create a goal with the name "First Triumvirate," add the description "What are the members of the first triumvirate?", organize the goal with the tag "Person," and specify the goal progress as "not started." If he is further interested in the Roman civil wars and the end of the Roman Republic, he might create several levels of subgoals. ELWMS offers the opportunity to

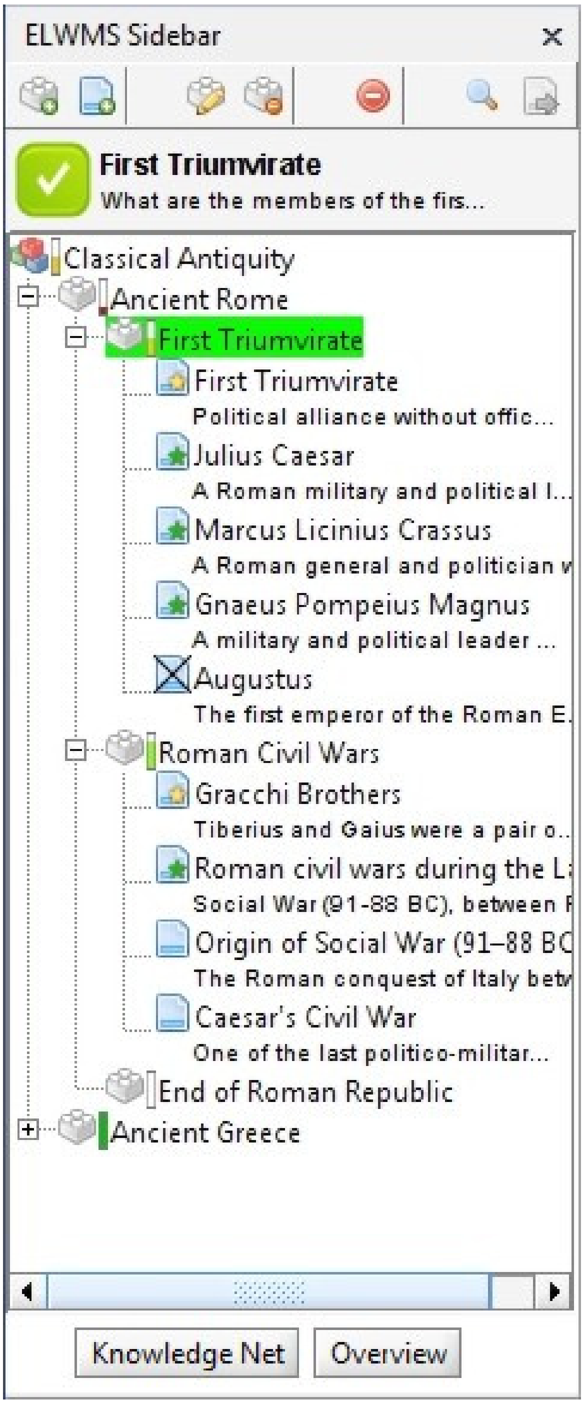


Figure 2. Screen shot of the ELWMS sidebar.

arrange goals in a hierarchical structure in order to organize the upcoming learning phase into sequences.

During the actual learning (action phase), the learner is scaffolded to monitor his learning, and if necessary, to engage in process-regulation. To stay aware of the goal that is currently pursued, the learner may apply the goal activation function. An activated goal appears in large letters at top of the sidebar. Further, ELWMS provides a function for handling resources, which allows the user to gather snippets of information by highlighting and importing words, phrases, or paragraphs from web pages. For each resource, a name and a tag may be defined, and its relevance may be judged. Snippets are automatically saved within the description of a resource and may be adapted by the user. The Uniform Resource Locator (URL) of the web page that the resource came from is automatically documented. It is further possible to bookmark whole web pages. Upon the import, resources are assigned to associated goals, and a goal-resource structure is created. This structure is displayed by three illustration tools. First, it may be viewed in detail in the sidebar on the left-hand side of the browser. In addition, we provided a knowledge net, which, similar to a mind map, presents defined goals and imported resources in a netlike overview. Its advantage lies in the illustration of tags, which are used across goal paths. We also provided an overview, which presents goals and resources in a hierarchical structure, but displays the full content of the persisted resources. It therefore offers a good basis for (re)viewing persisted content. Both knowledge net and overview are nonembedded (Clarebout & Elen, 2006) and may be accessed on demand by clicking on a button. Additionally, single goals and resources may be viewed by applying the viewing function. With an increasing number of relevant resources, the progress toward the completion of specific goals may be adapted. Whenever a goal is completed, the learner may switch the goal that is currently pursued by deactivating the current goal or by activating another goal. In the case of the need for a change in strategy during the action of learning at any time, new goals may be defined, activated goals may be switched or deactivated, and existing goals and resources may be edited, restructured, or deleted.

Toward the end of learning (postaction phase), reflection and modification processes are scaffolded. Previously defined goals and saved resources may be viewed in order to reflect on the search process and to review the results. Further, learners may reopen a web page that a resource was originally retained from. The sidebar, the knowledge net, and the overview are three illustrations of the created goal-resource structure, and also serve for elaboration purposes.

Table 1

Scaffolds Provided in the Standard Version of ELWMS: Functions and Supported SRL Processes

Phase	Preaction	Action		Postaction
Metacognitive processes	Goal setting & planning	Self-monitoring	Process regulation	Reflection & modification
Function	Defining goals[b]	Activating / deactivating goals[ab]	Defining new goals[b]	Viewing goals[b]
	Structuring goals[b]	Viewing goals[ab]	Redefining goals[b]	Viewing resources
		Assigning resources to goals[b]	Deleting goals[b]	Reopening web pages through resources
		Defining relevance of resources	Restructuring goals[b]	Viewing goal-resource structure[b]
		Viewing resources[a]	Adapting resources	Viewing knowledge net
		Defining progress toward goal completion[b]	Deleting resources	Viewing overview
		Viewing knowledge net	Restructuring resources	
		Viewing overview	Reopening web pages through resources	

Note. ELWMS = E-Learning knoWledge Management System; SRL = self-regulated learning.

[a]Function in the second generation of ELWMS.

[b]Function not available in the control version of ELWMS.

1.4 The Extended Version of ELWMS

To supplement our indirect scaffolding approach by additional prompting, in our previous study (Benz et al., 2010), we had extended ELWMS by adding two prompts. At the beginning of the preaction and postaction phases, the learners were invasively disrupted and directed to engage in goal setting, planning, and reflection. We found that the group that received prompts was more involved in reflection, but carried out fewer self-monitoring processes and created a structure of lower quality. We attributed the fact that we could not find differences between groups for the process of goal setting to failures in prerequisite processes, like the identification of knowledge gaps. However, referring to the effectiveness of the reflection prompt, we reasoned that learners tend to follow invasive and directive prompts. In addition, we supposed that the intensive deployment of self-monitoring processes of the group that did not receive prompts was responsible for the creation of a high-quality structure.

Aiming to provide learners with more intensive support to enable them to achieve a performance gain on the basis of our previous study (Benz et al., 2010), we identified nine processes that were considered to enhance the quality of the outcome in a web-based learning scenario (see Table 2). Those achievement-enhancing processes can be metaphorically referred to as bridges. If a learner does not manage to cross a bridge, the learning outcome cannot be affected in a positive manner. In the preaction phase, we consider it essential to follow a (2) *goal-oriented approach* in order to lead learning on the WWW in a beneficial direction. However, for this approach to have beneficial effects on the learning outcome, it is a precondition that learners engage in successful self-diagnosis to (1) *identify their knowledge gaps*. The transformation of those knowledge gaps then allows for the creation of—in that sense—(3) *relevant goals*. In the action phase, (4) *goal orientation during action* is inevitable for pursuing the goals that have been defined in advanced. Being aware of relevant goals enables learners to (5) *find relevant web pages* and to identify relevant information on those pages. When (6) *importing relevant information* with ELWMS, it is further essential that learners manage to (7) *assign relevant information to corresponding relevant goals*. Assuming that a learner has managed to establish a structure of high quality, it is essential that she elaborates on the created material in the postaction phase. Without (8) *learning the relevant information,* it will not be possible to (9) *retrieve the relevant information* when necessary.

For the current study, to enhance learners' deployment of the achievement-enhancing processes, we extended the standard version of ELWMS. Referring to our previous study (Benz et al., 2010), we created two prompts that invasively instructed participants to approach their learning in a goal-oriented way and to prepare for the posttest toward the end of the learning phase. In addition, we created four prompts that did not invasively disrupt the learning processes in terms of a pop-up screen, but were integrated into the ELWMS interface. When creating goals, learners were instructed to set relevant goals; and during the action of learning, they were instructed to activate their current goal, to check the relevance of the resource, as well as to check the fit of relevant information and relevant goals. We did not create a prompt that specifically fostered the finding of relevant web pages because prompting learners to reflect on each web page would have resulted in annoying disturbances of the learning process (see Table 2). The prompts are described in more detail in combination with the procedure of the study (see section 2.3).

Table 2

Achievement-Enhancing Processes and Prompts Provided in the Second Version of ELWMS

Achievement-enhancing processes	Prompted actions	Type
Identifying knowledge gaps[a]	-	-
Goal-oriented approach[b]	Prompt to set goals in preaction phase	Invasive & directive
Defining relevant goals[b]	Prompt to set relevant goals	Directive
Goal orientation during action	Prompt to activate current goal	Directive
Finding relevant web pages[c]	-	-
Importing relevant information	Prompt to check relevance of resources	Directive
Assigning relevant information to relevant goal	Prompt to check goal-resource fit	Directive
Learning relevant information[b]	Prompt to prepare for posttest in postaction phase	Invasive & directive
Retrieval of relevant information[c]	-	-

Note. ELWMS = E-Learning knoWledge Management System.

[a]Process not supported by ELWMS, but balanced across groups by the study design.

[b]Processes supported in previous version of ELWMS (Benz et al., 2010).

[c]Process not supported by ELWMS.

1.5 The Control Version of ELWMS

To be able to compare experimental and control groups on the basis of log data in the current study, we harmonized the design of the study with our method of collecting data. Aiming to simulate real-world WWW learning approaches, we created a control version of ELWMS, which offered the same functions as the standard version of ELWMS, but did not provide metacognitive support. More specifically, instead of offering a function for the management of goals, the control version provided a function for the management of folders.

Accordingly, participants in the control group had the opportunity to use folders to organize their imported resources, but neither were supported to (re)define goals, to (re)structure goals, to activate goals, to view goals, to assign resources to goals, to define the progress toward goal completion, to delete goals, or to deploy the achievement-enhancing processes (see Table 1).

1.6 Research Questions

Our research was based on our model of adaptive learner support for the enhancement of learning quality (see Figure 3). The model sketches the relation between stable and varying learner characteristics, situational parameters, and the effectiveness of applying certain types of support. In the case of a fit between the preconditions and the design of the scaffolds, the quality of the learning process can be enhanced. This may be indicated by the application of SRL strategies, or more specifically by the deployment of metacognitive, cognitive, or motivational/emotional processes. The quality of the learning process in turn has an impact on the achieved learning outcome, which may be affected at different levels of complexity.

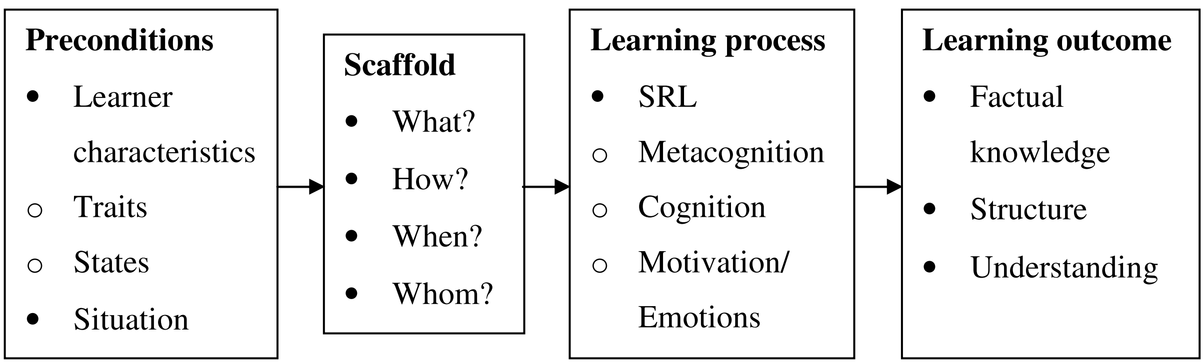


Figure 3. Model of adaptive learner support for the enhancement of learning quality.

With an elaborated study design, the current study served to provide further insight into the impact of our indirect scaffolding approach on the quality of learning on the WWW. More specifically, we examined whether optimizing the WWW as a learning environment by providing learners with scaffolds that combine functionality and metacognitive support enhances the quality of the learning process and the learning outcome of web-based learning. We also investigated whether providing learners with additional prompts that instructed them to deploy the achievement-enhancing processes would entail an increase in the quality of the learning processes and the learning outcome of web-based learning.

1.6.1 Quality of the Learning Process

To address our first research question, we hypothesized that learners who worked with the standard and the extended versions of ELWMS would carry out more achievement-enhancing processes, more SRL processes per se, and more SRL processes with relevance to the achievement test than learners who worked with the control version. To address our second research question, we hypothesized that the group that worked with the extended version of ELWMS would perform better than the group that was equipped with the standard version on the same measures described above.

1.6.2 Quality of the Learning Outcome

To address our third research question, we hypothesized that learners who worked with the standard and the extended version of ELWMS would create a structure of higher quality and would gain more factual knowledge than the group that worked with the control version. To address our fourth research question, again, we hypothesized that the group that worked with the extended version of ELWMS would perform better than the group that was equipped with the standard version on the same measures described above.

2. Method

2.1 Participants

Participants were N = 108 students from a German university. Eighty (74.1%) were students of psychology, 17 (15.7%) of education, and 11 (10.2%) of other majors that did not have a relation to *Classical Antiquity*. As an incentive for participating in the study, all students were provided with constructive feedback on their learning process. In addition, psychology students received credits, whereas students of pedagogy and other majors were given 20€ for their participation. On average, students had university for 3.1 semesters, with a mean age of 23.5 years. Seventy-eight (72.2%) participants were female and 30 (27.8%) were male. The native language of 96 (88.9%) participants was German. Thirty-seven (34.3%) had previously been enrolled in another major with five (4.6%) having graduated. Twenty-three (21.3%) had conducted a vocational apprenticeship prior to their studies.

2.2 Groups

To investigate our research questions, we randomly assigned participants to one of three conditions. Experimental group 1 (EG-Tool, n = 36) worked with the standard version of ELWMS, experimental group 2 (EG-Prompt, n = 38) with the extended version of ELWMS, and the control group (CG-Folder, n = 34) with the control version of ELWMS.

2.3 Procedure

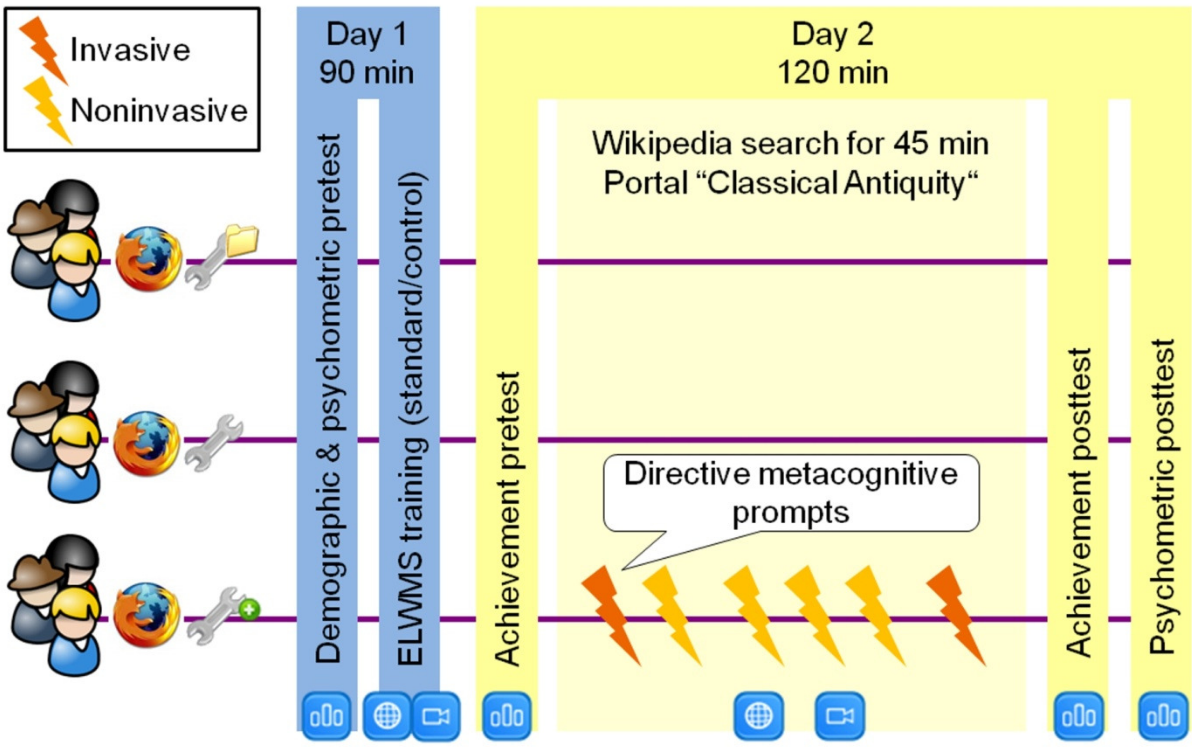


Figure 4. Overview of the design of the study.

The study was conducted in a laboratory, which was equipped with 30 up-to-date computers with internet access. To avoid close contact between learners and to keep disturbances at a low level, we utilized only a maximum of 17 computers per session and put up screens between the participants. In addition, we appointed an adequate number of supervisors to be able to calmly address the questions of participants throughout the study. Prior to each trial, we installed the required software on the computers. One third of the computers was prepared for EG-Tool, one third for EG-Prompt, and one third for CG-Folder. The study was carried out in two sessions of 90 and 120 min on two consecutive days (see Figure 4).

The purpose of the first session was to provide participants with an introduction to the study, to administer our pretests, and depending on the condition, to familiarize participants with either the control or the standard version of ELWMS. When entering the lab, participants were randomly assigned to a computer and thereby to one of the three conditions. After a short welcome, participants received a brief overview of the two sessions. Further, they were given instructions for how to fill out of the web-based questionnaires and how to handle the computers. We also asked them not to discuss this experiment with their fellow students, but to wait until the debriefing.

Next, participants worked at their own pace on a demographic pretest (see section 2.6.1) and on a psychometric pretest (see section 2.6.2) that assessed computer literacy and their context-specific SRL skills. After having completed the web-based questionnaires, each participant was automatically forwarded to a web-based training on ELWMS (see section 2.4). Once participants had reached a predefined criterion, the web-based training was completed and they were allowed to quietly leave the lab. During the first session, computer screens were recorded and all actions that were carried out during the web-based training were automatically registered in log files (see section 2.6.8).

The purpose of the second session was to have participants learn on the WWW with the standard, the extended, or the control version of ELWMS, and to provide data on the quality of the learning process and the learning outcome. To assure that participants worked with the same version of ELWMS throughout the whole study, when entering the lab, we asked them to go to the same computer that they had been using the day before. We controlled the seating arrangement with a seating plan that we had created during the first session and that was destroyed after the second session to guarantee anonymity. Again, after a short welcome, participants were provided with a brief overview of the second session and were reminded of the instructions that they had been given in the first session.

We started by administering a set of web-based questionnaires, which were presented by Microsoft Internet Explorer. After participants had reported their state motivation and state self-efficacy for conducting a web-based learning task (see section 2.6.3), they were automatically forwarded to a multiple-choice achievement test on *Classical Antiquity*. The latter served for evaluating participants' previous knowledge on the topic and for identifying each individual's specific knowledge gaps. To eliminate biased data due to lucky guesses, each question that assessed factual knowledge was supplemented by a question that assessed the certainty that the answer was given with. When constructing the test, we had assured that it was challenging enough for each participant to experience at least 10 knowledge gaps (see section 2.6.4). After having completed the achievement posttest, participants were asked again to indicate their state motivation and state self-efficacy. Further, to gain insight into students' achievement-goal orientation, we asked them to complete the Achievement Goal Questionnaire (AGQ, Elliot & McGregor, 2001) (see section 2.6.5).

After approximately 20 minutes, when participants had completed the first set of tests, they were given instructions for how to generate individual feedback for their performance on the achievement pretest. In the next step, each participant was provided with 10 multiple-choice questions that had been answered incorrectly or with uncertainty. To get a first impression of the feedback, participants were given 2 min to screen the page.

Next, participants were guided to maximize the Firefox web browser, which, depending on the condition was supplemented by the standard, the extended, or the control version of ELWMS. Firefox displayed a web page that informed all learners in all conditions that they would next be given 45 min to learn on Wikipedia and to prepare for an identical achievement posttest. Further, they were notified that their feedback on the achievement pretest would be accessible on Internet Explorer for the first 5 min of the learning period. We informed participants also that they would receive a reminder 5 min before the 45-min learning period was about to end, and that it would not be possible to use the established goal/folder-resource structure when working on the achievement posttest. In addition, EG-Prompt was informed that they were expected to follow several prompts during the learning period, and they were provided with an overview of the processes that they were going to be prompted to deploy. Before we started the learning period, participants were given the chance to raise a hand in order to call a supervisor and to quietly discuss questions.

We had added an *evaluation-menu* to the Firefox menu bar, which served to create timestamps in our log file for specific actions and to facilitate the navigation through the experiment. To begin the learning period, we instructed participants to click on *Start Learning* in the Firefox evaluation menu. In a first step, all groups were forwarded to a screen that informed them again that their feedback on the achievement pretest would be accessible for the next 5 min. The screen also provided a link to the Wikipedia portal *Classical Antiquity*, allowing each participant to start looking for information when they felt ready. In contrast to EG-Tool and CG-Folder, on this invasive screen, EG-Prompt was additionally provided with directive instructions to set their goals before starting the search. Further, in the extended version of ELWMS, on top of the window for defining a goal, we had integrated directive instructions that asked learners to assure that the goal they were creating related to their knowledge gaps. After 5 min, we instructed participants to close the feedback, which was provided by Internet Explorer, and to click on *Feedback Closed* in the Firefox evaluation menu. During learning on Wikipedia, in the extended version of ELWMS, prompts given at the top of the sidebar provided EG-Prompt with directive instructions to activate the goal that they were currently pursuing. In addition, when importing a resource, EG-Prompt received directive instructions at the top of the import window to check if the resource was relevant for the goal that it was being assigned to. Also, at the bottom of the windows for creating goals and importing resources, participants in the EG-Prompt group were instructed in a directive way to check whether the goal that was activated was still the goal that they were currently pursuing, and if not, to activate their current goal. Five min before the learning period was over, participants in all groups were informed about the time and instructed to click on *Five Minutes to Go* in the Firefox evaluation menu. For EG-Prompt, this click opened an invasive screen that displayed directive instructions to use the remaining time for reflection purposes and to prepare for the achievement posttest.

After the 45 min had passed, we asked participants to click on *End of Learning Period* in the Firefox evaluation menu. As a consequence, the ELWMS sidebar was automatically closed and participants were forwarded to the final set of web-based questionnaires, which they could work on at their preferred pace. Before and after completing the achievement posttest (see section 2.6.6), we again asked participants to indicate their state motivation and state self-efficacy. To evaluate how learners had perceived their learning on a psychometric posttest (see section 2.6.7), we also asked them whether they had employed the achievement-enhancing processes and SRL processes during the implementation of the task on Wikipedia. Once participants had completed this final set of web-based questionnaires, they were allowed to quietly leave the lab. During the second session, computer screens were recorded and relevant actions that were carried out during the 45-min learning period were automatically registered by log files (see section 2.6.8).

2.4 Computer-Based Training on ELWMS

To assure that each participant was capable of handling ELWMS in the appropriate version, we developed two versions of a web-based training using Adobe Captivate (Version 3). We created one version of the web-based training to familiarize the control group with the control version of ELWMS, and another version of the training to introduce both experimental groups to the standard version of ELWMS. Accordingly, during the web-based training, EG-Prompt was not provided with additional prompts, but learned how to handle the functions of the tool.

Both versions of the web-based training were parallel with regard to sequence and content. At first, participants watched a 30-min demo, which covered the topics *hypermedia*, *navigation in Wikipedia*, *handling of Firefox*, *management of resources in ELWMS*, *management of goals/folders in ELWMS,* and *illustration of the created goal/folder-resource structure in ELWMS*. The demo ended with a screen that instructed participants to work freely with ELWMS in the second part of the training. Aiming to ensure that all participants would sufficiently apply relevant functions of the software, we defined a criterion for the completion of the training. After participants had switched to the Wikipedia portal *Psychology*, which was already opened, they were required to perform at least the following actions: import three resources, edit one resource, open one resource, create three goals/folders, edit one goal/folder, delete one resource or one goal/folder, and restructure one resource or one goal/folder. Further, they were supposed to open the knowledge net and the overview at least once. In addition, experimental groups were asked to activate and deactivate at least one goal. If participants were unsure of how to conduct a specific action, they could return to the training at any time and watch the part of the demo that specifically referred to their question. Once participants had carried out all actions, they could complete the training by pushing the button "I have successfully completed all tasks."

Retrospective analyses showed that out of 103 participants, 92.2% had completed at least 78% of the required tasks. However, nearly all participants had practiced the application of the most central functions, such as the creation of goals and the importing of resources. We supposed that some learners did not apply some of the less relevant functions, such as restructuring an instance, because all functions had been introduced in detail during the demo, and being equipped with a decent degree of computer literary, it was not difficult to apply them. Further, 83.1% of the 71 participants in the experimental groups had practiced using the activation function.

2.5 Learning Environment

At the beginning of the learning period, participants were automatically forwarded to the Wikipedia portal *Classical Antiquity*. From there, they were allowed to navigate freely in the German Wikipedia, which provided multiple informational sources like text, photographs, as well as static and animated diagrams. As Wikipedia is not a standardized learning environment developed for experimental purposes, two disadvantages arose from its usage. First, as a web 2.0 technology, Wikipedia is a dynamic environment based on user-generated content. To assure that passages that were relevant for the achievement test were not substantially altered during the implementation of the study, before each trail, we checked the history of changes of relevant pages. Second, when applying hyperlinks to navigate through the environment, different pieces of information are not equally accessible. In our achievement test, we aimed to balance this unequal accessibility of resources (see section 2.6.4). In addition, in our web-based training (see section 2.4), we trained participants to navigate in the hypertext environment by applying the Wikipedia and the Firefox search functions. Thereby, it was possible to identify each piece of information on Wikipedia by conducting two actions. However, by utilizing Wikipedia, which is a learning environment with a great relevance for modern life, we were able to increase the external validity of our study.

2.6 Measures

To acquire the variables of interest, we collected *offline* and *online* measures. The demographic and the psychometric pretests, which were administered in the first session, were completed by all 108 participants. Two participants did not attend the second session, which resulted in 106 participants completing the achievement pretest and posttest and the AGQ (Elliot & McGregor, 2001). Further, one participant did not finish the psychometric posttest, resulting in 105 completions. All questionnaires were created using a web-based survey application called LimeSurvey (Version 1.7.2) and could be accessed by a browser through a URL. In addition, we tracked actions that were performed by participants when working with ELWMS in a log file. Due to technical difficulties, we lost data on the web-based training of five participants and on the learning task of one participant.

2.6.1 Demographic Pretest

We assessed the common demographic variables age, gender, native language, major, semester, career and/or apprenticeship before current major, and overall high school GPA. In addition, to gain insight into participants' preknowledge on the period of *Classical Antiquity*, we assessed whether previous careers and/or apprenticeships were related to the topic, as well as whether the individual had majored in History or Latin in school and how they would judge their knowledge in those subjects.

2.6.2 Psychometric Pretests

We applied two psychometric pretests. First, we assessed individuals' computer literacy using a shortened and revised version of an instrument that we had developed for our previous study (Benz et al., 2010). The questionnaire contained 12 items that constituted the three scales *general computer skills*, *web search skills*, and *experience with Firefox*.

Second, we assessed the SRL skills that participants were equipped with, applying a context-specific test that we constructed on the basis of the psychometric posttest of our previous study (Benz et al., 2010). We did not make use of one of the common instruments, like the Motivated Strategies for Learning Questionnaire (MSLQ; Pintrich, Smith, Garcia, & McKeachie, 1993) or the Volitional Components Questionnaire (VCQII; Kuhl & Fuhrmann, 1998), because, due to their assessment of learning skills in a general context, they are of limited value for web-based learning. Instead, to assess SRL skills in a web-based learning context, we applied a scenario that was related to the learning task in our study, but sufficiently dissimilar to prevent transfer affects. More specifically, when working on 96 items that were presented in three blocks referring to the three phases of learning, participants were asked to indicate on a 4-point scale whether they would carry out a specific process while learning in a hypermedia environment in order to prepare for a speech. The questionnaire was designed in such a way that two items assessed the same subprocess of SRL. On this basis, we were able to create self-report scales for the achievement-enhancing processes as well as for the SRL scales of *goal setting*, *planning*, *self-monitoring*, *process-regulation*, *motivation*, *subjective experience, reflection,* and *modification*. Accordingly, an overall scale for the implementation of metacognitive strategies that included the six metacognitive scales, and an overall scale for self-regulated learning that included all SRL scales, could be created.

2.6.3 State Measures for Motivation and Self-Efficacy

We used two items to assess the current state of participants' motivation and self-efficacy for learning on the WWW. All four items had been developed by the authors.

2.6.4 Achievement Pretest

With the goal of informing each participant about exactly 10 knowledge gaps that had been experienced on the achievement pretest, and thereby to compensate for individual differences in the ability to perform self-diagnoses, two requirements arose for our achievement test. First, we had to ensure that the test was challenging enough for participants to experience at least 10 knowledge gaps. Second, as we did not know which of the questions an individual would not be able to answer and therefore which questions would appear in the feedback, we had to assure that all the questions in the test would be of equal difficulty in order to provide all participants with feedback of equal difficulty. Whereas the first requirement implied the creation of questions that participants would not be able to answer on the basis of their previous knowledge, the second requirement entailed the construction of questions whose answers would be equally accessible on Wikipedia.

In a first step, we identified the difficulty of the 30 items of the achievement pretest that we had been using in our previous study (Benz et al., 2010). In a second step, we analyzed the accessibility of the correct answer for each question. To do so, we took into account the number of Wikipedia pages that contained information for answering a question, the number of clicks necessary to browse from the *Classical Antiquity* portal to the web page, the position of the information on a web page, as well as the number of questions that a web page provided information for. To get an impression of the relevance of the accessibility of an answer, we used the data from our previous study to correlate the frequency with which a resource had been found with three aspects of its accessibility. As we had assumed, resources were found more frequently if fewer clicks were necessary to navigate to the specific page from *Classical Antiquity* portal (r = -.146, p = .027, n = 176), and if the pertinent information was located in the introductory paragraph rather than in the body of a Wikipedia page (r = .232, p = .001, n = 176). However, as the number of answers that a Wikipedia page contained negatively affected the frequency with which a resource was found (r = -.131, p = 042, n = 176), learners seemed distracted by pages that presented a great amount of information rather than being able to profit from answers that were arranged in close proximity. However, despite those results, the true accessibility of a piece of information on Wikipedia depended on the search strategy conducted by the learner.

Taking into account the item difficulty and the variables that indicated the accessibility of a piece of information, we identified 13 questions from our previous achievement test (Benz et al., 2010) that served as the basis for constructing a revised achievement test. To enhance the probability that participants would experience at least 10 knowledge gaps, we added one further question. Of the 14 questions, two were ranking questions that asked learners to place items in the correct order. The remaining 12 questions were classical multiple-choice questions that followed the same construction scheme. We offered four possible answers that each comprised two statements. Two answers were paired and started with the same statement, but ended with a different one. Accordingly, there was one option that was designed to closely resemble the correct answer, and two further options that were related to each other, but not to the correct answer. Whereas it was not sufficient to know which of the two first statements was correct in order to choose the right answer, the identification of the correct second statement allowed for determining the correct option. To eliminate biases due to lucky guesses, we supplemented each question that assessed factual knowledge by a question that assessed the certainty that the answer was given with on a 4-point scale.

All knowledge questions on the test were very carefully constructed in correspondence to existing Wikipedia pages. We ensured that for each question, Wikipedia provided a piece of information that was sufficient for answering a question correctly by explicitly containing both statements of an answer. Also, in order to avoid confusing learners, we aligned the formulations of questions and answers to expressions that were used in Wikipedia, and ensured that Wikipedia did not offer support for any of the answers that we considered to be incorrect. Further, preventing dependency between questions, we made sure that answers of different questions were presented on Wikipedia at least 10 lines apart and that questions sufficiently differed with regard to the topic.

As Wikipedia is not a standardized learning environment developed for experimental purposes (section 2.5), it was not possible to construct a perfectly balanced test. Of the 14 questions, eight referred to Ancient Rome and six to Ancient Greece. The answers of 12 questions from the *Classical Antiquity* portal were accessible by clicking on a minimum of one hyperlink, whereas two answers required following at least two hyperlinks. Seven answers were located in an introductory paragraph and seven in the body of a Wikipedia page. Furthermore, one page provided the answers to six questions and another page the answers to two questions, whereas the remaining six answers were located on separate pages.

On the ranking questions, participants could receive proportional points for each item that was ranked in the correct spot, or one point if they had put the items in the correct order. For the multiple-choice questions, one point could be obtained for choosing the right answer, whereas all other options were not rewarded. However, the fact that a learner had answered a knowledge question correctly was only attributed to knowledge if a certainty of more than 50% had been indicated. All other combinations were considered knowledge gaps. Accordingly, on the pre-achievement test, a maximum of 14 points and a minimum of points could be achieved.

To be able to inform each participant about exactly 10 knowledge gaps before beginning the learning period, we had developed a mechanism that displayed all the questions that had been answered incorrectly or with a less-than-medium certainty on the achievement pretest. Each question was represented in a masked way by the name of a city. Without being aware of the purpose of the whole procedure, participants were guided to select 10 cities from the screen, thereby creating their individual feedback. To assure that all questions would appear on participants' feedback with equal frequency, we had created 10 versions of the achievement pretest, each of which provided the questions as well as the cities in a different order.

2.6.5 Achievement Goal Questionnaire

To gain insight into the way participants dealt with their knowledge gaps, we assessed their achievement-goal orientation. More specifically, we evaluated whether participants, when made aware of their knowledge gaps, were motivated to achieve a better performance or to avoid a poorer performance than their fellow students, as well as to gain mastery or to prevent a lack of mastery of the topic. We translated the 12 items of the AGQ (Elliot & McGregor, 2001) into German, and adjusted them to the scenario of our study.

2.6.6 Achievement Posttest

To assess the factual knowledge participants had gained by learning on Wikipedia for 45 min, we administered an achievement test that was identical to the one we had used as pretest (section 2.6.4).

2.6.7 Psychometric Posttest

To obtain retrospective self-reports on the SRL processes that individuals had deployed during the implementation of the web-based learning task, we administered the same questionnaire that we had used as a pretest (section 2.6.2). However, as we were evaluating how learners had perceived their learning, items were formulated in the past tense. Accordingly, we established self-report scales for the achievement-enhancing processes and for the SRL processes, as well as overall scales for metacognition and SRL. In addition, to evaluate the emotions that participants had experienced during the implementation of their learning task, we integrated the Positive Affect Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988) into the psychometric posttest. At the end of the test, we also asked participants if they had experienced ELWMS as a useful tool and if they would like to use it in their everyday lives.

2.6.8 Log Data

Aiming to avoid time-consuming analyses of screen recordings, we improved our methodology of collecting log data on the basis of the categories that we had applied in the video analyses of our previous study (Benz et al., 2010). We supplemented ELWMS by a Firefox add-on that automatically tracked three basic types of ELWMS activities: actions related to goals/folders, actions related to resources, and actions related to navigation. For each activity, we tracked the code of the participant who performed the activity, the condition the participant had been assigned to, and the time of occurrence. Each instance was assigned an ID, which allowed for pursuing how it was adapted throughout the period of learning. Goal/folder- and resource-related categories were specified by the features of the instance, such as the name, the description, the progress toward goal completion, the relevance of a resource, the website it had been obtained from, the tag, the goal/folder path, the time of the creation of the superior goal/folder, as well as the value of the instance before the action had been performed. If an instance was not deleted but was part of the final goal/folder-resource structure, its final state was indicated. Categories that described the navigation to web pages were supplemented by the corresponding URL, and when a Wikipedia search had been conducted, by the search term. Actions that occurred in the first 5 min were indicated to have been conducted in the preaction phase, actions that occurred between 5 to 40 min were indicated to have been carried out in the action phase, and actions that occurred in the last 5 min were indicated to have been executed in the postaction phase. Further, to be able to evaluate whether participants were pursuing a goal-oriented approach, each resource was checked to determine whether it had been imported before or after the creation of the goal that it had been assigned to. In addition, to investigate whether individuals were pursuing their activated goal, the currently activated goal was tracked, and it was determined whether the activated goal and the goal or the resource that an action was performed on were placed at the same location of the goal-resource tree. To be able to validate our log data, we also recorded computer screens using Camtasia Studio (Version 3).

Aiming to perform an additional qualitative analysis, goals, resources, and web pages were assigned a relevance rating with regard to the achievement test. To rate goals/folders and resources, we took into account the name, the description, and the goal/folder path of an instance. However, to keep work at a manageable level, we did not rate large resources of more than 1,150 characters or entire web pages. The rating 1 was assigned to a goal/folder if an instance had a clear relation to a specific question on the achievement test and represented both statements that were necessary to answer the question correctly (see section 2.6.4). In turn, resources were rated 1 if they contained both statements, and therefore explicitly identified the correct answer to the question. Goals/folders were given a 2 if they had a clear relation to a specific question, but either represented only the second statement of the correct answer or referred to another option. These instances did not fully represent a question, and as a consequence were not sufficient for finding a resource that was worthy of a relevance rating of 1. Correspondingly, resources were assigned a 2 if they contained only the second statement of an answer, and therefore, indicated the correct answer only implicitly. Goals/folders were rated 3 if they related to a specific question but hardly represented its meaning, if they were used for structuring, or if they related to more than one question. A resource was rated 3 if it contained the first statement of the correct answer, and therefore helped to exclude a minimum of one incorrect answer, but did not serve to identify the correct answer. Goals/folders and resources that were rated 4 did not relate to the achievement test. To rate entire web pages that learners had been navigating through, five categories were applied. Referring to the relevance of the resources that were provided by a page, we assigned either a 1, 2, or 3. Pages were given a 4 if their title was promising and a 5 if a visit did not make sense. For instances that had been rated 1, 2, or 3, we also identified the corresponding questions on the achievement test.

Coding was completed by the first author of this paper and a student assistant who had received intensive training. Both raters were equipped with detailed material about coding rules and assigned the relevance to each instance in consensus. To determine the reliability of the ratings, 100 goals/folders and 100 resources were rated twice. For goals/folders, we determined a Cohen's Kappa of $\kappa = .99$, and for resources of $\kappa = .90$. After the relevance rating had been completed, we fed the relevance of all instances into our database, applying the ID of an instance and the time of occurrence as mapping parameters. Among others, this procedure allowed for automatically checking whether a resource and the goal/folder that it had been assigned to referred to the same question on the achievement test. The latter relation was described as the *fit* between a goal/folder and a resource.

Besides applying our log data to evaluate participants' performances on the web-based training, we used the quantitative and qualitative *online* data to evaluate the quality of participants' learning processes during the implementation of the web-based learning task. We created scales for the achievement-enhancing processes and quantitative and qualitative scales for goal setting, planning, self-monitoring, process-regulation, and reflection, as well as overall scales for metacognition (Appendix). Further, an achievement measure was established by determining a value for the quality of the final goal/folder-resource structure that a learner had created. The value was generated by adding the number of goals/folders and resources that were part of the final structure and had received a relevance rating of 1, 2, or 3. Instances with a relevance of 1 were weighted by the factor 3, with a relevance of 2 by the factor 2, and with a relevance of 3 by the factor 1. Also the number goal/folder-resource fits that were part of the final structure were added.

2.7 Validating *Offline* and *Online* Measures

We had followed a synchronized multi-method approach to assess the quality of the learning process by *online* and *offline* measures. To investigate the validity of our measures, we correlated the established scales on the level of the achievement-enhancing processes and on the level of the metacognitive processes. In Tables 3 and 4, correlations that showed at least marginal ($\alpha = .10$) significance are presented. According to Campbell & Fiske's (1959) multitrait-multimethod approach, when examining correlations between similar and dissimilar measures, convergent validity, indicated by the correlations in the grey cells, is supposed to be high, whereas discriminant validity, represented by the correlations in the white cells, is supposed to be low. On the level of the achievement-enhancing processes, as well on the level of the metacognitive processes, we found indicators for both convergent and discriminant validity. Accordingly, *online* and *offline* measures assessed some similar and some different aspects of the process of learning on the WWW. This finding, which we had also found in our previous study (Benz et al., 2010), is a well-described pattern in the literature (Perry & Winne, 2006; Veenman, Prins, & Verheij, 2003; Winne & Jamieson-Noel, 2002). It also underlines the importance and appropriateness of our multi-method approach.

Table 3

Correlations of the Achievement-Enhancing Processes Determined Through Self-Reports and Log Data Analyses

	Self-report scales						
Scales from log data analyses	Goa	Drg	Goda	Frwp	Iri	Aritrg	Lri
Goa	$.18^{*}$	$.18^{*}$	$.17^{*}$	$.14^{\#}$		$.23^{*}$	
Drg	$.13^{\#}$	$.15^{\#}$		$.17^{*}$			$.16^{*}$
Goda	$.17^{*}$			$.14^{\#}$			
Frwp							
Iri	$.18^{*}$	$.18^{*}$	$.23^{**}$	$.19^{*}$	$.22^{*}$		
Aritrg	$.28^{**}$	$.24^{**}$	$.29^{**}$	$.21^{*}$	$.27^{**}$	$.18^{*}$	$.23^{**}$
Lri							

Note. N = 104; Goa = goal oriented approach; Drg = defining relevant goals; Goda = goal orientation during; Frwp = finding relevant web pages; Iri = importing relevant information; Aritrg = assigning relevant information to relevant goal; Lri = Learning relevant information. $^{\#}p < .10$, one tailed. $^{*}p < .05$, one-tailed. $^{**}p < .01$, one-tailed.

Table 4

Correlations of the Metacognitive Self-Report Scales and Quantitative and Qualitative Metacognitive Scales Attained from Log File Analyses

	Self-report scales						
Scales from log data analyses	SRL	Metacog	Goals	Plan	Selfm	Procreg	Refl
Quantitative							
Metacog							.13#
Goals	.16*	.13#	.20*				.17*
Plan							
Selfm							.14#
Procreg							
Refl				-.13#			
Qualitative							
Metacog	.14#	.14#				.13#	.21*
Goals	.17*	.16#	.18*	.16#	.14#		.18*
Plan							
Selfm							.17*
Procreg							
Refl							

Note. N = 104; SRL = self-regulated learning; Metacog = metacognition; Goals = goal setting; Plan = planning; Selfm = self-monitoring; Procreg = process-regulation; Refl = reflection.
#p < .10, one tailed. *p < .05, one-tailed.

3. Results

3.1 Quality of the Learning Process

To investigate our first two research questions, we carried out four sets of ANOVAs. We applied the three conditions EG-Tool, EG-Prompt, and CG-Folder as levels of the independent variable, and the achievement-enhancing and SRL processes, attained through self-reports and log data, as dependent variables. In a first contrast, we compared the two groups that had received metacognitive scaffolding to the control group. In a second contrast, we analyzed differences between learners who worked with the standard and the extended versions of ELWMS.

3.1.1 Achievement-Enhancing Processes

3.1.1.1 Log data. Table 5 presents the results of the first set of ANOVAs, which applied the scales of the achievement-enhancing processes, established on the basis of log data, as dependent variables. For the scales *goal-oriented approach* and *defining relevant goals*, we found highly significant differences between groups with both contrasts being significant. For the scale *assigning relevant information to relevant goal*, we also found significant differences between groups with the contrast between the two experimental groups being significant.

In an additional analysis, we further found that EG-Prompt (M = 5.14, SD = 5.93, n = 36) had activated significantly more relevant goals during the action phase than EG-Tool (M = 1.29, SD = 2.63, n = 35), $t(48.55) = -3.56$, $p = .001$, $\eta^2 = .21$.

Table 5

Differences Between Groups in the Implementation of the Achievement-Enhancing Processes Based on Log Data

Scales from log data analyses	Conditions EG-Tl (n = 35) *M* (*SD*)	EG-Ppt (n = 36) *M* (*SD*)	CG-Fo (n = 34) *M* (*SD*)	*F*	Planned comparisons: EG-Tool & EG-Ppt vs. CG-Fo[a]	EG-Tool vs. EG-Ppt[a]
Goa	5.97 (3.82)	8.78 (3.62)	5.59 (5.00)	$F(2, 102) = 6.12$, $p = .003$, $\eta^2 = .11$	$t(102) = 1.85$, $p = .035$, $\eta^2 = .03$	$t(102) = 3.18$, $p = .001$, $\eta^2 = .09$
Drg	4.06 (4.06)	5.64 (3.31)	3.26 (3.84)	$F(2, 102) = 3.94$, $p = .022$, $\eta^2 = .07$	$t(102) = 2.10$, $p = .019$, $\eta^2 = .04$	$t(102) = 1.85$, $p = .034$, $\eta^2 = .03$
Goda	2.43 (3.77)	2.61 (4.14)	1.79 (3.44)	$F(2, 102) = 0.44$, $p = .644$, $\eta^2 = .01$	$t(102) = 0.92$, $p = .181$, $\eta^2 = .01$	$t(102) = 0.20$, $p = .420$, $\eta^2 < .01$
Frwp	24.74 (11.25)	28.25 (12.62)	25.00 (9.34)	$F(2, 102) = 1.06$, $p = .342$, $\eta^2 = .02$	$t(102) = -0.64$, $p = .262$, $\eta^2 < .01$	$t(102) = -1.32$, $p = .095$, $\eta^2 = .02$

	Conditions				Planned comparisons	
	EG-Tl (*n* = 35)	EG-Ppt (*n* = 36)	CG-Fo (*n* = 34)			
Scales from log data analyses	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*F*	EG-Tool & EG-Ppt vs. CG-Fo[a]	EG-Tool vs. EG-Ppt[a]
Iri	4.11 (2.63)	4.69 (2.48)	4.26 (3.23)	*F*(2, 102) = 0.41, *p* = .663, η^2 = .01	*t*(102) = -0.24, *p* = .406, η^2 < .01	*t*(102) = -0.88, *p* = .192, η^2 = .01
Aritrg	4.69 (3.91)	6.81 (4.72)	4.56 (4.05)	*F*(2, 102) = 3.13, *p* = .048, η^2 = .06	*t*(102) = 2.37, *p* = .092, η^2 = .05	*t*(102) = 2.10, *p* = .014, η^2 = .04
Lri	0.46 (0.82)	0.94 (1.88)	0.65 (1.28)	*F*(2, 102) = 1.09, *p* = .339, η^2 = .02	*t*(68.2) = -0.19, *p* = .424, η^2 = .00	*t*(48.0) = -1.42, *p* = .081, η^2 = .04

Note. EG-Tl = experimental group working with the standard version of ELWMS; EG-Ppt = experimental group working with extended version of ELWMS; CG-Fo = control group working with control version of ELWMS; Goa = goal oriented approach; Drg = defining relevant goals; Goda = goal orientation during action; Frwp = finding relevant web pages; Iri = importing relevant information; Aritrg = assigning relevant information to relevant goal; Lri = Learning relevant information.

[a]one-tailed.

3.1.1.2 Self-reports. The second set of ANOVAs, which applied the self-report scales of the achievement-enhancing processes as dependent variables, did not reveal any significant results.

3.1.2 SRL Processes

3.1.2.1 Log data. Table 6 presents the results of the third set of ANOVAs, which applied the quantitative and qualitative metacognitive scales that were established on the basis of log data as dependent variables. For the quantitative scales, we found highly significant differences between groups for the overall scale of *metacognition*, and for the scales of *goal setting*, *planning*, and *reflection*, with both contrasts being significant. For qualitative scales, we found the same pattern. However, participants in the experimental groups only marginally carried out more reflection processes with relevance to the achievement test than participants in the control group.

Table 6

Differences Between Groups in the Implementation of the Metacognitive Processes Based on Log Data

	Conditions				Planned comparisons	
	EG-Tl (*n* = 35)	EG-Ppt (*n* = 36)	CG-Fo (*n* = 34)			
Scales from log data analyses	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*F*	EG-Tool & EG-Ppt vs. CG-Fo[a]	EG-Tool vs. EG-Ppt[a]
Quantitative						
Metacog	75.20 (29.43)	92.58 (30.79)	68.50 (27.14)	$F(2, 102) = 6.39$, $p = .002$, $\eta^2 = .11$	$t(102) = -2.53$, $p = .007$, $\eta^2 = .06$	$t(102) = -2.51$, $p = .007$, $\eta^2 = .06$
Goals	7.09 (4.91)	10.14 (4.51)	6.41 (5.76)	$F(2, 102) = 5.42$, $p = .006$, $\eta^2 = .10$	$t(102) = 2.08$, $p = .020$, $\eta^2 = .04$	$t(102) = 2.54$, $p = .007$, $\eta^2 = .06$
Plan	-4.54 (10,.57)	2.39 (11.78)	-8.06 (9.20)	$F(2, 102) = 8.86$, $p < .001$, $\eta^2 = .15$	$t(102) = -3.16$, $p = .001$, $\eta^2 = .09$	$t(102) = -2.76$, $p = .004$, $\eta^2 = .07$

	Conditions				Planned comparisons	
	EG-Tl (*n* = 35)	EG-Ppt (*n* = 36)	CG-Fo (*n* = 34)			
Scales from log data analyses	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*F*	EG-Tool & EG-Ppt vs. CG-Fo[a]	EG-Tool vs. EG-Ppt[a]
Selfm	46.06 (16.31)	49.00 (16.44)	42.79 (14.51)	$F(2, 102) = 1.35$, $p = .264$, $\eta^2 = .03$	$t(102) = -1.44$, $p = .077$, $\eta^2 = .02$	$t(102) = -0.79$, $p = .217$, $\eta^2 = .01$
Procreg	26.46 (16.92)	28.31 (17.49)	28.21 (17.61)	$F(2, 102) = 0.13$, $p = .882$, $\eta^2 < .01$	$t(102) = -0.23$, $p = .410$, $\eta^2 < .01$	$t(102) = 0.45$, $p = .327$, $\eta^2 < .01$
Refl	0.14 (3.67)	2.75 (4.56)	-0.85 (3.30)	$F(2, 102) = 8.05$, $p = .001$, $\eta^2 = .14$	$t(79.1) = -3.07$, $p = .002$, $\eta^2 = .11$	$t(66.7) = -2.66$, $p = .005$, $\eta^2 = .10$

	Conditions				Planned comparisons	
	EG-Tl (n = 35)	EG-Ppt (n = 36)	CG-Fo (n = 34)		EG-Tool &	
Scales from log data analyses	M (SD)	M (SD)	M (SD)	F	EG-Ppt vs. CG-Fo[a]	EG-Tool vs. EG-Ppt[a]
Qualitative						
Metacog	43.06 (22.50)	55.36 (22.26)	38.12 (19.80)	F(2, 102) = 5.97, p = .004, η^2 = .10	t(102) = -2.46, p = .008, η^2 = .06	t(102) = -2.40, p = .009, η^2 = .05
Goals	4.83 (4.13)	6.69 (4.29)	3.94 (4.10)	F(2, 102) = 3.99, p = .021, η^2 = .07	t(102) = -2.09, p = .020, η^2 = .04	t(102) = -1.88, p = .032, η^2 = .03
Plan	-1.57 (9.14)	1.94 (10.07)	-4.32 (5.70)	F(2, 102) = 4.73, p = .011, η^2 = .08	t(102) = -2.53, p = .007, η^2 = .06	t(102) = -1.73, p = .043, η^2 = .03
Selfm	24.40 (11.69)	27.31 (10.11)	22.50 (11.08)	F(2, 102) = 1.71, p = .186, η^2 = .03	t(102) = -1.47, p = .073, η^2 = .02	t(102) = -1.12, p = .134, η^2 = .01

	Conditions				Planned comparisons	
	EG-Tl (n = 35)	EG-Ppt (n = 36)	CG-Fo (n = 34)			
Scales from log data analyses	M (SD)	M (SD)	M (SD)	F	EG-Tool & EG-Ppt vs. CG-Fo[a]	EG-Tool vs. EG-Ppt[a]
Procreg	15.60 (12.21)	18.53 (12.73)	16.35 (10.65)	$F(2, 102) = 0.58$, $p = .562$, $\eta^2 = .01$	$t(102) = 0.29$, $p = .388$, $\eta^2 < .01$	$t(102) = 1.04$, $p = .152$, $\eta^2 = .01$
Refl	-0.20 (1.88)	0.89 (1.94)	-0.35 (2.32)	$F(2, 102) = 3.86$, $p = .024$, $\eta^2 = .07$	$t(102) = -1.36$, $p = .053$, $\eta^2 = .03$	$t(102) = -2.24$, $p = .014$, $\eta^2 = .05$

Note. EG-Tl = experimental group working with the standard version of ELWMS; EG-Ppt = experimental group working with extended version of ELWMS; CG-Fo = control group working with control version of ELWMS; SRL = self-regulated learning; Metacog = metacognition; Goals = goal setting; Plan = planning; Selfm = self-monitoring; Procreg = process-regulation; Refl = reflection.

[a]one-tailed.

3.1.2.2 Self-reports. The fourth set of ANOVAs, which applied self-report SRL scales as dependent variables, did not reveal any significant results.

3.2 Quality of the Learning Outcome

To investigate our third and fourth research questions, we contrasted the two experimental groups against the control group and the two experimental groups against each other. The quality of the established goal/folder-resource structure and the performance gain from the achievement pretest to the achievement posttest were applied as dependent variables.

3.2.1 Goal/Folder-Resource Structure

Applying the three conditions as levels of the independent variable and the quality of the established goal/folder-resource structure as the dependent variable, we conducted an ANOVA. As presented in Table 7, we did not find significant differences in the quality of the structure that participants had created to prepare for the achievement posttest. However, we found that the final structures of the experimental groups contained significantly more goals with a relevance of 1 and 2, but not more resources with a relevance of 1 and 2 than the final structures of the control group. This was also the case for participants who received intensive prompting during the period of learning compared to participants who worked with the standard version of ELWMS.

Table 7

Differences Between Groups in the Quality of the Final Goal/Folder-Resource Structure Based on Log Data

	Conditions				Planned comparisons	
	EG-Tl (n = 35)	EG-Ppt (n = 36)	CG-Fo (n = 34)			
Scales from log data analyses	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*F*	EG-Tool & EG-Ppt vs. CG-Fo[a]	EG-Tool vs. EG-Ppt[a]
Structure	33.71 (18.13)	39.06 (16.52)	32.26 (20.27)	$F(2, 102) = 1.34$, $p = .265$, $\eta^2 = .03$	$t(102) = 1.08$, $p = .142$, $\eta^2 = .01$	$t(102) = -1.23$, $p = .112$, $\eta^2 = .01$
Goals relevance 1&2	4.91 (3.84)	6.56 (3.88)	4.18 (4.22)	$F(2, 102) = 3.30$, $p = .041$, $\eta^2 = .06$	$t(102) = 1.88$, $p = .032$, $\eta^2 = .03$	$t(102) = -1.74$, $p = .043$, $\eta^2 = .03$
Resources relevance 1&2	4.63 (2.46)	4.75 (2.39)	4.85 (3.16)	$F(2, 102) = 0.06$, $p = .942$, $\eta^2 < .01$	$t(102) = 0.29$, $p = .386$, $\eta^2 < .01$	$t(102) = -0.19$, $p = .425$, $\eta^2 < .01$

Note. EG-Tl = experimental group working with the standard version of ELWMS; EG-Ppt = experimental group working with extended version of ELWMS; CG-Fo = control group working with control version of ELWMS.

[a]one-tailed.

3.2.2 Achievement Test

Differences between groups on the achievement posttest were analyzed by applying a one-factor ANCOVA with the three conditions EG-Tool, EG-Prompt, and CG-Folder as levels of the independent variable and achievement on the pretest (EG-Tool: $M = 0.37$, $SD = 0.81$, $n = 35$; EG-Prompt: $M = 0.65$, $SD = 1.42$, $n = 37$; CG-Folder: $M = 0.35$, $SD = 0.69$, $n = 34$) as the covariate. On the achievement posttest, EG-Tool attained a mean of 5.31 ($SD = 2.87$) answers that were correct and given with a certainty of more than 50%, whereas EG-Prompt and CG-Folder reached medium scores of 5.92 ($SD = 3.17$), and 6.09 ($SD = 3.19$). The covariate, achievement on the pretest, was significantly related to achievement on the posttest, $F(1,102) = 28.19$, $p < .001$, partial $\eta^2 = .32$. However, there was no significant effect of achievement on the posttest after controlling for achievement on the pretest, $F(2,102) = 0.78$, $p = .460$, partial $\eta^2 = .02$.

4. Discussion

4.1 Quality of the Learning Process

For our first research question, we aimed to investigate whether our scaffolding approach of optimizing the WWW as a learning environment by providing learners with scaffolds that combine functionality and metacognitive support enhances the quality of the learning process. Our second research question aimed to investigate the impact of additional intensive prompting of the six achievement-enhancing processes. Pursuing a synchronized multi-method approach, we evaluated the quality of the learning process by assessing the implementation of the achievement-enhancing processes and the SRL processes through self-reports and quantitative and qualitative log data.

4.1.1 Experimental Conditions Versus CG-Folder

In contrast to our first hypothesis, based on self-reports, on the level of the SRL processes, participants in the two scaffolding conditions did not report carrying out significantly more metacognitive and motivational processes, nor did they report experiencing the learning process significantly differently while working on the task than the control group. Also, based on self-reports, for the achievement-enhancing processes, no differences were found between the groups.

However, in line with our first hypothesis, based on quantitative and qualitative log data, learners who received scaffolding deployed significantly more metacognitive processes per se, significantly more metacognitive processes with relevance to the achievement test, and significantly more achievement-enhancing processes. In the preaction phase, the ELWMS goal-setting and planning support enhanced not only the learners' involvement in goal-setting and planning processes per se, but also in goal-setting and planning processes that related to their knowledge gaps. More specifically, learners approached the web-based learning task in a goal-oriented way and utilized the time while they still had access to the feedback on their knowledge gaps to define goals that served to enhance their achievement. Considering the results of our previous study (Benz et al., 2010), in which we found only marginal differences for the process of goal setting and significant differences for the process of planning between groups that received indirect scaffolding and groups that did not, the current findings show that revising our study design by providing individuals with feedback on their knowledge gaps enhanced the effectiveness of the ELWMS goal-setting support. These results also provide evidence that learners' identification of their knowledge gaps can be considered a prerequisite process for the successful deployment of goal-setting activities for web-based learning.

In the action phase, based on quantitative and qualitative log data, ELWMS self-monitoring and process-regulation support did not enhance the activation of self-monitoring and process-regulation processes. As a consequence, during the ongoing learning process, learners failed to detect ineffective and inefficient processes, and were unable to reestablish beneficial learning processes while trying to find relevant web pages and resources. Regarding the results of our previous study (Benz et al., 2010), in which we had been far from finding significant differences for the process of self-monitoring, we had upgraded ELWMS with a goal-activation function. We have to acknowledge that this approach did not enhance learners' awareness of the goal that they were currently pursuing. On the one hand, this finding underlines the difficulty of effectively assisting individuals during the ongoing learning process in an indirect way. On the other hand, as we calculated group comparisons, the effectiveness of the self-monitoring support was determined in relation to learners who worked with the control version of ELWMS. During the implementation of the learning task, those participants had access to the knowledge net and the overview, which are two elements that also may enhance self-monitoring. As a consequence, it is possible that we did not find the self-monitoring support to be effective because the control group also profited from functions that fostered self-monitoring. This matter is the trade off for having harmonized the design of our study with our method of collecting data by equipping the control group with a downgraded version of ELWMS, which we did in order to raise comparable log data for experimental and control groups. Furthermore, in contrast to our previous study, our process-regulation support did not prove to be effective. This finding as well may be attributed either to the approximation of experimental and control conditions or to the fact that this result in our previous study was based on self-reports of participants.

In the postaction phase, based on quantitative and qualitative log data, the ELWMS reflection support enhanced the participants' involvement in reflection processes per se, but not in reflection processes that related to learners' knowledge gaps. Accordingly, in line with our previous study (Benz et al., 2010), ELWMS is an effective tool for inducing reflection processes. However, for reflection processes to have the potential to enhance achievement, the successful completion of prerequisite processes is a precondition for becoming equipped with information that can serve to close a knowledge gap. Since learners did not carry out self-monitoring and process-regulation processes during the action phase, this precondition was apparently not met.

In sum, when providing learners with information on their knowledge gaps, ELWMS effectively supports individuals in approaching learning on the WWW in a high-quality way. However, during the actual learning process, it fails to help learners to stay on this promising path. As a consequence, even though it effectively fosters reflection processes, those reflection processes do not relate to learners' knowledge gaps. Altogether, in line with our previous study (Benz et al., 2010), our indirect scaffolding approach serves to enhance the quality of the learning process during web-based learning. These results are consistent with other studies that also have found beneficial effects of process support (Azevedo, Guthrie, & Seibert, 2004; Benz et al., 2010; Brush & Saye, 2001; Greene & Land, 2000).

4.1.2 EG-Tool Versus EG-Prompt

In contrast to our second hypothesis, based on self-reports, we did not find learners who received additional intensive prompting to have deployed more achievement-enhancing processes than learners who were free to apply the scaffolds of their own accord. Also, based on self-reports, no differences could be perceived with regard to SRL processes.

However, in line with our second hypothesis, based on quantitative and qualitative log data, learners who were intensively prompted to deploy the achievement-enhancing processes in addition to being equipped with our scaffolds deployed significantly more metacognitive processes per se, significantly more metacognitive processes with relevance to the achievement test, and significantly more achievement-enhancing processes. In the preaction phase, the invasive and directive prompt to set goals before starting the search effectively supported approaching the learning task in a goal-oriented way and thereby helped participants to utilize the feedback on the knowledge gaps while it was still accessible. In addition, the noninvasive and directive prompt that appeared at the top of the window for defining a goal and that instructed participants to check whether the goal that they were creating related to a knowledge gap effectively fostered the creation of goals that served to enhance achievement. As a result, in the preaction phase, learners were not only more involved in goal-setting and planning processes per se, but also in goal-setting and planning processes that were relevant for a gain in performance. Considering the results of our previous study (Benz et al., 2010), in which we had not found differences between experimental groups for the processes of goal setting, these findings show that revising our study design by providing individuals with feedback on their knowledge gaps improved the effectiveness of our goal-setting prompts. Again, we found evidence that learners' successful identification of their knowledge gaps is a precondition for successful goal-setting activities in web-based learning.

In the action phase, based on quantitative and qualitative log data, even though the noninvasive and directive prompt to activate one's current goal effectively supported the activation of goals that related to one's knowledge gaps, it did not sufficiently support goal orientation during learning on the WWW. As a consequence, learners did not manage to identify significantly more web pages that contained relevant information. In addition, the noninvasive and directive prompt that appeared at the top of the window for importing information and that instructed participants to check whether a resource was relevant for the goal that it was being assigned to did not effectively foster the import of resources that served to enhance achievement. However, it effectively supported the matching of resources and goals that related to the same knowledge gap. As a result, in the action phase, learners did not carry out more self-monitoring and process-regulation activities. Regarding the results of our previous study (Benz et al., 2010), in which learners who had received prompts deployed fewer self-monitoring processes than learners who did not, in the current study, we applied a more intensive prompting of self-monitoring processes. Our findings indicated that we triggered the group that received additional prompting to engage in self-monitoring in a manner that was equal to the group that was free to apply the scaffolds of their own accord. Accordingly, in line with the results of our previous study, we found evidence that supplementing our scaffolds by prompts in the action phase was not beneficial. Again, these results also underline the difficulty of effectively assisting individuals during the ongoing learning process in an indirect way.

In the postaction phase, based on quantitative and qualitative log data, the invasive and directive prompt to use the remaining time for reflection and preparation purposes did not effectively support the learning of information that served to enhance achievement. However, in line with our previous study (Benz et al., 2010), it effectively supported learners' involvement in reflection processes per se as well as reflection processes that related to their knowledge gaps. Considering that all groups were informed about the remaining time 5 min before the end of the learning period, and hence received an invasive but a nondirective prompt, the effectiveness of our reflection prompt is underlined. As a consequence, again, the lack of effectiveness of the reflection prompt seems to be due to failures in prerequisite processes in the action phase.

In sum, when providing learners with information on their knowledge gaps, the two prompts that we administered in the preaction phase are an additional help for individuals to approach learning on the WWW in a high-quality way. However, the noninvasive and directive prompts to pursue previously set goals only partly aid learners to stay on this promising path. As a consequence, even though the invasive and directive reflection prompt appears to be effective, failures in prerequisite processes seem to undermine its influence on achievement. Altogether, supplementing the scaffolds that are based on our indirect approach by intensive prompting enhances the quality of the learning process during web-based learning. These results are consistent with other studies that also have found beneficial effects of prompting (Aleven et al., 2003; Bannert, 2006; Horz et al., 2009; Kramarski, & Zeichner, 2001; Oliver & Hannafin, 2000; Schwonke et al., 2006).

4.2 Quality of the Learning Outcome

For our third research question, we aimed to investigate whether our scaffolding approach of optimizing the WWW as a learning environment by providing learners with scaffolds that combine functionality and metacognitive support enhances the quality of the learning outcome. Our fourth research question was aimed at investigating the impact of additional intensive prompting of the six achievement-enhancing processes. The quality of the learning outcome was evaluated on two levels. Based on log data, we created a value for the established goal/folder-resource structure. In addition, we assessed factual knowledge on the period of *Classical Antiquity* with a multiple-choice achievement pretest and posttest.

4.2.1 Experimental Conditions Versus CG-Folder

In contrast to our third hypothesis, learners of all conditions equally profited from the 45-min learning period by establishing a high-quality goal/folder-resource structure and by gaining factual knowledge on the period of *Classical Antiquity*. Nevertheless, further analyses revealed that the structures of learners who received scaffolding contained a larger number of goals with relevance to their knowledge gaps.

Our results indicate that the three versions of ELWMS are equally functional for conducting a learning task on the WWW. Considering that the two scaffolding conditions deployed learning processes of higher quality during learning on the WWW (see section 4.1.1), which, based on SRL research is supposed to entail a learning outcome of higher quality (Azevedo, Guthrie et al., 2004; Zimmerman & Schunk, 2001), at a first glance these results appear counterintuitive. However, SRL interventions with differential effects on academic achievement are a well-known phenomenon (Benz & Schmitz, 2009). In our approach (see section 1.2.1), we supposed that to achieve a gain in performance, metacognitive processes have to be carried out in a high quality way during all three phases of learning. In other words, the implementation of single metacognitive processes is an improvement in the quality of the learning process, but does not affect the quality of the learning outcome. Based on this assumption, regarding the results of our previous study in which ELWMS had not effectively supported goal setting and self-monitoring (Benz et al., 2010), we had revised our study design and the ELWMS software to enhance those processes in particular. However, as sketched in section 4.1.1, even though the second generation of ELWMS effectively supports learners to approach the implementation of a web-based task in a high quality way, during the ongoing learning process, it does not sufficiently aid learners in engaging in self-monitoring and process-regulation. As a result, learners do not manage to identify enough relevant pages, do not import sufficient relevant resources, and fail to create a goal/folder-resource structure that serves to enhance their performance on the achievement posttest. As a consequence, even though ELWMS effectively enhances reflection processes toward the end of learning, they miss the opportunity to achieve beneficial effects from its use. However, again, we have to acknowledge that we determined the quality of the learning outcome on the basis of group comparisons. Equipping the control group with a downgraded version of ELWMS to be able to attain comparable log data for all groups entailed the creation of a very strong control condition that provided learners with more functions than standard software, such as Firefox. In doing so, we might have undermined our effects.

In sum, our indirect scaffolding approach did not serve to enhance the quality of the learning outcome of web-based learning. However, we suppose that besides having implemented a strong control group, these findings are due to the ineffectiveness of the ELWMS self-monitoring and process-regulation support during the action phase.

4.2.2 EG-Tool Versus EG-Prompt

In contrast to our fourth hypothesis, learners of both experimental groups profited equally from the learning period by establishing a high-quality goal/folder-resource structure and by gaining factual knowledge on the period of *Classical Antiquity*. Nevertheless, further analyses revealed that the structures of learners who received prompting contained a larger number of goals with relevance to their knowledge gaps.

Reflecting on the results of our previous study (Benz et al., 2010), we had identified nine processes that were considered to enhance the quality of the outcome in web-based learning. We metaphorically referred to those achievement-enhancing processes as bridges, which have to be crossed to affect the outcome in a positive manner. To help learners to deploy six of those processes, we had extended the second generation of ELWMS by intensive prompting. However, as sketched in section 4.1.2, even though learners of EG-Prompt more often followed a goal-oriented approach and created more relevant goals than EG-Tool, they did not pursue their goals more intensively. As a consequence, they were unable to find more pages and more resources that served to enhance their performance. Accordingly, even though they were better at matching resources with goals that related to the same knowledge gap, they failed to create a goal/folder-resource structure of higher quality and did not learn more relevant information.

In sum, supplementing the scaffolds that are based on our indirect approach by intensive prompting did not serve to enhance the quality of the outcome of web-based learning. However, we suppose that those findings are due to the ineffectiveness of the prompts during the action phase. Metaphorically speaking, during the ongoing learning process, individuals failed to cross the bridge that would allow them to achieve a gain in performance.

4.3 Limitations

In this study we have shown that the quality of learning on the WWW can be enhanced by providing learners with indirect support as implemented by our EWLMS software. However, there are some open questions with respect to our measures. First, when evaluating differences in the quality of the learning process between groups, even though objective *online* data indicated that learners had deployed different processes, we did not find significant differences on the basis of learners' retrospective self-reports. Explaining this counterintuitive result is somewhat speculative. Besides having to consider biases due to subjective impressions of learners and memory effects, we have to question whether our instrument was not adequately sensitive for assessing slight variations in the learning process. However, to be able to collect comparable log data for all conditions, we equipped the control group with a downgraded version of ELWMS that was supposed to imitate standard software for web-based learning. Since participants in all conditions worked with a version of ELWMS, perhaps our conditions were too similar for learners to experience their learning in different ways. The assumption that the value of self-reports may decline with an increasing similarity of conditions is also supported by our previous study (Benz et al., 2010), in which self-reports obtained through adapted MSLQ (Pintrich et al., 1993) and VCQ II (Kuhl & Fuhrmann, 1998) scales, in contrast to objective *online* measures, did not provide evidence for groups that worked with different versions of ELWMS.

Second, we had revised our methodology of collecting log data, aiming to assess all actions that we had identified to be meaningful indicators of the quality of the learning process. However, due to technical reasons, we were not able to log all actions that we considered to be relevant. For example, we did not track the movement of the mouse pointer or scrolling. In addition, we could track only the viewing of a single instance if a participant applied the ELWMS viewing function. If, for self-monitoring or reflection purposes, a learner used the knowledge net or the overview, we could not relate the action to a single instance; and if a learner screened the structure that was presented on the ELWMS sidebar, we could not track the action at all. Accordingly, when composing our scales, we had to rely on a sample of relevant overt actions to infer the quality of the learning process.

Third, when correlating self-reports on the learning process with quantitative and qualitative scales that were obtained through log data analyses, we found indicators for convergent and for discriminant validity (Campbell & Fiske, 1959). Such diverging results between *offline* and *online* measures are a well-known phenomenon in current research on learning (Perry & Winne, 2006; Veenman et al., 2003; Winne & Jamieson-Noel, 2002). However, as each method has its advantages and shortcomings, we should not interpret moderate correspondence between measures as a drawback. Rather, we should be aware of the fact that each method is able to provide us with a specific piece of the "actual learning process" puzzle, and continue to pursue multi-method approaches.

4.4 Future Perspectives

Future research will have to deal with two major topics: (a) the design of support for web-based learning, and (b) the assessment of the quality of web-based learning as a precondition for evaluating the effectiveness of the support. With regard to the questions of *what* to support, *how* and *when* to support it, and to *whom* the support should be administered (Pea, 2004), this study provided further evidence for the effectiveness of our concept of designing scaffolds. Optimizing the WWW as a learning environment by integrating scaffolds, which offer functions that can be used to complete a task and at the same time support the six metacognitive processes, fostered the deployment of metacognitive processes, and more specifically of the processes of goal setting, planning, and reflection. However, in line with the results of our previous study (Benz et al., 2010), future research will have to deal with the question of why the application of scaffolds that are based on this powerful concept does not result in significant gains in factual knowledge.

Considering that following a goal-oriented approach is a precondition for deploying other processes of self-regulated learning and to achieve a performance gain, it is one of the major achievements of this study that we managed to create effective indirect support in the preaction phase. Referring to our previous study (Benz et al., 2010) in which ELWMS had effectively supported planning processes, but had not been effective for the process of goal setting, in the current study, we provided learners with feedback on their knowledge gaps. When equipped with this information, learners managed to successfully deploy goal setting and planning processes in the preaction phase. Accordingly, as this study indicated that learners seemed unable to perform successful self-diagnoses, future research that aims to support goal setting should support the identification of knowledge gaps as well.

Even though ELWMS can be considered to be an effective tool for helping learners to approach learning on the WWW in a high-quality way, since we could not find effects on self-monitoring and process-regulation during the action phase, it fails to help learners to stay on this promising path. In line with the results of our previous study (Benz et al., 2010), our results underline the difficulty of effectively assisting individuals during the ongoing learning process. However, as goal orientation during web-based learning can be considered a precondition for achieving a performance gain, our results also point out its key role in helping participants find relevant pages and resources. It will be the central task of future research to investigate ways to effectively support metacognitive processes during the action phase of web-based learning to help learners create a better goal/folder-resource structure. As it seems to be an ambitious approach to support self-monitoring and process-regulation in only an indirect manner, we suggest adding a direct training approach, such as a web-based training, for example. Informing participants about why they should deploy SRL processes during learning on the WWW and training them in how to do it (Brown, Campione, & Day, 1981) should help them overcome mediation deficiencies (Reese, 1962), and thereby should make indirect support more effective.

As both of our previous studies indicated (Benz et al., 2010), ELWMS effectively induces reflection processes in the postaction phase. If future research manages to effectively support learners during the action phase, and thereby to create a goal/folder-resource structure of high quality, we are confident that a gain in factual knowledge will be achieved. However, besides investigating changes in the independent variable to examine effects on the dependent variable, it might be beneficial to utilize learning outcomes of different complexities, such as in structure or understanding (Azevedo & Cromley, 2004; Azevedo, Cromley et al., 2004; Bannert, 2006).

The results of this study also suggest that, in the preaction and postaction phases, learners profit from supplementing our scaffolds by the prompting of the achievement-enhancing processes. With respect to our previous study (Benz et al., 2010), in which we did not find the invasive and directive goal setting and planning prompt in the preaction phase to be effective, we clarified that this was due to failures in the prerequisite process of self-diagnosing knowledge gaps. Hence, both of our studies provide evidence that learners follow invasive and directive prompts. However, in the action phase, an intensive administration of noninvasive and directive prompting did not enhance metacognitive processes, which again underlines the difficulty of effectively assisting individuals during the ongoing learning process. Future research will have to investigate how and to what extent prompting benefits the employment of metacognitive processes during the ongoing learning process. In addition, it will have to clarify the advantages and disadvantages of prompting of different strengths. One promising approach might be to apply invasive and directive prompting during the action phase. However, since we found in our previous study that learners tended to follow prompts superficially and suppress processes that were not supported to the same extent (Greene & Land, 2000), more external regulation might entail rather negative effects, not only in the cognitive system, but very likely also in the motivational system (Boekaerts, 1999).

Referring to the assessment of the quality of web-based learning, it will be necessary to reflect on the methods and instruments that are applied in current research. It has already been stated that questionnaires like the MSLQ (Pintrich et al., 1993) and the VCQ II (Kuhl & Fuhrmann, 1998), which assess general learning processes, have only limited value for evaluating web-based learning. For the current study, we developed an *offline* self-report pretest that requested learners to indicate if they would carry out the achievement-enhancing processes and SRL processes in a specific web-based learning scenario, and a posttest that assessed whether learners had carried out the identical processes during the implementation of the learning task. It will be the challenge of future research to construct specific questionnaires for evaluating processes that occur during learning on the WWW.

With regard to *online* measures of learning, to be able to collect comparable log data for all conditions, we harmonized the design of our study with our method of collecting data. However, by equipping the control group with a downgraded version of ELWMS, we created a very strong control condition that might have entailed a loss of effect on the dependent variables. It will be one of the central challenges of future research to find a way to generate comparable log data for all groups without establishing conditions that are too much alike. A solution to this problem might be to focus on other *online* measures that depend less on the functions that a particular type of software provides, like eye tracking or thinking-aloud protocols. With respect to the qualitative analyses of our log data, considering the great effort that it took to assign each goal, resource, and page a relevance rating with respect to the achievement test, both of our studies indicated that relying only on quantitative log data is not too much of a trade off. Further, in the current study, to be able to assess the same process by different instruments, we aligned our *offline* self-report pretest and posttest and our method of collecting *online* log data. We strongly recommend that this synchronized multi-method approach be applied in future research.

Acknowledgments

This research was carried out within the Interdisciplinary Research Training Group *Feedback Based Quality Management in eLearning* funded by the German Research Foundation (DFG). The authors would like to thank Julia Friedmann, Jianhua Lambrecht, Karim Hirsch, and Thomas Loesch for their assistance. The research was part of the doctoral dissertation of Bastian F. Benz. An earlier version of this article was presented at the 2009 biannual meeting of the European Association for Research on Learning and Instruction

(EARLI) in Amsterdam, Netherlands and at the 2009 biannual meeting of the Educational Psychology Section of the German Psychological Society (DGPs) in Saarbruecken, Germany.

References

Aleven, V., Stahl, E., Schworm, S., Fischer, F., & Wallace, R. M. (2003). Help seeking and help design in interactive learning environments. *Review of Educational Research, 73(2)*, 277-320. doi:10.3102/00346543073003277

Alexander, P. A. (1997). Mapping the multidimensional nature of domain learning: The interplay of cognitive, motivational and strategic forces. In M. L. Maehr & P. R. Pintrich (Eds.), *Advances in Motivation and Achievement* (Vol. 10, pp. 213-250). Greenwich, CT: JAI Press.

Azevedo, R., & Cromley, J. G. (2004). Does training on self-regulated learning facilitate students' learning with hypermedia? *Journal of Educational Psychology, 96(3)*, 523-535. doi:10.1037/0022-0663.96.3.523

Azevedo, R., Cromley, J. G., & Seibert, D. (2004). Does adaptive scaffolding facilitate students' ability to regulate their learning with hypermedia? *Contemporary Educational Psychology, 29*(3), 344-370. doi:10.1016/j.cedpsych.2003.09.002

Azevedo, R., Guthrie, J. T., & Seibert, D. (2004). The role of self-regulated learning in fostering students‘ conceptual understanding of complex systems with hypermedia. *Journal of Educational Computing Research, 30(1)*, 87-111. doi:10.1016/j.cedpsych.2003.09.002

Bannert, M. (2006). Effects of reflection prompts when learning with hypermedia. *Journal of Educational Computing Research, 35(4)*, 359-378. doi:10.2190/94V6-R58H-3367-G388

Benz, B. F., & Schmitz, B. (2009). *Self-regulated learning and academic success: Do self-regulated learning interventions enhance performance? A meta-analysis.* Manuscript submitted for publication.

Benz, B. F., Scholl, P., Boehnstedt, D., Schmitz, B., Rensing, C., & Steinmetz, R. (2010). *Improving the Quality of E-Learning. Scaffolding Self-Regulated Learning on the World Wide Web.* Manuscript submitted for publication.

Boekaerts, M. (1999). Self-regulated learning: Where we are today. *International Journal of Educational Research*, 31(6), 445-457. doi:10.1016/S0883-0355(99)00014-2

Brown, A., Campione, J. C., & Day, J. D. (1981). Learning to learn: On training students to learn from texts. *Educational Researcher, 10(2)*, 147-21. doi: 10.3102/0013189X010002014

Brush, T., & Saye, J. (2001). The use of embedded scaffolds with hypermedia-supported student-centered learning. *Journal of Educational Multimedia and Hypermedia, 10(4)*, 333-356.

Campbell, D. T; & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56(2)*, 81-105. doi:10.1037/h0046016

Camtasia (Version 3) [Computer software]. Okemos, MI: TechSmith Corporation.

Captivate (Version 3) [Computer software]. San Jose, CA: Adobe.

Clarebout, G., & Elen, J. (2006). Tool use in computer-based learning environments: Towards a research framework. *Computers in Human Behavior, 22(3)*, 389–411. doi:10.1037/h0046016

Dillon, A., & Gabbard, R. (1998). Hypermedia as educational technology: A review of the quantitative research literature on learner comprehension, control, and style. *Review of Educational Research, 68(3)*, 322-349. doi:10.2307/1170600

Elliot, A., & McGregor, H. A. (2001). A 2 x 2 achievement goal framework. *Journal of Personality and Social Psychology, 80(3)*, 501-519. doi:10.1037//0022-3514.80.3.501

Flavell, J. H. (1970). Developmental studies of mediated memory. In H. W. Reese & L. P. Lipsitt (Eds.), *Advances in child development and behaviour* (Vol. 5, pp. 181-211). New York: Academic Press.

Friedrich, H. F., & Mandl, H. (1992). Lern- und Denkstrategien – ein Problemaufriss [Learning and thinking strategies: A sketch of the problem]. In H. Mandl & H. F. Friedrich (Eds.), *Lern- und Denkstrategien. Analyse und Intervention* (pp. 3-54). Goettingen, Germany: Hogrefe.

Greene, B., & Land, S. (2000). A qualitative analysis of scaffolding use in a resource-based learning environment involving the World Wide Web. *Journal of Educational Computing Research, 23(2)*, 151-179. doi:10.2190/1GUB-8UE9-NW80-CQAD

Horz, H., Winter, C., Fries, S. (2009). Differential benefits of situated instructional prompts. *Computers in Human Behavior, 25(4)*, 818-828. doi:10.1016/j.chb.2008.07.001

Jonassen, D. (1996). *Computers as mind tools for schools.* Columbus, OH: Merril.

Jacobson, M., & Archodidou, A. (2000). The design of hypermedia tools for learning: Fostering conceptual change and transfer of complex scientific knowledge. *Journal of the Learning Sciences, 9(2)*, 149-199. doi: 10.1207/s15327809jls0902_2

Jonassen, D., & Reeves, T. (1996). Learning with technology: Using computers as cognitive tools. In D. Jonassen (Ed.), *Handbook of research for educational communications and technology* (pp. 694-719). NY: Macmillan.

Kramarski, B., & Zeichner, O. (2001). Using technology to enhance mathematical reasoning: effects of feedback and self-regulation learning. *Educational Media International, 38(2-3)*, 77-82. doi: 10.1080/09523980110041458

Kuhl, J., & Fuhrmann, A. (1998). Decomposing self-regulation and self-control: The volitional component inventory. In J. Heckhausen & C. S. Dweck (Eds.), *Motivation and self-regulation across the life span* (pp. 15–49). New York: Cambridge University Press.

Land, S. & Zembal-Saul, C. (2003). Scaffolding reflection and articulation of scientific explanations in a data-rich, project-based learning environment: An investigation of progress portfolio. *Educational Technology Research & Development, 51(4)*, 65-84. doi:10.1007/BF02504544

Oliver, K., & Hannafin, M.J. (2000). Student management of web-based hypermedia resources during open-ended problem solving. *Journal of Educational Research, 94(2)*, 75-93. doi: 10.1080/00220670009598746

Palincsar, A.S. (1998). Keeping the metaphor of scaffolding fresh—A response to C. Adison Stone's "The metaphor of scaffolding: Its utility for the field of learning disabilities". *Journal of Learning Disabilities, 31*(4), 370-373. doi: 10.1177/002221949803100406

Pea, R. D. (2004). The social and technological dimensions of scaffolding and related theoretical concepts for learning, education, and human activity. *Journal of the Learning Sciences, 13*(3), 423-451. doi:10.1207/s15327809jls1303_6

Perry, N. E., & Winne, P. H. (2006). Learning from learning kits: gStudy traces of students' self-regulated engagements using software. *Educational Psychology Review, 18(3)*, 211-228. doi:10.1007/s10648-006-9014-3

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In P. R. Pintrich, M. Boekaerts & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 452-502). San Diego, CA: Academic Press.

Pintrich, P. R., Smith, D. A., Garcia, T., & McKeachie, W. (1993). Reliability and predictive validity of the Motivated Strategies for Learning Questionnaire (MSLQ). *Educational and Psychological Measurement, 53(3)*, 801-813. doi:10.1177/0013164493053003024

Puntambekar, S., & Hubscher, R. (2005). Tools for scaffolding students in a complex learning environment: What have we gained and what have we missed? *Educational Psychologist, 40(1)*, 1-12. doi:10.1207/s15326985ep4001_1

Rakes, G. C. (1996). Using the internet as a tool in a resource-based learning environment. *Educational Technology, 36(5)*, 52-56.

Reese, H. W. (1962). Verbal mediation as a function of age level. *Psychological Bulletin*, *59(6)*, 502-509. doi:10.1037/h0040739

Schmitz, B., & Wiese, B.S. (2006). New perspectives for the evaluation of training sessions in self-regulated learning: Time-series analyses of diary data. *Contemporary Educational Psychology, 31(1)*, 64-96. doi:10.1016/j.cedpsych.205.02.002

Schmitz, C. (2009). Limesurvey (Version 1.72) [Computer software]. Hamburg, Germany.

Schwonke, R., Hauser, S., Nuckles, M., & Renkl, A. (2006). Enhancing computer-supported writing of learning protocols by adaptive prompts. *Computers in Human Behavior, 22(1),* 77-92. doi:10.1016/j.chb.2005.01.002

Shapiro, A., & Niederhauser, D. (2004). Learning from hypertext: Research issues and findings. In D. H. Jonassen (Ed.). *Handbook of Research for Education Communications and Technology* (2nd ed., pp. 605-620). Mahwah, NJ: Lawrence Erlbaum Associates.

United Nations. (2008). *The global information society: A statistical view*. Chile: United Nations.

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. doi:10.1007/s11409-007-9017-6

Veenman, M. V. J., Prins, F. J., & Verheij, J. (2003). Learning styles: Self-reports versus thinking-aloud measures. *British Journal of Educational Psychology, 73(3)*, 357-372. doi:10.1348/000709903322275885

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* Cambridge, MA: Harvard University Press.

Watson, D., Clark, L.A. & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology, 54(6)*, 1063-70. doi:10.1037/0022-3514.54.6.1063

Williams, M. (1996). Learner control and instructional technologies. In D. Jonassen (Ed.), *Handbook of research on educational communications and technology* (pp. 957-983. New York: Scholastic.

Winne, P., & Hadwin, A. (2008). The weave of motivation and self-regulated learning. In D. Schunk & B. J. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and applications* (pp. 297–314). New York: Taylor & Francis.

Winne, P. H., & Jamieson-Noel, D. (2002). Exploring students' calibration of self reports about study tactics and achievement. *Contemporary Educational Psychology, 27(4)*, 551-572. doi:10.1016/S0361-476X%2802%2900006-1

Wood, D. J., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry, 17(2)*, 89-100. doi:10.1111/j.1469-7610.1976.tb00381.x

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich & M. Zeidner (Eds.), *Handbook of self-regulation (pp. 13-39).* San Diego, CA: Academic Press.

Zimmerman, B. J., & Schunk, D. J. (Eds.). (2001). *Self-regulated learning and academic achievement: Theoretical perspectives*. New York: Lawrence Erlbaum Associates.

Appendix

Scheme for Establishing the Quantitative and Qualitative Metacognitive Scales and the Scales of the Achievement-Enhancing Processes Based on Categories from Log Data Analyses

		Quantitative and qualitative metacognitive scales				
Name of category	Phase	Goal setting	Planning	Self-monitoring	Process-regulation	Reflection
Goal/folder categories						
Defining goals/folders	Pre	+ qt[a] / r12[b]				
	Act				+ qt / r12	
	Post					
Redefining goals/folders (name, description, tag)	Pre	+ qt / r12				
	Act				+ qt / r12	
	Post					
Defining progress of goals	Pre		+ qt / r12			
	Act			+ qt / r12		
	Post					+ qt / r12
Deleting goals/folders	Pre	+ qt / r4				
	Act				+ qt / r4	
	Post					

Name of category	Phase	Quantitative and qualitative metacognitive scales: Goal setting	Planning	Self-monitoring	Process-regulation	Reflection
Restructuring goals/folders	Pre		+ qt / r12			
	Act				+ qt / r12	
	Post					
Viewing goals/folders	Pre		+ qt / r12			
	Act			+ qt / r12[c]		
	Post					+ qt / r12[g]
Resource categories						
Importing resources	Pre		- qt / r12			
	Act			+ qt / r12[e]		
	Post					- qt / r12
Redefining resources (name, description, tag)	Pre					
	Act				+ qt / r12	
	Post					
Defining relevance of resources	Pre					
	Act			+ qt / r12		
	Post					

		Quantitative and qualitative metacognitive scales				
Name of category	Phase	Goal setting	Planning	Self-monitoring	Process-regulation	Reflection
Deleting resources	Pre					
	Act				+ qt / r4	
	Post					
Restructuring resources	Pre					
	Act				+ qt / r12[f]	
	Post					
Viewing resources	Pre					
	Act			+ qt / r12		
	Post					+ qt / r12[g]
Reopening webpages through resources	Pre					
	Act				+ qt / r12	
	Post					+ qt / r12[g]
Navigation categories						
Viewing knowledge net and overview	Pre		+ qt			
	Act			+ qt		
	Post					+ qt[g]

		Quantitative and qualitative metacognitive scales				
Name of category	Phase	Goal setting	Planning	Self-monitoring	Process-regulation	Reflection
Entering new web page	Pre		- qt / r12			
	Act			+ qt / r12[d]		
	Post					- qt / r12
Entering previously visited web page	Pre					
	Act				+ qt / r12[d]	
	Post					

Note. Pre = preaction phase (< 5 min); Act = action phase (5 – 40 min); Post = postaction phase (> 40 min). qt = quantitative (all actions are considered); r12 = all actions with the relevance of 1 and 2 are considered; r4 = all actions with the relevance 4 are considered; + = value is added; - = value is subtracted. Quantitative scales were established by summing across all quantitative values. Qualitative scales were established by summing the number of all actions with the relevance 1, 2, or 4.

[a]Values summed to establish the scale *goal-oriented approach*. [b]Values summed to establish the scale *defining relevant goals*. [c]Values summed to establish the scale *goal orientation during action*. [d]Values summed to establish the scale *finding relevant pages*. [e]Values summed to establish the scale *importing relevant information*. [f]Values summed and added to the number of goal/folder-resource *fits* to establish the scale *assigning relevant information to relevant goal*. [g]Values summed to establish the scale *learning relevant information*

Part 3: Appendices

Contents

Appendix A: Study 1

Appendix A1

Variables Assessed for Descriptive purposes, Proposed Moderators and Appurtenant Categories

Variable	Categories of variable	Application of variable
Author	-	Descriptive
Year	-	Descriptive
Title	-	Descriptive
Country of origin	-	Descriptive
Publication type	-	Descriptive
Review status	• Peer-reviewed • Non-peer-reviewed	Proposed moderator
Size of sample	-	Descriptive
Relevant data sets	-	Descriptive
Percentage of female participants	-	Descriptive
Age of participants	• 9 - 13years • 14 - 18 years • 19 - 37years	Proposed moderator
Research design	• Experimental • Quasi-experimental	Proposed moderator
Experimental condition	-	Descriptive
SRL layer of intervention	• Metacognitive • Metacognitive, cognitive • Metacognitive, motivational • Metacognitive, cognitive, motivational	Proposed moderator

Variable	Categories of variable	Application of variable
SRL level of intervention	• Micro • Mid • Micro & mid	Proposed moderator
Type of support	• Strategy instruction • Process support • Strategy instruction • & process support	Proposed moderator
Instance of delivery of intervention	• Human (researcher) • Human (teacher) • Computer • Paper • Human & paper	Proposed moderator
Hours of intervention	• 0 - 1 hours • 1.1 - 2 hours • 6 - 9 hours • 11 - 16 hours • 20+ hours	Proposed moderator
Length of intervention	• 1 day • 3 - 6 weeks • 2 - 7 months	Proposed moderator
Domain of learning	• Mathematics • Language • Science • Other	Proposed moderator
Control condition	-	Descriptive

Variable	Categories of variable	Application of variable
Measure of academic achievement	• Grade, undefined • Problem solving • Knowledge multimedia-based • Comprehension multimedia-based • Writing quality	Proposed moderator
Mean of experimental group(s)	-	
Standard deviation of experimental group(s)	-	
Size of experimental group(s)	-	
Mean of control group(s)	-	
Standard deviation of control group(s)	-	
Size of control group(s)	-	

Appendix B: Study 2

Appendix B1

Multiple Regression Analyses

Table B1 - 1

Multiple Regression of the Metacognitive Processes Established on the Basis of Quantitative Video Analyses and Gain in Factual Knowledge

Quantitative variables from video analysis	*B*	*SE B*	β
Constant	1.80	2.71	
Goal setting	-0.10	0.39	-.07
Planning	-0.08	0.49	-.04
Monitoring	0.05	0.10	.16
Process-regulation	0.04	0.09	.11
Reflection	-0.01	0.26	-.01

Note: R^2 = .04.

Table B1 - 2

Multiple Regression of the Metacognitive Processes Established on the Basis of Qualitative Video Analyses and Gain in Factual Knowledge

Qualitative variables from video analysis	*B*	*SE B*	β
Constant	1.34	2.48	
Goal setting	1.74	0.84	.38*
Planning	-1.10	0.79	-.26
Monitoring	0.10	0.14	.14
Process-regulation	-0.19	0.21	-.17
Reflection	-0.47	0.46	-.19

Note: R^2 = .24.

*p < .05.

Table B1 - 3

Multiple Regression of the Metacognitive Processes Established on the Basis of Self-Reports and Gain in Factual Knowledge

Self-reports scales	*B*	*SE B*	β
Constant	4.62	2.87	
Goal setting	1.85	0.98	.35#
Planning	0.17	0.95	.03
Monitoring	0.55	1.09	.08
Process-regulation	-1.24	0.92	-.21
Reflection	-1.32	1.24	-.15

Note: R^2 = .10.

# $p < .10$.

Appendix C: Study 3

Appendix C1

Categories of Log File Analyses for Coding Basic Types of ELWMS Activities

Table C1 - 1

Categories of Goal/Folder Activities.

Goal/folder categories	Parameters: Code	Group	t action	IDg/f	Ioda	Name	Descr	Goalprog	Tag	Maing/f	Path	Prevval	t creatg/f	Ratrel	Quest	Finalstat	Actgoal	Pactgoal	SRL-Poa	Lengtha	t cioda
Creatg/f	X	X	X	X	-	X	X	X	X	-	X	-	-	X	X	X	X	X[a]	X	-	-
Redefg/f	X	X	X	X		X	X	X	X	X	X	-	X	X	X	X	X	X[c]	X	-	-
Redefg/f-n	X	X	X	X	-	X[b]	X	X	X	X	X	X	X	X	X	-	X	X[c]	X	-	-
Redefg/f-d	X	X	X	X	-	X	X[b]	X	X	X	X	X	X	X	X	-	X	X[c]	X	-	-
Redefprog	X	X	X	X	-	X	X	X[b]	X	X	X	X	X	X	X	-	X	X[c]	X	-	-
Redefg/f-t	X	X	X	X	-	X	X	X	X[b]	X	X	X	X	X	X	-	X	X[c]	X	-	-
Viewg/f	X	X	X	X	-	X	X	X	X	X	X	-	X	X	X	-	X	X[c]	X	X	-
Activgoal	X	X	X	X	-	X	X	X	X	X	X	-	X	X	X	-	-	-	X	X	-
Deactivgoal	X	X	X	X	-	X	X	X	X	X	X	-	X	X	X	-	-	-	X	-	-
Delg/f	X	X	X	X	-	X	X	X	X	-	X	-	X	X	X	-	X	X[c]	X	-	-
Delg/f-g/f	X	X	X	X	X	X	X	X	X	-	X	-	X	X	X	-	X	X[d]	X	-	X
Restg/f-wg/f	X	X	X	X	-	X	X	X	X	-	X	-	X	X	X	-	X	X[c]	X	-	-

Goal/folder categories	Parameters																				
	Code	Group	t action	IDg/f	Ioda	Name	Descr	Goalprog	Tag	Maing/f	Path	Prevval	t creatg/f	Ratrel	Quest	Finalstat	Actgoal	Pactgoal	SRL-Poa	Lengtha	t cioda
Restg/f-bg/f	X	X	X	X	-	X	X	X	X	-	X[b]	X	X	X	X	-	X	X[c]	X	-	-
Restg/f-g/f	X	X	X		X	X	X	X	X	-	X[b]	X	X	X	X	-	X	X[d]	X	-	X

Note. Code = code of participant; Group = experimental condition of participant; t action = timestamp of action; IDg/f = ID of goal/folder; Ioda = instance of direct action; Name = name of goal/folder; Descr = description of goal/folder; Goalprog = goal progress; Tag = tag for goal/folder; Maing/f = main goal/folder?; Path = goal/folder- path; Prevval = previous value; t creatg/f = timestamp creation of goal/folder; Ratrel = rated relevance in qualitative analyses; Quest = question in achievement test; Finalstat = final status of goal/folder?; Actgoal = activated goal; Pactgoal = pursuing activated goal?; SRL-Poa = SRL-Phase of action; Lengtha = length of action; t cioda = timestamp creation of instance of direct action; Creatg/f = creating goal/folder; X = respective information is collected for this category; - = respective information is not collected for this category; Redefg/f = redefining goal/folder (includes all redefinitions of name, description, progress and tag); Redefg/f-n = redefining name of goal/folder; Redefg/f-d = redefining description of goal/folder; Redefprog = redefining progress of goal; Redefg/f-t = redefining tag for goal/folder; Viewg/f = viewing goal/folder; Activgoal = activating goal; Deactivgoal = deactivating goal; Delg/f = deleting goal/folder; Delg/f-g/f = deleting subgoal/subfolder by deleting goal/folder; Restg/f-wg/f = restructuring goal/folder within goal/folder; Restg/f-bg/f = restructuring goal/folder between goals/folders; Restg/f-g/f = restructuring subgoal/subfolder by restructuring goal/folder.

[a]Comparison of activated goal with the goal that is one level higher. [b]Value adapted by action, previous value is kept in proper column.

[c]Comparison of activated goal with last goal in path. [d]Comparison of activated goal with goal that the action is performed on.

Table C1 - 2

Categories of Resource Activities.

Resource categories	Parameters																										
	Code	Group	t action	TabID	IDres	Ioda	Name	Descr	Rel	Tag	Path	Website	Prevval	t import	n char	Less1150	Ratrel	Quest	Fit	Finalstat	Actgoal	Pactgoal	SRL-Poa	Lengtha	t cioda	t cs/gf	SRL-Pcs/gf	Ib/acsg/f
Impres	X	X	X	-	X	-	X	X	X	X	X	X	-	-	X	X	X	X	X	X	X	X[a]	X	-	-	X	X	X
Redefres	X	X	X	-	X	-	X	X	X	X	X	X	-	X	X	X	X	X	X	X	X	X[a]	X	-	-	X	X	X
Redefres-n	X	X	X	-	X	-	X[b]	X	X	X	X	X	X	X	X	X	X	X	X	-	X	X[a]	X	-	-	X	X	X
Redefres-d	X	X	X	-	X	-	X	X[b]	X	X	X	X	X	X	X	X	X	X	X	-	X	X[a]	X	-	-	X	X	X
Redefres-r	X	X	X	-	X	-	X	X	X[b]	X	X	X	X	X	X	X	X	X	X	-	X	X[a]	X	-	-	X	X	X
Redefres-t	X	X	X	-	X	-	X	X	X	X[b]	X	X	X	X	X	X	X	X	X	-	X	X[a]	X	-	-	X	X	X
Viewres	X	X	X	-	X	-	X	X	X	X	X	X	-	X	X	X	X	X	X	-	X	X[a]	X	X	-	X	X	X
Delres	X	X	X	-	X	-	X	X	X	X	X	X	-	X	X	X	X	X	X	-	X	X[a]	X	-	-	X	X	X
Delres-g/f	X	X	X	-	X	X	X	X	X	X	X	X	-	X	X	X	X	X	-	-	X	X[c]	X	-	X	X	X	X
Restres-wg/f	X	X	X	-	X	-	X	X	X	X	X	X	-	X	X	X	X		X	-	X	X[a]	X	-	-	X	X	X
Restres-bg/f	X	X	X	-	X	-	X	X	X	X	X[b]	X	X	X	X	X	X	X	X	-	X	X[a]	X	-	-	X[d]	X[d]	X[d]
Restres-g/f	X	X	X	-	X	X	X	X	X	X	X[b]	X	-	X	X	X	X	X	-	-	X	X[c]	X	-	X	X	X	X
Reopres	X	X	X	X	X	-	X	X	X	X	X	X	-	X	X	X	X	X	X	-	X	X[a]	X	-	-	X	X	X

Note. Code = code of participant; Group = experimental condition of participant; t action = timestamp of action; TabID = tab ID in Firefox; IDres = ID of resource; Ioda = instance of direct action; Name = name of resource; Descr = description of resource; Rel = relevance of resource; Tag = tag for resource; Path = goal/folder-resource path; Website = website contained from; Prevval = previous value; t import = timestamp import; nchar = number of characters; Less1150 = less than 1150 characters?; Ratrel = rated relevance in qualitative analyses; Quest = question in achievement test; Fit = goal/folder-resource fit?; Finalstat = final status of resource?; Actgoal = activated goal; Pactgoal = pursuing activated goal?; SRL-Poa = SRL-Phase of action; Lengtha = length of action; t cioda = timestamp creation of instance of direct action; t cs/gf = timestamp of creation of superior goal/folder; SRL-Pcs/gf = SRL-Phase creation of superior goal/folder; Ib/acsg/f = import before/ after creation of superior goal/folder?; Impres = importing resource; X = respective information is collected for this category; - = respective information is not collected for this category; Redefres = redefining resource (includes all redefinitions of name, description, relevance and tag); Redefres-n = redefining name of resource; Redefres-d = redefining description of resource; Redefres-r = redefining relevance of resource; Redefres-t = redefining tag for resource; Viewres = viewing resource; Delres = deleting resource; Delres-g/f = deleting resource by deleting goal/folder; Restres-wg/f = restructuring resource within goal/folder; Restres-bg/f = restructuring resource between goals/folders; Restres-g/f = restructuring resource by restructuring goal/folder; Reopres = reopening webpage through resource.
[a]Comparison of activated goal with last goal in path. [b]Value adapted by action, previous value is kept in proper column. [c]Comparison of activated goal with goal that the action is performed on. [d]New superior goal/folder is used for comparison.

Table C1 - 3

Categories of Navigation Activities.

Navigation categories	Parameters: Code	Group	t action	TabID	Website	Searcht	Ratrel	Quest	Actgoal	Pactgoal	SRL-Poa	Lengtha
Viewkn	X	X	X	-	-	-	-	-	X	X[a]	X	X
Viewo	X	X	X	X	-	-	-	-	X	X[a]	X	-
Entnp	X	X	X	X	X	-	X	X	X	X[a]	X	-
Entpvp	X	X	X	X	X	-	X	X	X	X[a]	X	-
Searchwik	X	X	X	-	-	X	-	-	X	X[a]	X	-
Beginpre	X	X	X	-	-	-	-	-	X	-	X	-
Begina	X	X	X	-	-	-	-	-	X	-	X	-
Beginpost	X	X	X	-	-	-	-	-	X	-	X	-
Endpost	X	X	X	-	-	-	-	-	X	-	X	-

Note. Code = code of participant; Group = experimental condition of participant; t action = timestamp of action; TabID = Tab ID in Firefox; Website = website visited; Searcht = term used for Wikipedia search; Ratrel = rated relevance in qualitative analyses; Quest = question in achievement test; Actgoal = activated goal; Pactgoal = pursuing activated goal; SRL-Poa = SRL-Phase of action; Lengtha = length of action; X = respective information is collected for this category; - = respective information is not collected for this category; Viewkn = viewing knowledge net; Viewo = viewing overview; Entnp = entering new web page; Entpvp = entering previously visited web page; Searchwik = searching in Wikipedia; Beginpre = begin preaction phase; begina = begin action phase; Beginpost = begin postaction phase; Endpost = end postaction phase.

[a]Comparison of activated goal with last goal in path.

Appendix C2

Results of the Web-Based Training

When contrasting groups in terms of their achievement on the training (EG-Tool: M = 87.62, SD = 16.34, n = 35; EG-Prompt: M = 87.65, SD = 19.84, n = 36; CG-Folder: M = 88.54, SD = 17.28, n = 32; $F(2, 102) = 0.03$, $p = .972$, $\eta^2 = .00$) as well as their time spent for practicing with ELWMS (EG-Tool: M = 48.08, SD = 7.84, n = 35; EG-Prompt: M = 46.92, SD = 3.12, n = 36; CG-Folder: M = 46.03, SD = 0.71, n = 32; $F(2, 102) = 1.45$, $p = .239$, η^2 = .03), we did not find proof for differences in the capability of handling ELWMS.

Appendix C3

Results of the Achievement Pretest

When contrasting the three conditions in terms of performance on the achievement pretest, we did not find proof for differences in previous knowledge on the topic of *Classical Antiquity* (EG-Tool: *M* = 0.37, *SD* = 0.81, *n* = 35; EG-Prompt: *M* = 0.65, *SD* = 1.42, *n* = 37; CG-Folder: *M* = 0.35, *SD* = 0.69, *n* = 34; *F*(2, 105) = 0.92, *p* = .400, η^2 = .02). We also did not find evidence for differences in time that participants of the three groups had spent working on the achievement pretest (EG-Tool: *M* = 11.62, *SD* = 3.74, *n* = 35; EG-Prompt: *M* = 11.50, *SD* = 3.57, *n* = 35; CG-Folder: *M* = 11.55, *SD* = 2.79, *n* = 34; *F*(2, 103) = 0.11, *p* = .989, η^2 = .00).

We also retrospectively analyzed if our feedback mechanism had met our requirements. We did not find significant differences between the frequencies that questions were statistically expected to appear on the feedback and their actual appearance. In addition, we evaluated if participants of different conditions had received a feedback of different complexity. Regarding the accessibility of an answer on Wikipedia, we determined the difficulty of a question by adding the number of resources that had been found by participants in the current study for each question. We applied the relevance ratings from the qualitative analyses of our log files to account for the quality of the resources. Instances with a relevance of 1 were weighted by the factor 3, with a relevance of 2 by the factor 2, and with a relevance of 3 by the factor 1. On this basis, we generated a value that represented the difficulty of the feedback for each participant. When contrasting groups, we did not find proof for differences in the difficulty of the feedback (EG-Tool: *M* = 1074.83, *SD* = 77.96, *n* = 35; EG-Prompt: *M* = 1073.89, *SD* = 99.73, *n* = 37; CG-Folder: *M* = 1089.53, *SD* = 79.33, *n* = 34; *F*(2, 105) = 0.36, *p* = .701, η^2 = .01).

Appendix C4

Results of the AGQ[1]

When contrasting groups, we could not find significant differences between participants of the three conditions in terms of their achievement goal orientation (Performance-approach: EG-Tool: $M = 2.00$, $SD = 0.72$, $n = 35$; EG-Prompt: $M = 2.13$, $SD = 0.76$, $n = 37$; CG-Folder: $M = 2.03$, $SD = 0.81$, $n = 34$; $F(2, 105) = 0.30$, $p = .740$, $\eta^2 = .01$; Performance-avoidance: EG-Tool: $M = 2.52$, $SD = 0.62$, $n = 35$; EG-Prompt: $M = 2.62$, $SD = 0.62$, $n = 37$; CG-Folder: $M = 2.83$, $SD = 0.59$, $n = 34$; $F(2, 105) = 2.27$, $p = .108$, $\eta^2 = .04$; Mastery-approach: EG-Tool: $M = 2.87$, $SD = 0.56$, $n = 35$; EG-Prompt: $M = 2.66$, $SD = 0.43$, $n = 37$; CG-Folder: $M = 2.74$, $SD = 0.71$, $n = 34$; $F(2, 105) = 1.21$, $p = .303$, $\eta^2 = .03$; Mastery-avoidance: EG-Tool: $M = 2.55$, $SD = 0.78$, $n = 35$; EG-Prompt: $M = 2.29$, $SD = 0.64$, $n = 37$; CG-Folder: $M = 2.33$, $SD = 0.70$, $n = 34$; $F(2, 105) = 1.42$, $p = .247$, $\eta^2 = .02$).

[1] Achievement Goal Questionnaire (Elliot & McGregor, 2001)

Appendix C5

Differences Between Groups in the Conduction of the Achievement Enhancing Processes Based on Self-Reports

Self-report scales	EG-Tl (*n* = 35) *M* (*SD*)	EG-Ppt (*n* = 36) *M* (*SD*)	CG-Fo (*n* = 34) *M* (*SD*)	*F*
Goa	3.34 (0.55)	3.42 (0.64)	3.32 (0.49)	$F(2, 102) = 0.27, p = .766, \eta^2 = .01$
Drg	3.29 (0.61)	3.28 (0.71)	3.31 (0.69)	$F(2, 102) = 0.02, p = .980, \eta^2 < .01$
Goda	3.10 (0.48)	3.09 (0.53)	3.11 (0.52)	$F(2, 102) = 0.01, p = .994, \eta^2 < .01$
Frwp	3.34 (0.59)	3.35 (0.67)	3.36 (0.61)	$F(2, 102) = 0.01, p = .993, \eta^2 < .01$
Iri	3.32 (0.64)	3.40 (0.53)	3.24 (0.65)	$F(2, 102) = 0.60, p = .549, \eta^2 = .01$
Aritrg	2.91 (0.77)	2.94 (0.69)	2.93 (0.83)	$F(2, 102) = 0.02, p = .984, \eta^2 < .01$
Lri	3.20 (0.69)	3.37 (0.57)	3.24 (0.71)	$F(2, 102) = 0.62, p = .539, \eta^2 = .01$

Note. EG-Tl = experimental group working with the standard version of ELWMS; EG-Ppt = experimental group working with extended version of ELWMS; CG-Fo = control group working with control version of ELWMS; Goa = goal oriented approach; Drg = defining relevant goals; Goda = goal orientation during action; Frwp = finding relevant web pages; Iri = importing relevant information; Aritrg = assigning relevant information to relevant goal; Lri = Learning relevant information.

[a]one-tailed.

Appendix C6

Differences Between Groups in the Conduction of SRL Processes Based on Self-Reports

	Conditions			
	EG-Tl	EG-Ppt	CG-Fo	
Self-report	(*n* = 35)	(*n* = 36)	(*n* = 34)	
scales	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*F*
SRL	2.95 (0.41)	2.90 (0.44)	2.88 (0.46)	$F(2, 102) = 0.23, p = .799, \eta^2 < .01$
Metacog	2.94 (0.43)	2.88 (0.49)	2.85 (0.47)	$F(2, 102) = 0.38, p = .683, \eta^2 = .01$
Goals	3.23 (0.48)	3.29 (0.58)	3.21 (0,53)	$F(2, 102) = 0.20, p = .822, \eta^2 < .01$
Plan	2.52 (0.62)	2.27 (0.60)	2.26 (0.59)	$F(2, 102) = 2.01, p = .139, \eta^2 = .04$
Selfm	2.87 (0.47)	2.76 (0.51)	2.71 (0.57)	$F(2, 102) = 0.89, p = .415, \eta^2 = .02$
Procreg	2.97 (0.43)	2.92 (0.56)	2.96 (0.48)	$F(2, 102) = 0.08, p = .925, \eta^2 < .01$
Refl	3.10 (0.51)	3.11 (0.56)	3.03 (0.63)	$F(2, 102) = 0.22, p = .807, \eta^2 < .01$
Mot	3.14 (0.44)	3.09 (0.41)	3.17 (0.64)	$F(2, 102) = 0.20, p = .818, \eta^2 < .01$
Subexp	2.75 (0.78)	2.86 (0.71)	3.00 (0.75)	$F(2, 102) = 0.95, p = .390, \eta^2 = .02$
PANASpos	2.81 (0.40)	2.81 (0.48)	2.82 (0.55)	$F(2, 102) = 0.01, p = .992, \eta^2 < .01$
PANASneg	1.36 (0.39)	1.38 (0.43)	1.29 (0.43)	$F(2, 102) = 0.49, p = .614, \eta^2 = .01$
Msmot	4.35 (0.91)	4.04 (0.97)	4.47 (0.92)	$F(2, 102) = 2.04, p = .135, \eta^2 = .04$
Msse	4.18 (0.72)	4.19 (0.66)	4.16 (0.84)	$F(2, 102) = 0.01, p = .986, \eta^2 < .01$

Note. EG-Tl = experimental group working with the standard version of ELWMS; EG-Ppt = experimental group working with extended version of ELWMS; CG-Fo = control group working with control version of ELWMS; SRL = self-regulated learning; Metacog = metacognition; Goals = goal setting; Plan = planning; Selfm = self-monitoring; Procreg = process-regulation; Refl = reflection; Mot = motivation; Subexp = subjective experience; PANASpos = positive Items of the PANAS (Watson, Clark, & Tellegen, 1988); PANASneg = negative Items of the PANAS; Msmot = mean state motivation; Msse = Mean state self-efficacy.

[a]one-tailed.

Appendix C7

Further Results Regarding the Quality of the Learning Outcome

In further analyses, we split the questions of the achievement test into half by determining how often a question on the posttest had been solved correctly and with a certainty of more than 50%. To examine whether groups had established a goal/folder-resource-structure of different quality for easy and difficult questions, we created two different values for the quality of the structure: one that accounted for the goals and resources that were relevant for the easy questions and one that accounted for instances that were relevant for the difficult questions. We used the same formula like for calculating the general value of the structure, but did not add the number of goal/folder-resource fits that were part of the final structure. We found evidence that experimental groups had established a more sophisticated basis for questions that were more likely to be answered correctly on the posttest than the control group. This was not the case for the difficult questions. The same pattern was perceived when comparing experimental groups.

Table C7

Differences Between Groups in the Quality of the Final Goal/Folder-Resource Structure Based on Log Data

	Conditions				Planned comparisons	
	EG-Tl (n = 35)	EG-Ppt (n = 36)	CG-Fo (n = 34)			
Scales from log file analyses	M (SD)	M (SD)	M (SD)	F	EG-Tool & EG-Ppt vs. CG-Fo[a]	EG-Tool vs. EG-Ppt[a]
Structure – easy questions	17.54 (9.68)	22.06 (8.55)	16.00 (10.38)	$F(2, 102) = 3.83$, $p = .025$, $\eta^2 = .07$	$t(102)$ = ,1.91 $p = .030$, $\eta^2 = .03$	$t(102) = 1.99$, $p = .025$, $\eta^2 = .04$
Structure – difficult questions	12.63 (6.40)	13.47 (7.06)	12.68 (7.49)	$F(2, 102) = 0.16$, $p = .850$, $\eta^2 < .01$	$t(102) = 0.26$, $p = .399$, $\eta^2 < .01$	$t(102) = 0.51$, $p = .306$, $\eta^2 < .01$.

Note. EG-Tl = experimental group working with the standard version of ELWMS; EG-Ppt = experimental group working with extended version of ELWMS; CG-Fo = control group working with control version of ELWMS.

[a]one-tailed.

Appendix C8

Multiple Regression Analyses

We further investigated whether the conduction of the achievement enhancing processes as well as the metacognitive processes during the 45 minute learning period, enhanced participants gain of factual knowledge. We conducted five regression analyses, with the achievement enhancing processes measured by self-reports and log files, as well as the metacognitive processes established on the basis of self-reports and quantitative and qualitative log files as independent variables, and the difference between the questions that participants had answered correctly and with a certainty of more than 50% from the achievement pre- to the posttest as dependent variable. For all five regressions, we used the forced entry method.

Table C8 - 1

Multiple Regression of the Crucial Processes Established on the Basis of Log Data and Gain in Factual Knowledge

Scales from log file analyses	*B*	*SE B*	β
Constant	3.07	0.66	
Goal oriented approach	-.02	0.08	-.02
Defining relevant goals	.15	0.10	.20
Goal orientation during action	-0.18	0.06	-.25**
Finding relevant pages	-.011	0.02	-.04
Importing relevant information	.59	0.11	.60**
Assigning relevant information to relevant goal	-.04	0.07	-.07
Learning relevant information	.05	0.16	.024

Note: $R^2 = .39$.

**$p < .01$.

Table C8 - 2

Multiple Regression of the Crucial Processes Established on the Basis of Self-Reports and Gain in Factual Knowledge

Self-report scales	*B*	*SE B*	β
Constant	-2.79	1.76	
Goal oriented approach	-0.20	0.70	-.04
Defining relevant goals	0.86	0.50	.21#
Goal orientation during action	0.56	0.71	.10
Finding relevant pages	0.69	0.61	.16
Importing relevant information	0.45	0.65	.10
Assigning relevant information to relevant goal	-0.68	0.383	-.19#
Learning relevant information	0.72	0.45	.17

Note: $R^2 = .23$.

#$p < .10$.

Table C8 - 3

Multiple Regression of the Metacognitive Processes Established on the Basis of Quantitative Log Data and Gain in Factual Knowledge

Scales from log file analyses - quantitative	*B*	*SE B*	β
Constant	2.73	1.00	
Goal setting	0.22	0.06	.42**
Planning	-0.06	0.03	-.24#
Self-monitoring	0.03	0.02	.15
Process-regulation	-0.02	0.02	-.11
Reflection	-0.08	0.07	-.12

Note: $R^2 = .15$.

#$p < .10$, **$p < .01$.

Table C8 - 4

Multiple Regression of the Metacognitive Processes Established on the Basis of Qualitative Log Data and Gain in Factual Knowledge

Scales from log file analyses - qualitative	*B*	*SE B*	β
Constant	2.62	0.82	
Goal setting	0.23	0.08	.37**
Planning	-0.08	0.04	-.25*
Self-monitoring	0.06	0.03	.23*
Process-regulation	-0.01	0.02	-.02
Reflection	0.06	0.13	.05

Note: R^2 = .18.

*p < .05, **p < .01.

Table C8 - 5

Multiple Regression of the Metacognitive Processes Established on the Basis of Self-Reports and Gain in Factual Knowledge

Self-report scales	*B*	*SE B*	β
Constant	-1.67	1.77	
Goal setting	1.90	0.78	.36*
Planning	0.17	0.62	.04
Self-monitoring	-1.09	1.07	-.21
Process-regulation	0.44	0.77	.08
Reflection	0.70	0.78	.14

Note: R^2 = .16.

*p < .05.

Eidesstattliche Erklärung

Ich versichere hiermit, dass ich die vorliegende Arbeit selbstständig verfasst, keine anderen, als die angegebenen Hilfsmittel verwendet und die Stellen, die anderen Werken im Wortlaut oder dem Sinne nach entnommen sind, mit Quellenangaben kenntlich gemacht habe. Dies gilt auch für Zeichnungen, Skizzen, Ton- und Bildträger sowie bildliche Darstellungen.

Die Arbeit wurde bisher keiner anderen Prüfungsbehörde vorgelegt.

Darmstadt, den 25.03.2010

…………………………………………

Dipl.-Psych. Bastian Frithjof Benz

Curriculum Vitae

Personal Information

Last name	Benz
Prenames	Bastian Frithjof
Date of birth	21.11.1979
Place of birth	Frankfurt, Germany
Nationality	German

Academic Background

July 1996 – June 1997	Exchange year in Milwaukee, USA; Senior student at Milwaukee Trade and Technical High School
August 1997 – June 2000	High School Michelstadt, Germany; Degree: German High School Diploma (Abitur)
October 2000 – September 2006	Study of Psychology at Technische Universität Darmstadt, Germany; Majors: Educational Psychology, Industrial and Organizational Psychology; Minor: Basics of Business Administration; Degree: Diplom-Psychologe
April 2003 – February 2006	Formation “Train the Trainer“ at Technische Universität Darmstadt, Germany
March 2004 – March 2005	Academic exchange year in Buenos Aires, Argentina; Student at Universidad Católica Argentina for one semester; Intern at the Argentinian subsidiary of Henkel KGaA for six months
November 2005 – April 2006	Student assistant at the department of Educational Psychology at Technische Universität Darmstadt, Germany
October 2006 – June 2010	PhD student at the Department of Educational Psychology at Technische Universität Darmstadt, Germany; Scholarship holder in the "Interdisciplinary Research Training Group on Feedback Based Quality Management in eLearning" funded by the German Research Foundation (DFG); Spokesperson of the PhD-students in the training group

Teaching Experience

Summer term 2007	Course “Psychological Aspects of eLearning”
Winter term 2007/2008	Course “Quality improvement in eLearning through Scaffolding”

Publications

2010 Benz, B. F., Scholl, P., Boehnstedt, D., Schmitz, B., Rensing, C., & Steinmetz, R. (2010). *Improving the Quality of E-Learning. Scaffolding Self-Regulated Learning on the World Wide Web.* Manuscript submitted for publication.

Benz, B. F., Scholl, P., Boehnstedt, D., Schmitz, B., Rensing, C., & Steinmetz, R. (2010). *Improving the Quality of Learning on the World Wide Web by Scaffolding Self-Regulated Learning.* Manuscript submitted for publication.

Benz, B. F. (2010). Improving the quality of learning on the World Wide Web: Seamless and disruptive metacognitive support. In M. Muehlhaeuser, W. Sesink, A. Kaminski, & S. Steimle (Eds.), *Interdisciplinary approaches to technology-enhanced learning.* Manuscript in preparation.

2009 Benz, B. F. & Schmitz, B. (2009). *Self-regulated learning and academic success: Do self-regulated learning interventions enhance performance? A meta-analysis.* Manuscript submitted for publication.

Benz, B. F. (in press). Nonreaktive Methoden: Vermeidung reaktiver Effekte in der psychologischen Forschung [Preventing reactive effects in psychological research]. In H. Holling & B. Schmitz (Eds.), *Handwoerterbuch der Psychologie: Methoden und Evaluation*. Germany: Hogrefe.

Scholl, P., Benz, B. F., Boehnstedt, D., Rensing, R., Schmitz, B., & Steinmetz, R. (2009). Implementation and evaluation of a tool for setting goals in self-regulated internet research. In U. Cress, V. Dimitrova, & M. Specht (Eds.), *Learning in the synergy of multiple disciplines.* Germany: Springer.

2008 Boehnstedt, D., Scholl, P., Benz, B. F., Rensing, C. Steinmetz, R., & Schmitz, B. (2008). Einsatz persoenlicher Wissensnetze im Ressourcen-basierten Lernen [Implementation of personal knowledge nets in resource-based learning]. In S. Seehusen, U. Lucke & S. Fischer (Eds.), *DeLFI 2008: 6. e-Learning Fachtagung Informatik (pp.113-124)*. Germany: Köllen.

Scholl, P., Benz, B. F., Boehnstedt, D., Rensing, C. Steinmetz, R., & Schmitz, B. (2008). Einsatz und Evaluation eines Zielmanagement-Werkzeugs bei der selbstregulierten Internet-Recherche [Self-regulated internet search: Implementation and evaluation of a goal management tool]. In S. Seehusen, U. Lucke & S. Fischer (Eds.), *DeLFI 2008: 6. e-Learning Fachtagung Informatik (pp.125-136)*. Germany: Köllen.

2007 Benz, B. F., Polushkina, S., Schmitz, B., & Bruder, R. (2007). Developing learning software for the self-regulated learning of mathematics. In M. B. Nunes. & M. McPherson (Eds.), *IADIS International Conference e-Learning (pp.200-204)*. Portugal: IADIS Press.

Scholl, P., Benz, B. F., Mann, D., Rensing, C., Steinmetz, R., & Schmitz, B. (2007). Scaffolding von selbstreguliertem Lernen in einer Rechercheumgebung fuer internetbasierte Ressourcen [Scaffolding self-regulated learning in a search environment for internet based resources]. In Roessling, G. & Rensing, C. (Eds.), *Proceedings der Pre-Conference Workshops der 5. e-Learning Fachtagung Informatik - DeLFI 2007 (43-50)*. Germany: Logos.

Symposia, Presentations, Posters

2010 Benz, B. F. & Schmitz, B. (2010, May). *Methods for analyzing metacognitive processes.* Paper presented at the biennial meeting of the European Association for Research on Learning and Instruction (EARLI) Metacognition SIG, Muenster, Germany.

2009 Hollender, N. & Benz, B. F. (2009, September). *Foerderung relevanter Lernprozesse im E-Learning* [Fostering relevant learning processes in elearning]. Symposium conducted at the 12th biennial meeting of the Educational Psychology Section of the German Psychological Society (DGPs), Saarbruecken, Germany.

Benz, B. F. & Schmitz, B. (2009, September). *Unterstuetzung des ressourcenbasierten Lernens im World Wide Web* [Supporting resource-based learning on the World Wide Web]. Paper presented at the 12th biennial meeting of the Educational Psychology Section of the German Psychological Society (DGPs), Saarbruecken, Germany.

Benz, B. & Schmitz, B. (2009, September). Scaffolding self-regulated learning on the World Wide Web. In D. Albert (Chair), *Sustainable life-long learning and digital media*. Symposium conducted at the 3rd meeting on the Challenge of Demographic Change of the Japanese-German Center Berlin (JDZB), Berlin, Germany.

Benz, B. F. (2009, August). *Supporting self-regulated learning and metacognition in computer-based environments*. Symposium conducted at the biennial meeting of the European Association for Research on Learning and Instruction, Amsterdam, The Netherlands.

Benz, B. & Schmitz, B. (2009, August). *Scaffolding self-regulated learning on the World Wide Web*. Paper presented at the biennial meeting of the European Association for Research on Learning and Instruction, Amsterdam, The Netherlands.

2008 Benz, B. F. & Schmitz, B. (2008, July). *Self-regulated learning and academic achievement: A meta-analysis.* Paper presented at the 29th International Congress of Psychology (ICP), Berlin, Germany.

Benz, B. F. & Schmitz, B. (2008, July). *Fostering self-regulated web search.* Poster session presented at the 29th International Congress of Psychology (ICP), Berlin, Germany.

Benz, B. F. & Schmitz, B. (2008, May). *Fostering self-regulated web search: Evaluating a goal management tool.* Paper presented at the biennial meeting of the European Association for Research on Learning and Instruction (EARLI) Metacognition SIG, Ioannina, Greece.

2007 Benz, B. F., Polushkina, S., Schmitz, B. & Bruder, R. (2007, July). *Developing learning software for the self-regulated learning of mathematics.* Paper presented at the International Association for Development of the Information Society (IADIS) International Conference e-Learning, Lisbon, Portugal.

Benz, B. F. & Schmitz, B. (2007, September). *Durch Selbstregulation zum Erfolg. Eine Metaanalyse ueber den Zusammenhang zwischen Selbstregulationsinterventionen und akademischer Leistung* [Self-regulated learning and academic success: A meta-analysis on the effectiveness of SRL interventions]. Paper presented at the 11th biennial meeting of the Educational Psychology Section of the German Psychological Society (DGPs), Berlin, Germany.

Benz, B. F. & Schmitz, B. (2007, September). *Qualitaetsverbesserung im E-Learning durch Scaffolding des selbstregulierten Lernens* [Quality improvement in elearning by scaffolding self-regulated learning], Poster session presented at the 11th biennial meeting of the Educational Psychology Section of the German Psychological Society (DGPs), Berlin, Germany.

Bastian F. Benz,
Darmstadt, 25.03.2010